OPEN ACCESS Freely available online

PLOS ONE

Most Human Proteins Made in Both Nucleus and Cytoplasm Turn Over within Minutes

Sabyasachi Baboo[1], Bhaskar Bhushan[2], Haibo Jiang[3], Chris R. M. Grovenor[3], Philippe Pierre[4,5,6], Benjamin G. Davis[2], Peter R. Cook[1]*

1 Sir William Dunn School of Pathology, University of Oxford, Oxford, United Kingdom, **2** Chemistry Research Laboratory, Department of Chemistry, University of Oxford, Oxford, United Kingdom, **3** Department of Materials, University of Oxford, Oxford, United Kingdom, **4** Centre d'Immunologie de Marseille-Luminy, Aix-Marseille Université, Marseille, France, **5** Institut National de la Santé et de la Recherche Médicale, U1104, Marseille, France, **6** Centre National de la Recherche Scientifique, Unités Mixtes de Recherche 7280, Marseille, France

Abstract

In bacteria, protein synthesis can be coupled to transcription, but in eukaryotes it is believed to occur solely in the cytoplasm. Using pulses as short as 5 s, we find that three analogues – L-azidohomoalanine, puromycin (detected after attaching fluors using 'click' chemistry or immuno-labeling), and amino acids tagged with 'heavy' ^{15}N and ^{13}C (detected using secondary ion mass spectrometry) – are incorporated into the nucleus and cytoplasm in a process sensitive to translational inhibitors. The nuclear incorporation represents a significant fraction of the total, and labels in both compartments have half-lives of less than a minute; results are consistent with most newly-made peptides being destroyed soon after they are made. As nascent RNA bearing a premature termination codon (detected by fluorescence *in situ* hybridization) is also eliminated by a mechanism sensitive to a translational inhibitor, the nuclear turnover of peptides is probably a by-product of proof-reading the RNA for stop codons (a process known as nonsense-mediated decay). We speculate that the apparently-wasteful turnover of this previously-hidden ('dark-matter') world of peptide is involved in regulating protein production.

Citation: Baboo S, Bhushan B, Jiang H, Grovenor CRM, Pierre P, et al. (2014) Most Human Proteins Made in Both Nucleus and Cytoplasm Turn Over within Minutes. PLoS ONE 9(6): e99346. doi:10.1371/journal.pone.0099346

Editor: Thomas Preiss, The John Curtin School of Medical Research, Australia

Received January 15, 2014; **Accepted** May 13, 2014; **Published** June 9, 2014

Funding: SB was supported by the Felix Scholarship Trust of Oxford University and The Sir William Dunn School of Pathology, BB is supported by a Rhodes Scholarship, and HJ by The China Scholarship Council. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: peter.cook@path.ox.ac.uk

Introduction

In prokaryotes, protein synthesis can be coupled to transcription [1]. In eukaryotes, it was debated whether some translation might also be coupled [2], but the discovery of introns seemed to provide the reason why eukaryotes should be different; if nuclear ribosomes translated introns with their many termination codons, too many truncated peptides would be produced, and some of these would be toxic. Clearly, restricting intron-containing RNA to nuclei, and translation to the cytoplasm, prevents such lethal consequences. The debate then fizzled out, but was reignited by the discovery that some nonsense-mediated mRNA decay (NMD) occurs in nuclei [3]. This process involves scanning mRNAs for inappropriately-placed ('premature') termination codons (PTCs), and – if found – destruction of the faulty message. As a translating ribosome is the only known mechanism for detecting a termination codon, this places an active ribosome in the nucleus. Despite subsequent reports pointing to some nuclear translation [4–8], the consensus remains that proteins are made only in the cytoplasm [9]; most have half lives of many hours [10–12], although some turn-over within minutes [13–14].

Using pulses as short as 5 s, here we show that three analogues – L-azidohomoalanine, puromycin (detected after attaching fluors using 'click' chemistry or immuno-labeling), and amino acids tagged with 'heavy' ^{15}N and ^{13}C (detected using secondary ion mass spectrometry) – are incorporated into both nucleus and cytoplasm in a process sensitive to translational inhibitors. With all three approaches, substantial signal is seen in both compartments. However, our extraordinary finding is that most signal in both nucleus and cytoplasm disappears within minutes. As these structurally-different analogues (detected in different ways) give similar results, it seems that most newly-made peptide – like newly-made RNA [15] – is destroyed almost as soon as it is made.

We then examined what use – if any – the cell might make of this apparently-wasteful turnover. In the case of the nuclear turnover, we tested whether it was involved in NMD. We find that nascent RNA bearing a PTC is eliminated by a mechanism sensitive to a translational inhibitor; this points to an active ribosome proof-reading the nascent RNA prior to its destruction by NMD. [Here we apply the term 'nascent' to RNA (and peptide) still associated with the polymerase (or ribosome).] Results are consistent with considerable amounts of translation occurring in nuclei, where a 'pioneer round of translation' proofreads the nascent transcript. However, we can only speculate on why so many nascent peptides made in the cytoplasm are degraded.

Results

Aha incorporation

Aha is an analogue of methionine (Met), an essential amino acid, and it is incorporated into proteins both at the N terminus and internally; as it possesses a reactive azide group, 'click' chemistry is widely used to conjugate alkyne-containing fluors on to Aha-containing peptides, before localization of those peptides [16–17]. All reports indicate that Aha-bearing peptides behave much like their Met-containing counterparts; for example, zebrafish larvae develop normally when grown in Met-free and Aha-containing medium for two days [18]. [However, Aha is not incorporated as efficiently as Met by the bacterial Met-tRNA synthetase [19].] When HeLa cells are starved of Met to deplete endogenous pools, grown in Aha for 2 min, fixed, and Alexa555 'clicked' on to incorporated Aha, fluorescence is seen in both cytoplasm and nucleus (**Fig. 1Aii**). Pre-incubation with the translational inhibitor, anisomycin [20], reduces Aha incorporation (**Fig. 1Aiii**; see also the legend to **Fig. S1**). [The anisomycin concentration and pre-incubation time applied here are routinely used to inhibit translation to the levels we see [21]; shorter pre-incubation times of 15–30 min will be used in critical experiments, and other inhibitors (e.g., puromycin, cycloheximide) give similar results (**Fig. S1A**, legend).] Although signal is brightest in nuclei, quantitative analysis shows that slightly more is found in the larger area of the cytoplasm (**Fig. 1Aiv**). Since starvation stresses cells and this might have unforeseen results, the experiment was repeated without prior starvation; although signals are now fainter, again nuclei are the brightest and integrated signal over the cytoplasm is the highest (**Fig. S1A**). Similar results are obtained with primary (diploid) human umbilical vein endothelial cells (HUVECs; **Fig. S1A**), so results are not peculiar to a transformed cell like HeLa.

A pulse-chase experiment confirms that Aha is incorporated by a metabolic process, and that newly-made peptides are degraded rapidly. After a 2-min Aha pulse, and regrowth in Met for up to 5 min; nuclear and cytoplasmic signals fall with half-lives of $<$1 min at 37°C (**Fig. 1Av** and **Fig. 1B**) – but not at 4°C (**Fig. S1B**). Pre-treatment with the proteasomal inhibitor, MG132 [22], increases nuclear and cytoplasmic fluorescence (**Fig. 1C**). Essentially no signal remains after a 5-min chase (**Fig. 1Av**). This contrasts with the known half-lives of human proteins (measured using pulses lasting days) of ~20 h (the range of half-lives covers minutes to many tens of hours; [11–12]). However, a 60-min pulse gives sufficient incorporation to allow some signal to survive a 5-min chase (**Fig. S1C**). These results are consistent with most newly-made Aha-tagged peptides being destroyed within minutes by the proteasome (so inhibiting it increases signal), and with a tiny fraction surviving to contribute to the 'mature' proteome (so only the latter is detected using long pulses). Of course, this does not exclude the possibility that other proteases contribute to the turnover [23]. In the unlikely event that Aha is incorporated by some unknown non-ribosomal mechanism, the resulting product must then be degraded by the proteasome. This makes it unlikely that the signal could just be due to incorporation of Aha into an amino-acyl tRNA.

A ribosome polymerizes ~5 amino acids per second *in vivo* [24], so a typical human protein with ~400 residues is polymerized in ~80 s. Therefore, the nuclear signal could result from cytoplasmic synthesis followed by rapid nuclear import. The use of a 5-s pulse eliminates this possibility, as signals are still seen in both compartments (**Fig. 1D**, **Fig. S1D**). Since only one in sixteen proteins (i.e., 400/[5×5]) will be completed during this pulse, most (>90%) nuclear signal must arise from peptides made in nuclei.

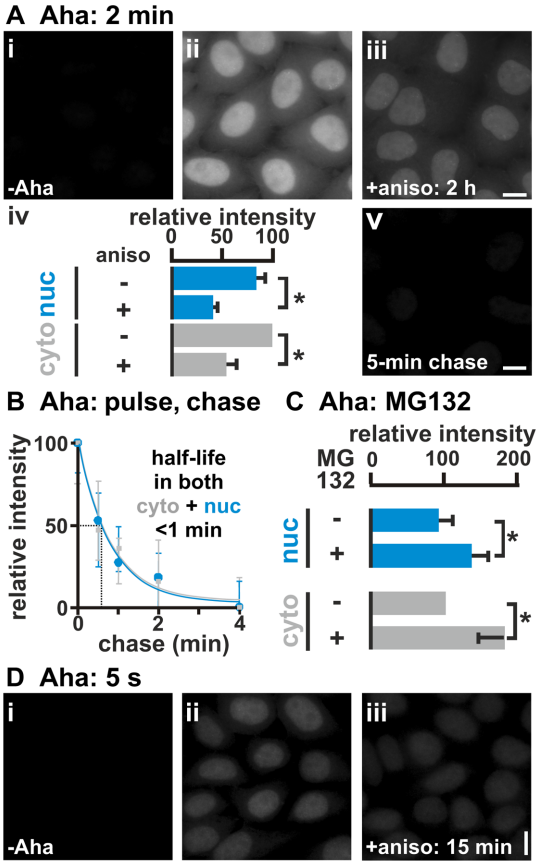


Figure 1. Aha incorporation. HeLa cells were starved of Met (15 and 30 min for 2-min and 5-s pulses, respectively), pulsed ±2 mM Aha, and chased (0–5 min; 0.2 mM Met without Aha). After fixation and 'clicking' on Alexa555, DNA was counterstained with DAPI, images collected using a wide-field microscope, and fluorescence intensities (± SD) in the cytoplasm (*cyto*) and nucleus (*nuc*) normalized relative to values in the untreated cytoplasm. *: $P<0.0001$ (Student's two-tailed *t* test, $n=20$ cells). Bars: 10 µm. (**A**) 2-min Aha pulse. (i,ii) Aha labels both nucleus and cytoplasm, with the nucleus being the brightest. (iii) Pretreatment with anisomycin (aniso; 100 µg/ml; 2 h) reduces signals in both nucleus and cytoplasm. (iv) Slightly more signal is found in the larger area of the cytoplasm. (v) Signals in both compartments disappear during a chase. (**B**) After a 2-min pulse, signals in both compartments disappear quickly during a chase. Best fits of simple exponential curves to the data are included, but note that we do not know how many kinetic populations there might be (Materials and Methods S1). (**C**) MG132 (100 µg/ml; 2 h) increases signal given by a 2-min pulse. (**D**) 5-s Aha pulse. (i-iii) Signal is sensitive to anisomycin (aniso; 100 µg/ml; 15 min).
doi:10.1371/journal.pone.0099346.g001

To confirm that the signal seen during a 5-s pulse did not result just from aminoacylation of tRNA (rather than a subsequent incorporation by a ribosome into peptide), we also examined whether destruction of RNA reduced the signal. In contrast to such an expectation, treatment with a cocktail of RNases (after fixation and immediately prior to 'clicking') increased the relative intensity in nucleus and cytoplasm to 189±26% and 139±22%, respectively (**Fig. S1**, legend). This is consistent with RNase-sensitive material in both nucleus and cytoplasm (i.e., RNA) preventing access of the 'click' reagent to Aha incorporated into RNA-free complexes (i.e., peptide) in both compartments. [See the legend of **Figure S1** for an analogous control involving DNase that demonstrates that Aha is likely to be incorporated into DNA-free complexes.]

Puromycin incorporation

Here we modify an approach applied previously to show that some translation occurs in nuclei [5]; however, we use shorter pulses and unpermeabilized cells. Cells are first treated with cycoheximide to 'freeze' ribosomes, and then incubated with puromycin – a structural mimic of aminoacyl-tRNA which is incorporated into the C-termini of nascent peptides; finally, the puromycylated and still-nascent peptides are immuno-localized using an anti-puromycin antibody [5], [25]. After growth in cycloheximide for 15 min and puromycin for 5 s, puromycylated peptides are seen in bright nucleoplasmic foci in 'confocal' images (**Fig. 2A,B**; **Fig. S2A** compares 'wide-field' and 'confocal' images). After 30 s in puromycin, the nuclear signal decreases and becomes more diffuse (**Fig. 2C**). After 60 s, the peri-nuclear region becomes the brightest (**Fig. 2D,E**). These changes mimic those seen previously and are simply interpreted in light of the known behaviour of puromycylated peptides: after dissociating from ribosomes, many accumulate at exit sites on the smooth endoplasmic reticulum (SER) before passing through the Golgi apparatus to the exterior [5], [25–27]. Additional results obtained using short pulses and the longer ones applied by David et al. [5] are also consistent with this interpretation (**Fig. S3A,B**). Moreover, similar results are obtained when pre-incubation with cycloheximide is omitted, so the labeling again cannot result from some unknown stress-induced response (**Fig. S3C**). If nuclear translation is coupled to transcription [4], inhibiting transcription should reduce nuclear (but not cytoplasmic) incorporation, and it does (**Fig. S3D**).

NanoSIMS

NanoSIMS (high-resolution secondary-ion mass spectrometry [28]) involves raster-scanning the surface of a fixed specimen with a focused ion beam to vaporize the surface of the specimen, using a mass spectrometer to measure the mass/charge ratio of secondary ions ejected during the bombardment, and then creating an image that reflects the isotopic composition of the surface. It combines high sensitivity and mass resolution, with a spatial resolution of ~100 nm in the *x* and *y* axes, and a few nanometers in the *z* axis. HeLa were grown for 2 min in 'heavy' [^{15}N]Lys and [^{15}N]Arg, and **Figure 3Ai** illustrates an image depicting the distribution of ('light') $^{12}C^{14}N^-$ ions in a typical section through the middle of a nucleus (prepared as for conventional electron microscopy). Signal due to ('heavy') $^{12}C^{15}N^-$ ions is lower (**Fig. 3Aii**); it contains contributions from ^{15}N naturally present in the biosphere, plus some from the tagged amino acids. To compare levels in different samples, signal due to heavy $^{12}C^{15}N^-$ ions is expressed as a ratio relative to both heavy and light ions, and normalized relative to the natural abundance of ^{15}N (**Fig. 3Aiii**). Then, a control unexposed to heavy medium yields a ratio characteristic of the natural abundance (*row 1*). A 10-s pulse in heavy medium increases ratios by ~6% (*row 2*); anisomycin reduces this increase (*row 3*). A 120-s pulse yields still higher ratios (*row 4*) – which are reduced by a chase at 37°C (*row 5*), but not 4°C (*row 6*). These results confirm that some peptides are made in both nucleus and cytoplasm, and that they turn over rapidly.

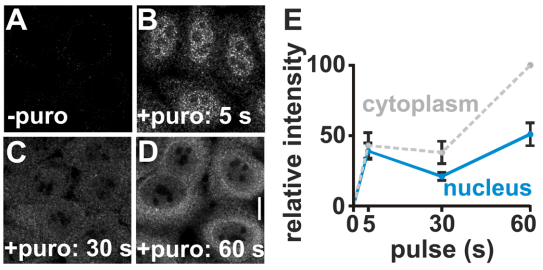


Figure 2. Puromycin incorporation. HeLa cells were pre-treated with cycloheximide (100 μg/ml; 15 min) to slow ribosomes, pulsed with puromycin (puro; 91 μM; 0–60 s), fixed, puromycylated peptides immuno-labeled with Cy3, DNA stained with DAPI, and images collected. (**A-D**) Typical confocal sections through the centre of nuclei. A 5-s pulse gives bright nuclear foci. After 30 s, cytoplasmic and nuclear signals are more similar and diffuse. After 60 s, the peri-nuclear region is the brightest. Bar: 10 μm. (**E**) Using wide-field images, Cy3 intensities (±SD; $n = 20$ cells) are expressed relative to the cytoplasmic value after a 60-s pulse.
doi:10.1371/journal.pone.0099346.g002

We next assessed how rapidly peptides made during a 2-min pulse with either [^{15}N]Met (which is incorporated by the ribosome into both the amino terminus and internally), or [^{15}N]Lys plus [^{15}N]Arg (which are not incorporated by the ribosome at the amino terminus). Most peptides tagged with [^{15}N]Met disappear from both nucleus and cytoplasm within several tens of seconds (**Fig. 3Bi**). Peptides tagged with [^{15}N]Lys plus [^{15}N]Arg disappear slightly more slowly (**Fig 3Bii**). These results are consistent with a scenario in which most translating ribosomes abort soon after initiation (with rapid destruction of the resulting peptides), and with only a minority going on to generate longer peptides that contribute to the 'mature proteome'. Then, during a pulse with an end-biased label like [^{15}N]Met, most label is incorporated into short (aborting) peptides that turn over quickly. But after a pulse with labels that can only be incorporated by ribosomal polymerization into the middle of a peptide (i.e., [^{15}N]Lys, [^{15}N]Arg), a larger minority is incorporated into longer peptides that turn over more slowly.

We further confirmed that much of the signal due to heavy label results from the formation of peptide bonds – the covalent linkage of a (carboxyl) carbon in one amino acid with the (amino) nitrogen in another. HeLa cells were starved of all amino acids for 15 min, and then incubated for 2 min in both [^{13}C]amino acids and [^{15}N]amino acids (each set being uniformly-labeled with the respective heavy atom). Protein synthesis will covalently link some ^{13}C atoms with ^{15}N atoms; this should then yield (heavy-heavy) $^{13}C^{15}N^-$ ions during NanoSIMS [29] (**Fig. 3Ci**). Most ions of this type arise from incomplete fragmentation of one peptide, or indirectly by complete fragmentation of one peptide to monatomic ^{13}C and ^{15}N which then recombine to give the diatomic ion; a minor fraction is derived by complete fragmentation of two different molecules followed by recombination of ^{13}C from one with ^{15}N from the other [29]. Then, detection of heavy-heavy ions at levels above those found naturally in the biosphere should primarily reflect the presence of ^{13}C and ^{15}N atoms lying close together within the same molecule in the sample.

To compare levels in different samples, signal due to the (heavy-heavy) $^{13}C^{15}N^-$ ions is expressed as a ratio relative to their (light-light) $^{12}C^{14}N^-$ counterparts (**Fig. 3Cii**). Then, a control unexposed to heavy medium yields the ratio characteristic of the sum of the natural abundance of both heavy isotopes (*row 1*). A 120-s pulse in heavy-heavy medium increases ratios in both nucleus and cytoplasm (*row 2*). The scale of this increase is reduced by two ribosomal inhibitors (anisomycin and cycloheximide; *rows 3,4*), and by a chase at 37°C (*row 5*) – whilst a chase at 4°C has less effect (*row 6*). These results are consistent with a significant fraction of the $^{13}C^{15}N^-$signal resulting from the formation of new peptide bonds in both nucleus and cytoplasm.

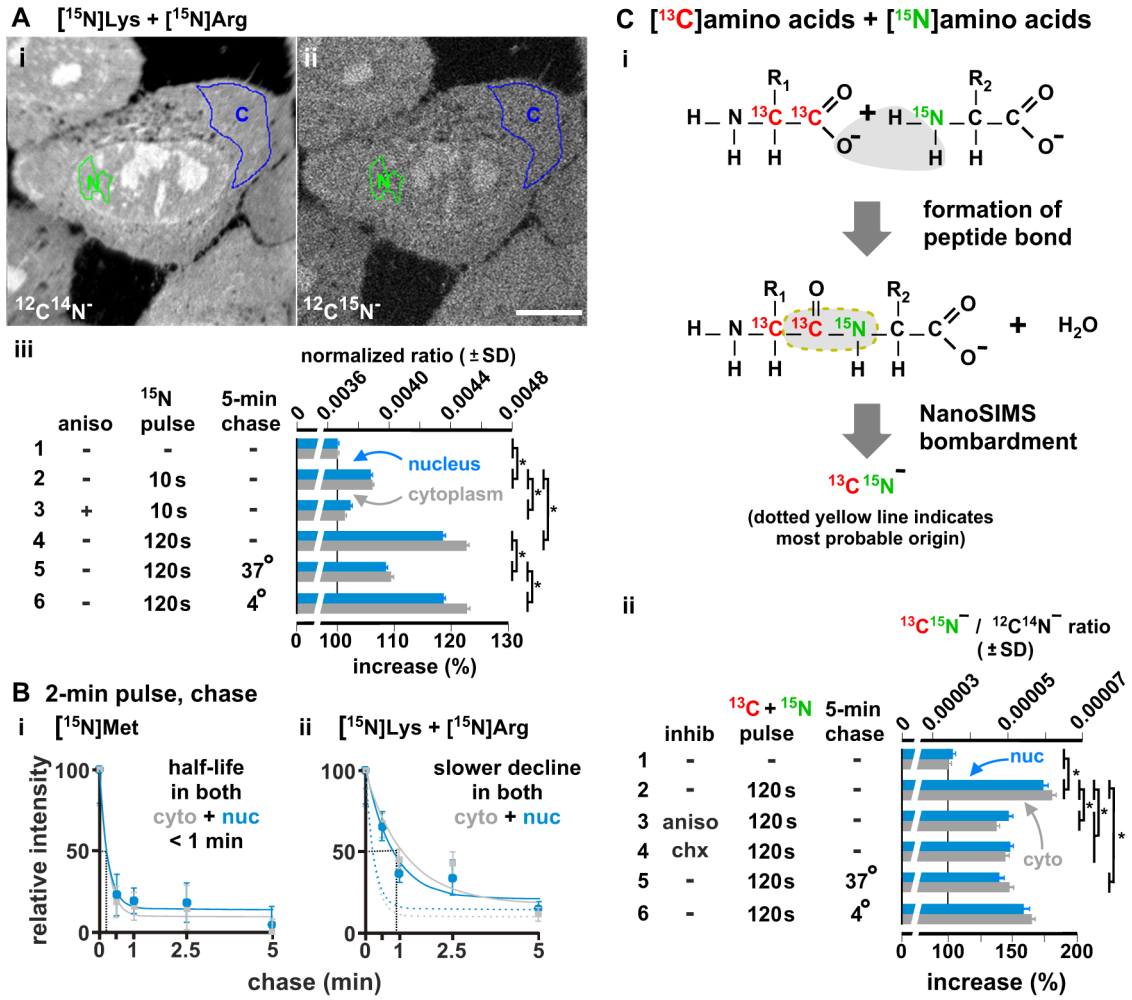


Figure 3. Incorporation of heavy amino acids detected using NanoSIMS. HeLa were starved of relevant amino acid(s), grown (10, 120 s) in heavy amino acids, fixed, sectioned, images collected, and distributions of light ($^{12}C^{14}N^-$), heavy ($^{12}C^{15}N^-$), or heavy-heavy ions ($^{13}C^{15}N^-$) measured. In some cases, inhibitors were present during starvation and the pulse, or chases followed the pulse. (A) An example after 30-min starvation of Lys + Arg, and 2-min growth in 8 mM [^{15}N]Lys +4 mM [^{15}N]Arg. (i,ii) Images show distributions of light and heavy ions in one section (regions of interest in the nucleus, N, and cytoplasm, C, indicated). Bar: 10 μm. (iii) Signals due to heavy ions are expressed as ratios (± SD) relative to total numbers of ions (i.e., $^{12}C^{15}N^-/[^{12}C^{15}N^- + ^{12}C^{14}N^-]$) normalized to the natural abundance of ^{15}N. *: $P<0.0001$ (Student's two-tailed t test; $n = 19–26$ cells). *Row 1*: cells unexposed to the heavy labels give a ratio characteristic of the natural abundance. *Rows 2–3*: a 10-s pulse in heavy medium increases ratios, and anisomycin (aniso; 100 μg/ml; 30 min) reduces this. *Rows 4–6*: a 120-s pulse yields higher ratios, which are reduced by a chase at 37°C (but not 4°C). (B) After 30-min starvation in the absence of either Met (or Lys + Arg), and a 2-min pulse in 2 mM [^{15}N]Met (or 8 mM [^{15}N]Lys +4 mM [^{15}N]Arg), signals in both compartments decline quickly during a chase. Best fits of simple exponential curves to the data are included, but note that we do not know how many kinetic populations there might be (Materials and Methods S1). (i) With [^{15}N]Met (which is incorporated into the N-terminus and internally), essentially all label disappears by 5 min. (ii) After labeling with [^{15}N]Lys + [^{15}N]Arg (which are never incorporated into the N-terminus), the decline is slower (which is consistent with these internal labels being preferentially incorporated into a minor, longer-lived, fraction). For comparison, curves (dotted lines) are reproduced from panel (i). (C) Signal is due to the formation of peptide bonds. Cells were starved (15 min) of all amino acids, and grown (2 min) in [^{13}C]amino acids + [^{15}N]amino acids ±100 μg/ml anisomycin or 100 μg/ml cycloheximide. (i) Some peptide bonds are then formed by a ribosome covalently linking a carboxyl ^{13}C atom in one [^{13}C]amino acid to an amino ^{15}N atom in a [^{15}N]amino acid; the NanoSIMS bombardment generates (directly or indirectly) $^{13}C^{15}N^-$. R_1 and R_2: different residues. (ii) After collecting images, distributions of $^{13}C^{15}N^-$ and $^{12}C^{14}N^-$ ions in regions of interest in the nucleus (nuc) and cytoplasm (cyto) were measured; signals due to (heavy-heavy) $^{13}C^{15}N^-$ are expressed as ratios (± SD) relative to those due to (light-light) $^{12}C^{14}N^-$. *: $P<0.0001$ (Student's two-tailed t test; $n = 11–18$ cells). *Row 1*: cells unexposed to heavy labels give ratios characteristic of the natural abundance. *Row 2*: growth in both heavy labels increases ratios. *Rows 3,4*: 100 μg/ml anisomycin (aniso) or cycloheximide (chx) reduces ratios. *Rows 5,6*: a chase at 37°C also reduces ratios, whilst one at 4°C gives less reduction.

doi:10.1371/journal.pone.0099346.g003

A possible role for the nuclear turnover in proof-reading nascent RNA

Finally, we examined what use the cell might make of this apparently-wasteful turnover. As a nuclear ribosome might proof-read nascent RNA for PTCs [3–4], turnover could reflect the destruction of peptides produced as a by-product. A test *Cd2* gene ± PTC (**Fig. S4**) was expressed using a multi-copy system; on transfection, the vector encoding *Cd2* replicates to generate several thousands of 'mini-chromosomes' that are co-transcribed by cellular polymerases in discrete foci [30]. *Cd2* expression is driven by a promoter normally switched off by a modified 'Tet' repressor (introduced by co-transfection with a second plasmid), but it can

be switched on by doxycycline; then, levels of nascent *Cd2* RNA are monitored by fluorescence *in situ* hybridization (FISH) using probes targeting intronic regions of the RNA. As expected, the (control) PTC̄ vector expresses high levels of these intronic regions in many nucleoplasmic foci (**Fig. 4A,B**). The PTC$^+$ vector yields less nuclear signal (**Fig. 4Ci**); presumably, because the RNA has been destroyed by NMD. Significantly, cycloheximide increases signal above that seen in the control (**Fig. 4Ci**) – consistent with it preventing ribosomes from detecting nascent PTC$^+$ transcripts (so they are no longer destroyed by NMD). This places active ribosomes close to pre-mRNA which is found only in nuclei. [Ribosomes, NMD components, and proteasomes are all found in the nucleoplasm (often at transcription sites) at concentrations that roughly equal those found in the cytoplasm [31–32]].

The experiment described above involved an artificially-introduced PTC; but how might PTCs arise normally? One possibility is that they are introduced accidentally by faulty splicing; if a splice-site is missed, the intron remains in the message and any stop codons in it will be detected as PTCs [33]. Then, inhibiting splicing with spliceostatin A (SSA; [34]) should generate unspliced transcripts with intronic PTCs, and these should be degraded by NMD to reduce the FISH signal. Consistent with this, spliceostatin reduces the signal; moreover, anisomycin reverses this reduction (**Fig. 4Cii**). In other words, inhibiting a translating ribosome prevents detection of PTCs, and subsequent degradation of intron-containing RNA [21]. These results are consistent with a translating ribosome scanning nascent transcript for PTCs, and – once found – faulty transcripts and truncated peptides produced as by-products being degraded. As the peptide turnover is significant, it follows that many nascent RNAs must encode PTCs.

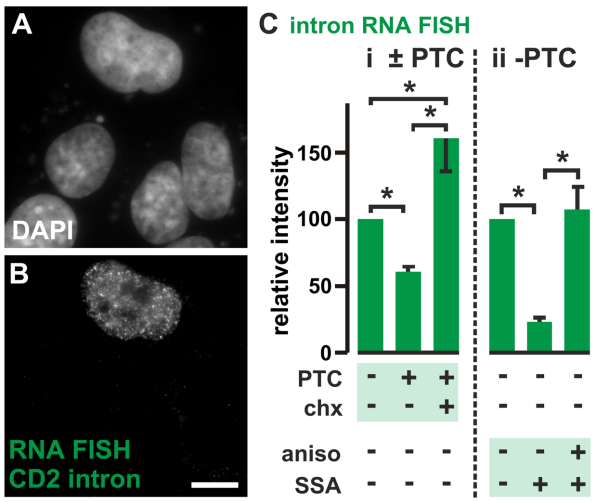


Figure 4. Effects of PTCs on nascent *Cd2* RNA. Cos-7 cells (which encode the SV40 T antigen) were co-transfected with constructs encoding the 'Tet' repressor and a test vector (with the 'Tet' promoter driving *Cd2* ± a PTC). By 24 h, the test vector (which also encodes the SV40 *ori*) replicates to give ~8,000 mini-chromosomes/cell; the 'Tet' promoter is silent and no *Cd2* RNA is detected. Now doxycycline (10 μM; 45 min) is added, cells fixed 26 h after transfection, intronic *Cd2* RNA detected by RNA FISH, DNA stained with DAPI, and images collected using a wide-field microscope. (**A,B**) Two views of one field after co-transfecting the PTC$^-$ vector (only the cell at the top was transfected and expresses intronic *Cd2* RNA in nuclear foci). Bar: 10 μm. (**C**) After subtracting background, intensities (± SD) seen in nuclei are expressed relative to the value found in untreated cells transfected with the PTC̄ vector. *: $P<0.0004$ (Student's two-tailed *t* test, $n=20$ cells). (**i**) A PTC reduces levels of intronic RNA, but cycloheximide (chx; 100 μg/ml; 2 h) more than reverses the effect. (**ii**) After transfecting the PTC̄ vector, SSA (100 ng/ml; 45 min) reduces levels of intronic RNA, and anisomycin (aniso; 100 μg/ml; 2 h) reverses this effect.
doi:10.1371/journal.pone.0099346.g004

Discussion

Initially, our aim was to re-examine whether any translation occurs in the nucleus, and to do so we applied three approaches utilizing different analogues and detection methods. In one, cells (transformed HeLa or diploid HUVECs) are pulse-labeled with Aha (a Met analogue), fixed, and a fluor 'clicked' on to the (nascent) Aha-labeled peptides [17]. In the second, cells are pre-treated with cycloheximide to 'freeze' ribosomes, pulse-labeled with puromycin (a structural mimic of aminoacyl-tRNA), fixed, and the now-puromycylated peptides immuno-localized using an anti-puromycin antibody [25]. In the third, we used a classical approach: cells are incubated in 'heavy' amino acids, and incorporated label analyzed by mass spectrometry (in our case, the spectrometer is incorporated into a microscope). To ensure most signal marks nascent peptides, we use pulses as short as 5 s (when ~25 residues are incorporated into a typical protein that will contain ~400 amino acids when complete). All three different methods give signal in both nucleus and cytoplasm, with the incorporation being sensitive to translation inhibitors (**Figs 1–3**). Surprisingly, the total amount of nuclear signal is only slightly less than the total amount seen in the cytoplasm. Extraordinarily, most signal in both compartments turns over in less than a minute (**Figs 1B**, **3B**) – and not over many hours as we expected [11–12].

We were concerned that the signal seen stems either from an adventitious binding of our labels to some unknown site within cells, and/or the incorporation of our labels by some unknown enzyme activity – and not from peptide bond formation by the ribosome. Results from the experiment described in **Figure 3C** make both possibilities unlikely. Thus, a peptide bond results from the covalent linkage of a (carboxyl) carbon in one amino acid with the (amino) nitrogen in another. Therefore, we incubated cells in both (heavy) [^{13}C]amino acids and (heavy) [^{15}N]amino acids; then, protein synthesis will covalently link some ^{13}C atoms with ^{15}N atoms in peptide bonds. Previous work has shown that heavy-heavy ions detected by mass spectrometry are generated with a higher probability when ^{13}C and ^{15}N atoms are directly bonded to each other in the same molecule [29]. After 2-min pulses, we find heavy-heavy ions in both nucleus and cytoplasm (**Fig. 3C**), consistent with the incorporation of these heavy labels into peptide bonds. As the relative levels of the heavy-heavy ions, as well as the effects of ribosomal inhibitors and chases on those levels, are essentially the same as those seen with all the other labels and the shorter pulses (i.e., the heavy Met/Lys/Arg, as well as Aha and puromycin; **Figs 1–3**), we think it fair to generalize that all our labels are incorporated by the ribosome into peptides in the expected manner.

Various additional arguments support this generalization. For example, could the incorporation of we see result from some non-ribosomal activity (e.g., through the action of a tRNA amino-acyl synthetase, or through α-amidation, trans-glutamination, or enzymes like an Arg-tRNA transferase or bacterial L/F transferase)? Probably not, as none of these non-ribosomal activities are sensitive to translational inhibitors [35]. Could our analogues be incorporated into (or attach to) molecules like DNA and RNA (e.g., through amino-acylation of a tRNA)? If so, the signal should be sensitive to DNase and RNase, and translation and proteasomal inhibitors should have no effect – but none of these apply (**Fig. 1C, Fig. S1**). Could the different analogues bind to other unknown

molecules? Again probably not, as binding would have to be sensitive to translation inhibitors, the resulting complex would have to be degraded by the proteasome, and the concentration of that molecule would have to change in extraordinary ways during pulse-chases (**Figs 1B** and **3B**).

But if ribosomes are responsible, why have the synthesis and turnover not been detected previously? Consider an analogous situation. About 95% newly-made nuclear RNA is degraded almost as soon as it is made. This turnover was only discovered in the 1950s once labeling periods shorter than the half-life of the newly-made RNA were introduced, and no credible reason for why it might occur could be proposed at the time [15]. [Only now do we now know it results from the combined destruction of prematurely-terminating (abortive) transcripts and intronic plus non-coding RNA, and only now have 'RNA-seq' techniques sensitive enough to allow accurate measurement of these unstable RNAs been developed.] Although such synthesis and destruction of RNA appears wasteful, it is nevertheless an integral part of cell metabolism. By analogy, we suggest the extraordinary turnover of most newly-made peptides went undetected because pulses longer than the half-life were used. [Previous work puts a lower limit of ~30% on the fraction of newly-made peptide that is destroyed within minutes [10], [36].] We also suggest that if many ribosomes abort soon after initiation, it is likely that the resulting short peptides will be missed using conventional approaches. Some possible reasons for this include: (i) translation was traditionally monitored using amino acids tagged with radiolabels, followed by acid precipitation of any now-tagged proteins – but short peptides are not precipitated using standard conditions [37], (ii) peptides containing fewer than ~10 amino-acid residues go undetected in conventional proteomic screens, and (iii) current biochemical approaches are insufficiently sensitive to permit detection of analogues incorporated during the short pulses used here (see **Materials and methods** for one approach we tried).

These results beg the question: does the cell utilize this apparently-wasteful turnover? In the case of the nuclear turnover, it seems it does. Thus, early results [3–4], and those in **Figure 4**, are consistent with a nuclear ribosome proof-reading nascent RNA for PTCs, and – if found – degradation of the RNA by NMD. Then, a truncated peptide is an inevitable by-product, and – as this may be toxic – we suggest it is quickly degraded.

But how might the cell utilize the cytoplasmic turnover? We can only speculate. First, nuclear proof-reading may be so error-prone the system has a second go with a cytoplasmic ribosome at weeding out unwanted PTC^{+} transcripts. Second, more proteins than hitherto expected may terminate prematurely and/or misfold, and these are probably degraded quickly [10], [13–14], [38]. Third, and in the special case of antigen-presenting cells, some newly-made peptides may be used to fight infection [39]. Finally, the apparently-wasteful turnover could play an important role in regulating translation. Here, RNA synthesis again provides a precedent. Thus, for every ~100 RNA polymerases that initiate at a promoter, ~99 will abort to release transcripts of 2–15 nucleotides [40]. These transcripts are not detected using high-throughput sequencing, as they are too short to be mapped. Subsequently, many of the polymerases that manage to 'escape' without aborting will now terminate within 20–500 nucleotides to give the sense and anti-sense transcripts copied from in/around promoters that can be detected by high-throughput sequencing [41]. As a result, the production of each completed transcript is regulated by the synthesis of hundreds of shorter sense and anti-sense products. Although this synthesis and destruction may appear energetically wasteful to us, the associated costs must be marginal – simply because the turnover is an integral part of metabolism, and because the system then goes on to discard nine-tenths of the resulting full-length transcript when introns are removed to give the mature mRNA. We suggest the same general principles apply to translation. Here, we also note that ~50% human mRNAs encode upstream open-reading frames (uORFs; [42]) that ribosome profiling indicates are translated [43–44] – presumably into peptides containing only a handful of residues that we also assume must be quickly degraded [13–14], [44–45]. Moreover, many of these uORFs are both conserved and play a role in regulating translation of their associated ORFs [43–45]. If translation is like transcription, then – during our short pulses – many ribosomes will terminate soon after initiation (both at the end of uORFs, and close to the beginning of uORFs and ORFs), most peptides generated in this way will be degraded quickly, and only a minority of ribosomes will go on to generate the peptides that contribute to the 'mature' proteome (**Fig. 5**). [Many of these ribosomes that terminate quickly may only be weakly associated with the mRNA, and so be lost during the cell lysis and treatment with the RNase used during ribosome profiling. This could explain why profiling uncovers only a small peak of bound ribosomes near initiation codons, and why translation inhibitors increase the height of this peak (i,e., by 'freezing' an unstable fraction on the message) [43], [46].] Cycles of initiation and abortion will lead to preferential incorporation of Aha or [^{15}N]Met, and to the astonishing turnover we see (because these are end labels, as well as internal labels); in contrast, incorporation of [^{15}N]Lys and [^{15}N]Arg will label more of the 'mature' proteome (because they are incorporated by the ribosome only internally). And conventional pulses lasting more than a few minutes (with end or internal labels) will completely miss the soon-to-terminate fraction.

Then, protein synthesis is like RNA synthesis: pulse-labels of different times highlight different metabolic pools, most newly-made material is degraded as soon as it is made, and the apparently-wasteful turnover is a by-product of an inefficient process. Whilst most of this turnover probably reflects 'noise', it seems likely that some will be exploited by the system to regulate production. Then, a hitherto unrecognized – and so 'dark-matter'

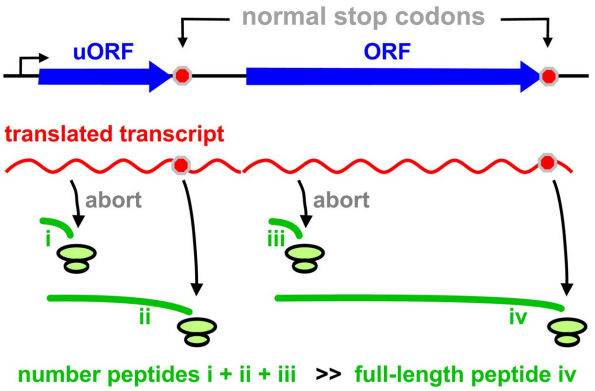


Figure 5. A model illustrating how 'dark-matter' peptides (green lines i-iii) and the 'mature' proteome (iv) arise. Most initiated ribosomes terminate prematurely (giving i and iii), and some translate to the end of an uORF (giving ii); the resulting peptides are rapidly degraded (half-life <1 min), to give rise to the astonishing turnover seen using short pulses. A minority of ribosomes translate the whole ORF (giving iv); such peptides are the ones detected conventionally using long pulses (they are generally stable and constitute the 'mature' proteome). During long pulses, most peptides i-iii are degraded and so are not detected.
doi:10.1371/journal.pone.0099346.g005

– world of short peptide will coexist with its RNA counterpart [47].

Materials and Methods

Cell culture and fixation

General conditions are described, with exceptions indicated in the legends to individual Figures. HeLa and Cos 7 cells were cultured in DMEM (PAA) supplemented with 1% penicillin/streptomycin (PAA) and FBS (Bioline; 5% for HeLa, 10% for Cos 7); HUVECs (human umbilical vein endothelial cells; Lonza; note that the supplier were responsible for obtaining the necessary ethical provision) were cultured in EGM-2 (Lonza). Prior to imaging, cells were cultured on clean glass coverslips (22×22 mm; no. 1.5) etched with 1% hydrofluoric acid (Merck). 2 mM Aha (Invitrogen, or synthesized as described in Materials and Methods S1), 100 μg/ml anisomycin (Sigma), 100 μg/ml MG132 (Enzo), 91 μM puromycin (Sigma), 100 μM DRB (Sigma), 10 μM doxycycline (Calbiochem), 100 μg/ml cycloheximide (Sigma), and 100 nM spliceostatin A (a gift of M Yoshida) were added to live cells at the times indicated. Prior to pulsing with Aha, HeLa cells were grown in the absence of Met (i.e., in DMEM̄, which is DMEM without Met, cysteine, and L-glutamine; Sigma) for 15–30 min at 37°C (to deplete endogenous pools of Met). Prior to pulsing with puromycin, cells were grown in the presence of cycloheximide (in DMEM supplemented with 5% FBS and 1% penicillin/streptomycin for HeLa, and EGM-2 for HUVECs) for 15 min at 37°C (to slow translating ribosomes). For pulses of 5–120 s in Aha or puromycin (and for pulse-chase experiments), cells on a coverslip were dipped in medium (DMEM̄ for HeLa and EGM-2 for HUVECs) containing the analogue (and then into medium lacking the analogue during a chase) in a 37°C room for the required time before plunging the coverslip immediately into ice-cold fixative. When a pulse was followed by chase, DMEM̄ + 2 mM Aha was replaced by DMEM (which contains ~200 μM Met) supplemented with 5% FBS and 1% penicillin/streptomycin, or EGM-2 in the case of HUVECs. Cells were fixed with 4% paraformaldehyde (Electron Microscopy Sciences) in 250 mM HEPES (pH 7.6; PAA) for 20 min, usually at 20°C. For transfection, Cos 7 were grown in DMEM supplemented with 10% FBS (tetracycline free; Clontech) to 80–90% occupancy in 6-well plates and transfected using 8 μl FuGENE HD (Roche/Promega) with 2 μg DNA that included 0.2 μg of Vector 1−3+ 0.2 μg pTetOn Advanced (Clontech) +1.6 μg sheared salmon sperm DNA (Invitrogen). Vectors are described in Materials and Methods S1. After 12 h, cells were trypsinized, washed, and replated on coverslips in 6-well plates (to reduce background caused by input DNA). Cells were then grown for 12 h, induced with doxycycline in the presence or absence of inhibitors, and fixed.

Fluorescent detection of Aha by 'click' chemistry

After fixation, cells were washed in PBS (PAA; 5 min; 20°C), permeabilized with 0.5% Triton X-100 (Sigma), washed in 3% BSA in PBS (5 min; 20°C), an alkyne-conjugated fluor (i.e., 100 nM Alexa555; Invitrogen) attached (using the copper-catalyzed azide-alkyne cycloaddition reaction, CuAAC) to incorporated Aha using the 'Click-iT Cell Reaction Buffer kit' (Invitrogen) as instructed by the supplier (30 min; 20°C; vigorous shaking), washed in 3% BSA in PBS (5 min; 20°C), and rinsed in PBS. After counterstaining nuclei with DAPI (Sigma) and mounting in Vectashield (Vector Laboratories), images were acquired images using a wide-field or confocal microscope.

Immuno-labeling

Fixed cells were washed with PBS (5 min; 20°C), permeabilized with 0.5% Triton X-100 and 0.5% saponin (Sigma), washed with 0.05% Tween 20 (Sigma) in PBS (10 min; 20°C), and blocked with 3% BSA and 0.2% cold water fish skin gelatin (Sigma). Puromycylated peptides were indirectly immuno-labeled by incubation (1 h; 20°C) with a mouse monoclonal antibody, 12D10 (1/100 dilution), cells washed 4x with 0.05% Tween 20 in PBS (1 min each; 20°C), incubated (30 min; 20°C) with a Cy3-conjugated donkey anti-mouse (Jackson ImmunoResearch, 715-165-150; 1/200 dilution), and washed 3x with 0.05% Tween 20 in PBS (10 min each; 20°C) followed by a wash with PBS (10 min; 20°C). After staining with DAPI and mounting in Vectashield, images were acquired with a confocal or wide-field microscope.

Fluorescence microscopy

Microscopes used were: (a) Wide-field fluorescence microscope (Zeiss; upright – Axioplan 2e microscope), equipped with a 175 W Xenon arc lamp (Perkin Elmer); images were acquired using a 63x Zeiss Plan-APOCHROMAT oil-immersion objective (numerical aperture 1.4), optical filters (Chroma), and a CoolSNAP$_{HQ}$ camera (Photometrics) running under MetaMorph software (Molecular Devices). (b) Confocal microscope (Olympus, FV1000, IX81); images were acquired using a 100x Olympus UPlanSApo oil-immersion objective (numerical aperture 1.4), optical filters (Olympus), a confocal aperture of 175 μm and scanning at 10 μs/pixel, 405 and 559 nm 'diode-pumped solid-state' or argon (488 nm) lasers, and FLUOVIEW v2.1b software (Olympus). Images were analyzed using ImageJ [48].

The fluorescence intensity over nucleus and cytoplasm (images obtained using a wide-field microscope) was determined as follows. (i) The fluorescent intensity over the whole cell and nucleus (area defined by DAPI staining) was measured, and that over the cytoplasm calculated by subtraction. (ii) Background in each compartment was subtracted; backgrounds determined separately for nucleus and cytoplasm (i.e., whole cell minus nucleus) over non-pulsed, 0-s pulsed, or untransfected cells as appropriate. (iii) Values from each experiment were exported to Excel, and mean and standard deviation (SD) calculated. *P* values (two-tailed) from unpaired Student's *t*-test were calculated using GraphPad (http://www.graphpad.com/quickcalcs/). For **Figure 4**, the mean intensity within a randomly-selected 30×30 pixel area over the nucleoplasm was measured.

Detecting Aha-tagged peptides biochemically

It is generally accepted that fluorescence microscopy can allow detection of fewer tagged molecules than conventional biochemical approaches. Nevertheless, we attempted to detect nascent Aha-tagged after running them on gels, but failed to do so successfully. For example, one robust approach we tried involved combining two analogues: after 60-s growth in Aha (with puromycin added during the last 5 s), clicking on a cleavable biotin (i.e., compound 14b from [49]), selection on Neutravidin beads, release of biotinylated peptides (using 0.5 M dithiothreiotol, and 'Western' blotting using the anti-puromycin antibody, we failed to distinguish between signal and background.

NanoSIMS

HeLa cells were grown on coverslips. For labeling with heavy Lys and Arg (**Fig. 3A**, **Fig. 3Bii**), cells were grown in custom-made DMEM without Lys + Arg (Thermo Scientific) for 30 min, regrown in DMEM (without Lys + Arg) supplemented with 8 mM [^{15}N^{13}C]Lys +4 mM [^{15}N^{13}C]Arg (Cambridge Isotopes Labora-

tories; ^{15}N 99% enriched) for 10 or 120 s at 37°C, immediately plunged into ice-cold fixative, and incubated in fixative for 20 min at 4°C (see below). [The ^{13}C tag is not used here.] For labeling with heavy Met (**Fig. 3Bi**), cells were grown in DMEM without Met and cystine (Sigma) for 30 min, regrown in the same medium supplemented with 2 mM [^{15}N]Met (Cambridge Isotopes Ltd; ^{15}N 97–99%) for 2 min, and fixed (see below). For labeling with [^{13}C]amino acids plus [^{15}N]amino acids (**Fig. 3C**), cells were starved for 15 min in custom-made DMEM lacking any amino acids (Thermo Scientific), regrown in the same medium supplemented with [^{13}C]algal amino acid mixture (Isotec; uniformly ^{13}C-labeled, 98 atom % ^{13}C) plus [^{15}N]cell-free amino acid mixture (Cambridge Isotope Labs Inc; uniformly ^{15}N-labeled, 96.98%), and fixed (see below). Here, we assume each of the 20 amino acids has a molecular weight of 136, and that each kind summed to give a final concentration of 1 mM during the pulse. In some cases, (i) 100 μg/ml anisomycin or 100 μg/ml cycloheximide was present during both the starvation and the pulse, and (ii) the pulse was followed by chases in DMEM plus 5% fetal bovine serum at 37°C or 4°C.

After the pulse, or pulse-chase, sections of cells were prepared using a procedure commonly used for electron microscopy. Cells were fixed in 4% paraformaldehyde (Electron Microscopy Sciences) and 1% glutaraldehyde (Agar) in PBS (PAA). Samples were further fixed (1 h; room temperature) with 2.5% glutaraldehyde in 0.1 M sodium cacodylate buffer (pH 7), washed in the buffer 3x (10 min each), incubated (1 h) in 1% osmium tetroxide in the same buffer, and washed with water for 20 min. Samples were dehydrated using a graded ethanol series (50%, 70%, 90%, 95% and 100%), and embedded in Agar100 epoxy resin; after removing the coverslip, semi-thin sections (0.5 μm) were cut using an ultramicrotome (Leica UC7) with a diamond knife (Diatome), mounted on platinium-coated coverslips, and coated with 5-nm platinum.

SIMS images were acquired on a CAMECA (Gennevilliers) NanoSIMS 50 to measure the ^{15}N and ^{13}C signals in the nucleus and cytoplasm. The NanoSIMS uses a 16 keV Cs^+ ion beam to bombard the sample surface, and was tuned to detect $^{12}C^{14}N^-$, $^{12}C^{15}N^-$, and $^{13}C^{15}N^-$ ions using a Mauttach-Herzog mass analyser with electrostatic sector and asymmetric magnet configuration. The instrument was tuned to avoid mass interference from $^{12}C^{14}N^-$ and $^{12}C^{15}N^-$ ions as appropriate. Images were analysed using the Image J plugin OpenMIMS (MIMS, Harvard University, www.nrims.harvard.edu). Regions of interest (ROIs) were selected in the nucleus and cytoplasm that avoided nucleoli and obvious cytoplasmic vesicles or oil droplets, and ratios calculated as indicated (i.e., $^{12}C^{15}N^-/[^{12}C^{15}N^-+^{12}C^{14}N^-]$, and $^{13}C^{15}N^-/^{12}C^{14}N^-$). The control sample in **Figure 3Aiii** *row 1* had a $^{12}C^{15}N^-/[^{12}C^{15}N^-+^{12}C^{14}N^-]$ ratio of 0.003838, which compares well with the expected value of 0.003667 (i.e., the ratio of $^{15}N/^{14}N$ in the biosphere); therefore other ratios were normalized relative to this expected value.

RNA fluorescent *in situ* hybridization (RNA FISH)

A set of forty-eight 20-mer probes (designed using Probe Designer; http://singlemoleculefish.com/) were synthesized (Biosearch Technologies, USA) targeting the intronic region (~1.6 kbp) of rat *Cd2* DNA present in Vectors 1–3 (**Fig. S4**; see Materials and Methods S1 for vector construction, and **Table S1** for the sequences of primers used during construction). At the 3′-end of each 20-mer, an mdC(TEG-amino) modification was added. The amino group was subsequently labeled [50] using the ARES Alexa Fluor 488 DNA labeling kit (Invitrogen) according to the manufacturers' instructions. Labeling efficiency was found to be ~5 fluors per 100 nucleotides. Fixed cells were processed for RNA FISH (probe sequences are listed in **Table S2**), and imaged as described [50] with the following exceptions: cells were immediately permeabilized as above for immuno-fluorescence, and the hybridization buffer contained 10% formamide. When RNA FISH was performed on EGFP-expressing cells, Alexa488 fluor was further detected by immuno-fluorescence as follows. After performing RNA FISH as described above, coverslips were washed 3x in 2xSSC (10 min each; 37°C), washed with 0.05% Tween 20 in PBS (3 min; 20°C), and blocked as above. The *Cd2* intronic RNA was detected by incubating the cells with primary antibody against Alexa 488 Fluor (rabbit polyclonal; Invitrogen, A11094; 1 μg/ml; 1 h; 20°C), washed 4x with 0.05% Tween 20 in PBS (1 min each; 20°C), then labeled further by incubating these cells with Cy3-conjugated donkey anti-rabbit (Jackson ImmunoResearch, 711-165-152; 1/2000 dilution; 30 min; 20°C), washed 4x with 0.05% Tween 20 in PBS (5 min each; 20°C) followed by a rinse with PBS at (5 min; 20°C). After DAPI/Vectashield counterstaining and mounting, images were collected using a wide-field microscope.

Supporting Information

Figure S1 Aha incorporation; some controls.
(TIF)

Figure S2 Comparison of images obtained using wide-field and confocal microscopes.
(TIF)

Figure S3 Puromycin incorporation: some controls.
(TIF)

Figure S4 CD2-EGFP expression constructs.
(TIF)

Table S1 Oligonucleotide DNA primers.
(PDF)

Table S2 RNA FISH probes against rat *Cd2* intron-2[a].
(DOCX)

Materials and Methods S1
(DOCX)

Acknowledgments

We thank Garry Brown (for pLdPDH), Minoru Yoshida (for SSA), Daniela Dieterich (for the cleavable biotin tag), David Vaux, Jordan Raff, and Katie Moore (for advice), Argyris Papantonis and Joshua Larkin (for discussion), and Jon Bartlett and Errin Johnson (for help).

Author Contributions

Conceived and designed the experiments: SB BB PRC BGD CRMG HJ. Analyzed the data: SB BB HJ CRMG PP BGD PRC. Wrote the paper: SB BB HJ CRMG PP BGD PRC. Performed experiments using Aha, puromycin, heavy amino acids, and RNA FISH: SB. Made reagents used for click chemistry and performed Aha labeling with SB: BB. Performed NanoSIMS experiments: HJ. Made the anti-puromycin antibody: PP.

References

1. Miller OL Jr, Hamkalo BA, Thomas CA Jr (1970) Visualization of bacterial genes in action. Science 169: 392–395. doi: 10.1126/science.169.3943.392.

2. Goidl JA, Allen WR (1978) Does protein synthesis occurs within the nucleus? Trends Biochem Sci 3: N225–N228.

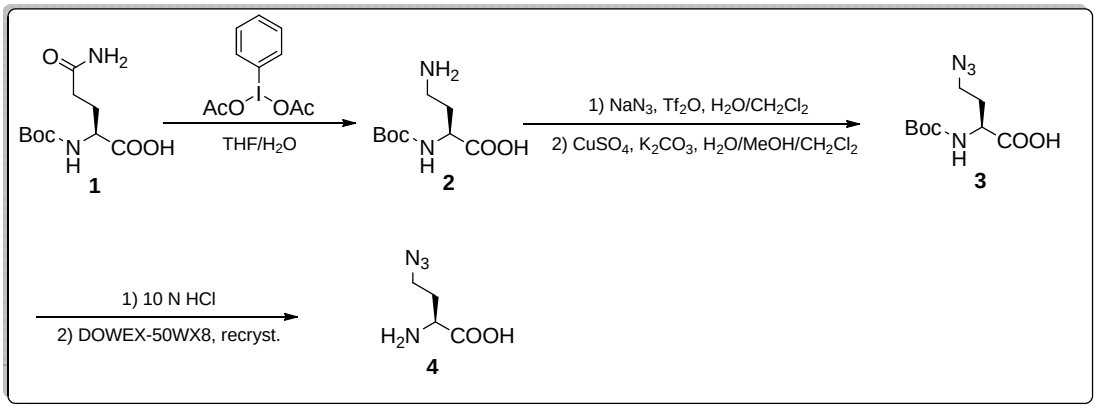
3. Maquat LE (2002) NASty effects on fibrillin pre-mRNA splicing: another case of ESE does it, but proposals for translation-dependent splice site choice live on. Genes Dev 16: 1743–1753. doi: 10.1101/gad.1014502.
4. Iborra FJ, Jackson DA, Cook PR (2001) Coupled transcription and translation within nuclei of mammalian cells. Science 293: 1139–1142. doi: 10.1126/science.1061216.
5. David A, Dolan BP, Hickman HD, Knowlton JJ, Clavarino G, et al. (2012) Nuclear translation visualized by ribosome-bound nascent chain puromycylation. J Cell Biol 197: 45–57. doi: 10.1083/jcb.201112145.
6. Al-Jubran K, Wen J, Abdullahi A, Roy Chaudhury S, Li M, et al. (2013) Visualization of the joining of ribosomal subunits reveals the presence of 80S ribosomes in the nucleus. RNA 19: 1669–1683. doi: 10.1261/rna.038356.113.
7. Apcher S, Millot G, Daskalogianni C, Scherl A, Manoury B, et al (2013) Translation of pre-spliced RNAs in the nuclear compartment generates peptides for the MHC class I pathway. Proc Natl Acad Sci USA 110: 17951–17956. doi: 10.1073/pnas.1309956110.
8. Yewdell JW, David A (2013) Nuclear translation for immunosurveillance. Proc. Natl. Acad. Sci. USA 110: 17612–17613. doi: 10.1073/pnas.1318259110.
9. Dahlberg J, Lund E (2012) Nuclear translation or nuclear peptidyl transferase? Nucleus 3: 320–321. doi: 10.4161/nucl.20754.
10. Wheatley DN, Giddings MR, Inglis MS (1980) Kinetics of degradation of "short-" and "long-lived" proteins in cultured mammalian cells. Cell Biol Int Rep 4: 1081–1090. doi: 10.1016/0309-1651(80)90045-4.
11. Doherty MK, Hammond DE, Clague MJ, Gaskell SJ, Beynon RJ (2009) Turnover of the human proteome: determination of protein intracellular stability by dynamic SILAC. J Proteome Res 8: 104–112. doi: 10.1021/pr800641v.
12. Boisvert FM, Ahmad Y, Gierlinski M, Charriere F, Lamont D, et al. (2012) A quantitative spatial proteomics analysis of proteome turnover in human cells. Mol Cell Proteomics 11: M111011429. doi: 10.1074/mcp.M111.011429.
13. Bachmair A, Finley D, Varshavsky A (1986) *In vivo* half-life of a protein is a function of its amino-terminal residue. Science 234: 179–186. doi: 10.1126/science.3018930.
14. Turner GC, Varshavsky A (2000) Detecting and measuring cotranslational protein degradation in vivo. Science 289: 2117–2120. doi: 10.1126/science.289.5487.2117.
15. Harris H (1994) An RNA heresy in the fifties. Trends Biochem Sci 19: 303–305. doi: 10.1016/0968-0004(94)90011-6.
16. van Kasteren SI, Kramer HB, Jensen HH, Campbell SJ, Kirkpatrick J, et al. (2007) Expanding the diversity of chemical protein modification allows post-translational mimicry. Nature 446: 1105–1109. doi: 10.1038/nature05757.
17. Tom Dieck S, Muller A, Nehring A, Hinz FI, Bartnik I, et al. (2012) Metabolic labeling with noncanonical amino acids and visualization by chemoselective fluorescent tagging. Curr Protoc Cell Biol 7: 11. doi: 10.1002/0471143030.cb0711s56.
18. Hinz FI, Dieterich DC, Tirrell DA, Schuman EM (2012) Non-canonical amino acid labeling in vivo to visualize and affinity purify newly synthesized proteins in larval zebrafish. ACS Chem Neurosci 3: 40–49. doi: 10.1021/cn2000876.
19. Kiick KL, Saxon E, Tirrell DA, Bertozzi CR (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. Proc Natl Acad Sci USA 99: 19–24. doi: 10.1073/pnas.012583299.
20. Hansen JL, Moore PB, Steitz TA (2003) Structures of five antibiotics bound at the peptidyl transferase center of the large ribosomal subunit. J Mol Biol 330: 1061–1075. doi: 10.1016/s0022-2836(03)00668-5.
21. Carter MS, Doskow J, Morris JP, Li S, Nhim RP, et al. (1995) A regulatory mechanism that detects premature nonsense codons in T-cell receptor transcripts in vivo is reversed by protein synthesis inhibitors *in vitro*. J Biol Chem 270: 28995–29003. doi: 10.1074/jbc.270.48.28995.
22. Palombella VJ, Rando OJ, Goldberg AL, Maniatis T (1994) The ubiquitin-proteasome pathway is required for processing the NF-κB1 precursor protein and the activation of NF-κB. Cell 78: 773–785. doi: 10.1016/s0092-8674(94)90482-0.
23. Reits E, Griekspoor A, Neijssen J, Groothuis T, Jalink K, et al. (2003) Peptide diffusion, protection, and degradation in nuclear and cytoplasmic compartments before antigen presentation by MHC class I. Immunity 18: 97–108. doi: 10.1016/s1074-7613(02)00511-3.
24. Nielsen PJ, McConkey EH (1980) Evidence for control of protein synthesis in HeLa cells via the elongation rate. J Cell Physiol 104: 269–281. doi: 10.1002/jcp.1041040302.
25. Schmidt EK, Clavarino G, Ceppi M, Pierre P (2009) SUnSET, a nonradioactive method to monitor protein synthesis. Nat Methods 6: 275–277. doi: 10.1038/nmeth.1314.
26. Redman CM, Sabatini DD (1966) Vectorial discharge of peptides released by puromycin from attached ribosomes. Proc Natl Acad Sci USA 56: 608–615. doi: 10.1073/pnas.56.2.608.
27. Siuta-Mangano P, Lane MD (1981) Very low density lipoprotein synthesis and secretion. Extrusion of apoprotein B nascent chains through the membrane of the endoplasmic reticulum without protein synthesis. J Biol Chem 256: 2094–2097.
28. Boxer SG, Kraft ML, Weber PK (2009) Advances in imaging secondary ion mass spectrometry for biological samples. Annu Rev Biophys 38: 53–74. doi: 10.1146/annurev.biophys.050708.133634.
29. McMahon G, Saint-Cyr HF, Lechene C, Unkefer CJ (2006) CN^- secondary ions form by recombination as demonstrated using multi-isotope mass spectrometry of ^{13}C- and ^{15}N-labeled polyglycine. J Am Soc Mass Spectrom 17: 1181–1187. doi: 10.1016/j.jasms.2006.04.031.
30. Xu M, Cook PR (2008) Similar active genes cluster in specialized transcription factories. J Cell Biol 181: 615–623. doi: 10.1083/jcb.200710053.
31. Iborra FJ, Escargueil AE, Kwek KY, Akoulitchev A, Cook PR (2004) Molecular cross-talk between the transcription, translation, and nonsense-mediated decay machineries. J Cell Sci 117: 899–906. doi: 10.1242/jcs.00933.
32. Catic A, Suh CY, Hill CT, Daheron L, Henkel T, et al. (2013) Genome-wide map of nuclear protein degradation shows NCoR1 turnover as a key to mitochondrial gene regulation. Cell 155: 1380–1395. doi: 10.1016/j.cell.2013.11.016.
33. Hentze MW, Kulozik AE (1999) A perfect message: RNA surveillance and nonsense-mediated decay. Cell 96: 307–310. doi: 10.1016/s0092-8674(00)80542-5.
34. Kaida D, Motoyoshi H, Tashiro E, Nojima T, Hagiwara M, et al. (2007) Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. Nat Chem Biol 3: 576–583. doi: 10.1038/nchembio.2007.18.
35. Finking R, Marahiel MA (2004) Biosynthesis of nonribosomal peptides. Annu Rev Microbiol 58: 453–488. doi: 10.1146/annurev.micro.58.030603.123615.
36. Dolan BP, Bennink JR, Yewdell JW (2011) Translating DRiPs: progress in understanding viral and cellular sources of MHC class I peptide ligands. Cell Mol Life Sci 68: 1481–1489. doi: 10.1007/s00018-011-0656-z.
37. Greenberg NA, Shipe WF (1979) Comparison of the abilities of trichloroacetic, picric, sulfosalicylic, and tungstic acids to precipitate protein hydrolysates and proteins. J Food Sci 44: 735–737. doi: 10.1111/j.1365-2621.1979.tb08487.x.
38. Yewdell JW, Nicchitta CV (2006) The DRiP hypothesis decennial: support, controversy, refinement and extension. Trends Immunol 27: 368–373. doi: 10.1016/j.it.2006.06.008.
39. Apcher S, Manoury B, Fahraeus R (2012) The role of mRNA translation in direct MHC class I antigen presentation. Curr Opin Immunol 24: 71–76. doi: 10.1016/j.coi.2012.01.007.
40. Goldman SR, Ebright RH, Nickels BE (2009) Direct detection of abortive RNA transcripts in vivo. Science 324: 927–928. doi: 10.1126/science.1169237.
41. Seila AC, Calabrese JM, Levine SS, Yeo GW, Rahl PB, et al. (2008) Divergent transcription from active promoters. Science 322: 1849–1851. doi: 10.1126/science.1162253.
42. Jackson RJ, Hellen CU, Pestova TV (2010) The mechanism of eukaryotic translation initiation and principles of its regulation. Nat Rev Mol Cell Biol 11: 113–127. doi: 10.1038/nrm2838.
43. Ingolia NT, Lareau LF, Weissman JS (2011) Ribosome profiling of mouse embryonic stem cells reveals the complexity and dynamics of mammalian proteomes. Cell 147: 789–802. doi: 10.1016/j.cell.2011.10.002.
44. Fritsch C, Herrmann A, Nothnagel M, Szafranski K, Huse K, et al. (2012) Genome-wide search for novel human uORFs and N-terminal protein extensions using ribosomal footprinting. Genome Res 22: 2208–2218. doi: 10.1101/gr.139568.112.
45. Calvo SE, Pagliarini DJ, Mootha VK (2009) Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc Natl Acad Sci USA 106: 7507–7512. doi: 10.1073/pnas.0810916106.
46. Lee S, Liu B, Lee S, Huang SX, Shen B, et al. (2012) Global mapping of translation initiation sites in mammalian cells at single-nucleotide resolution. Proc Natl Acad Sci USA 109: E2424–E2432. doi: 10.1073/pnas.1207846109.
47. Kapranov P, St Laurent G (2012) Dark matter RNA: existence, function, and controversy. Front Genet 3: 60. doi: 10.3389/fgene.2012.00060.
48. Abramoff MD, Magahaes PJ, Ram SJ (2004) Image processing with ImageJ. Biophotonics Int 11: 36–42.
49. Szychowski J, Mahdavi A, Hodas JJ, Bagert JD, Ngo JT, et al. (2010) Cleavable biotin probes for labeling biomolecules via azide-alkyne cycloaddition. J Am Chem Soc 132: 18351–18360. doi: 10.1021/ja1083909.
50. Papantonis A, Larkin JD, Wada Y, Ohta Y, Ihara S, et al. (2010) Active RNA polymerases: mobile or immobile molecular machines? PLoS Biol 8: e1000419. doi: 10.1371/journal.pbio.1000419.

SUPPLEMENTARY INFORMATION

MATERIALS AND METHODS

Synthesis of L-azidohomoalanine (Aha)

Aha was synthesized in three steps: (i) *N*-tert-butoxycarbonyl-(S)-2,4-diaminobutanoic acid was synthesized (adapted from [51]) from *N*-tert-butoxycarbonyl-(*S*)-glutamine (Sigma). (ii) *N*-tert-butoxycarbonyl-(*S*)-azidohomoalanine was synthesized (adapted from [52]) from *N*-tert-butoxycarbonyl-(S)-2,4-diaminobutanoic acid. (iii) Aha was synthesized [53] from *N*-tert-butoxycarbonyl-(*S*)-azidohomoalanine. Properties of the products in steps 2 and 3 were compared with those made previously [54].



Melting points (m.p.) were recorded on a Leica Galen III hot stage microscope equipped with a Testo 720 thermocouple probe and are uncorrected. Proton nuclear magnetic resonance (^{1}H NMR) spectra were recorded on a Bruker AVII500 (500 MHz) or on a Bruker DQX400 (400 MHz) spectrometer, as indicated. Carbon nuclear magnetic resonance (^{13}C NMR) spectra were recorded on a Bruker AVII500 (125 MHz) or on a Bruker DQX400 (100 MHz) spectrometer, as indicated. NMR Spectra were fully assigned using COSY, HSQC, HMBC, and NOESY. All chemical shifts are quoted on the δ scale in ppm using residual solvent as the internal standard

(^{1}H NMR: $CDCl_3$ = 7.26, CD_3OD = 4.87; DMSO-d_6 = 2.50 and ^{13}C NMR: $CDCl_3$ = 77.0; CD_3OD = 49.0; DMSO-d_6 = 39.5). Coupling constants (*J*) are reported in Hz with the following splitting abbreviations: s = singlet, d = doublet, t = triplet, q = quartet, quin = quintet, and a = apparent.

Infrared (IR) spectra were recorded on a Bruker Tensor 27 Fourier Transform spectrophotometer using thin films on NaCl plates for liquids and oils and KBr discs for solids and crystals. Absorption maxima (υ_{max}) are reported in wavenumbers (cm^{-1}). For compound characterization low resolution mass spectra (LRMS) were recorded on a Waters Micromass LCT Premier TOF spectrometer using electrospray ionization (ESI) and high resolution mass spectra (HRMS) were recorded on a Bruker MicroTOF ESI mass spectrometer. Nominal and exact *m/z* values are reported in Daltons. Optical rotations were measured on a Perkin–Elmer 241 polarimeter with a path length of 1.0 dm and are reported with implied units of 10^{-1} deg cm^2 g^{-1}. Concentrations (c) are given in g/100 ml. Thin layer chromatography (TLC) was carried out using Merck aluminium backed sheets coated with 60F254 silica gel. Visualization of the silica plates was achieved using a UV lamp (λ_{max} = 254 nm), and/or ammonium molybdate (5 % in 2 M H_2SO_4), and/or potassium permanganate (5 % $KMnO_4$ in 1 M NaOH with 5 % potassium carbonate). Flash column chromatography was carried out using BDH 40–63 μm silica gel (VWR). Mobile phases are reported in relative composition (*e.g.* 1:2:4 H_2O/*i*PrOH/EtOAc). Anhydrous solvents were purchased from Fluka or Acros. Triethylamine was stored over NaOH pellets. All other solvents were used as supplied (Analytical or HPLC grade), without prior purification. Distilled water was used for chemical reactions and Milli–QR purified water for protein manipulations. Reagents were purchased from Aldrich and used as supplied, unless otherwise indicated. 'Petrol' refers to the fraction of light petroleum ether boiling in the range 40–60 ºC. All reactions using anhydrous conditions were performed using flame-dried apparatus under an atmosphere of argon or nitrogen. Brine refers to a saturated solution of sodium chloride.

Anhydrous magnesium sulfate ($MgSO_4$) was used as drying agents after reaction workup, as indicated. DOWEX 50WX8 (H^+ form) was conditioned as follows: 100 g of the commercial resin was placed in a 500 mL sintered filter funnel and allowed to swell with 200 mL of acetone for 5 minutes. The solvent was removed by suction and the resin was washed successively with 800 mL of acetone, 500 mL methanol, 500 mL 5 M HCl, and then 1 L of water or until the pH of filtrate was ~ 7, as indicated by pH paper. The resin was partially dried on the filter and then stored and used as needed.

The three steps were carried out as follows:

1. *N*-tert-butoxycarbonyl-(S)-2,4-diaminobutanoic acid (**2**): The reaction was adapted from literature [51]. *N*-tert-butoxycarbonyl-(*S*)-glutamine (**1**; Sigma; 5.0 g, 20.3 mmol) was dissolved in THF (48 mL) and water (12 mL). The solution was cooled in an ice bath, and (diacetoxyiodo)benzene (7.8 g, 24.2 mmol) was added. The reaction was monitored by TLC (4:1:1 n-butanol: acetic acid: water). After 6 hours, the organic portion was evaporated *in vacuo*. The aqueous portion was diluted with water, and extracted with ethyl acetate (3 x 30 mL). The aqueous portion was then partially evaporated, frozen, and lyophilised. The product *N*-tert-butoxycarbonyl-(S)-2,4-diaminobutanoic acid was obtained as an orange-yellow solid (3.1 g, 71%). IR (ν_{max}, film): 3416, 2977, 2931, 1698, 1589, 1561, 1529, 1499, 1439, 1392, 1366, 1252, 1161, 1052, 1029, 949, 907, 868, 802. ^{1}H NMR (400 MHz, D_2O) δ: 3.86 (m, 1H, C*H*), 2.94 (t, *J* = 7.8 Hz, 2H, NH_2CH_2), 1.9 (m, 1H, CH_2), 1.85 (m, 1H, CH_2), 1.31 (s, 9H, $C(CH_3)_3$). ^{13}C (100 MHz, D_2O) δ: 28.3 ($(\underline{C}H_3)_3$), 29.9 (Cβ), 39.1 ($\underline{C}H_2N_3$), 57.1 (Cα), 80.9 ($(CH_3)_3\underline{C}CONH$)**,** 157.9 ($(CH_3)_3C\underline{C}ONH$), 175.6 (COOH), LRMS (ESI) Calculated: [M-H] $C_9H_{18}N_2O_4$ (m/z): 218, Obtained: $[M-H]^-$: 217.12, $[2M-H]^-$: 435.26, $[M+H]^+$: 219.16, Melting pt. = 196 ºC, Lit. [51] 190-191 ºC, $[\alpha]^{20}_D$ = -8.6º (c = 1.0, H_2O), Lit. [51] $[\alpha]^{20}_D$ = -9.6 (c = 1.0, H_2O).

2. *N*-tert-butoxycarbonyl-(*S*)-azidohomoalanine (**3**): Triflyl azide was prepared *in situ*, and the reaction was adapted from the literature [52]. Sodium azide (9.8 g, 150 mmol) was dissolved in a mixture of water (22 mL) and dichloromethane (39 mL). The contents were cooled in an ice bath, and triflic anhydride (5.1 mL) was added via syringe over 5 min. The contents were allowed to stir for 30 min, after which the layers were separated. The aqueous layer was extracted using dichloromethane (2 x 50 mL). The organic fractions were pooled, and washed twice with 5% aq. sodium carbonate (30 mL). The triflyl azide containing organic layer was then added to a solution of *N*-tert-butoxycarbonyl-(S)-2,4-diaminobutanoic acid (3.1 g, 14.2 mmol) in water (39 mL) and methanol (90 mL). To this, potassium carbonate (3.2 g) and copper sulphate pentahydrate (38.4 mg) were added. Stirring was continued for 18 h, after which the organic fraction was evaporated *in vacuo*. The aqueous portion was extracted with dichloromethane (2 x 50 mL), diluted with water, then acidified to pH 2.0 using 1.0 N hydrochloric acid. The acidified aqueous phase was then extracted with dichloromethane (3 x 50 mL). The organic phase was then washed with brine (2 x 80 mL), dried with magnesium sulphate, filtered, and evaporated. The product was obtained as orange oil (3.2 g, 92%). IR (ν_{max}, film): 3319, 2980, 2935, 2100, 1690, 1513, 1455, 1399, 1369, 1236, 1189, 1154, 1093, 1057, 1029, 940, 853. ^{1}H NMR (400 MHz, $CDCl_3$) δ: 5.26 (br d, *J* = 8.8 Hz, 1H, N<u>H</u>), 4.41 (m, 1H, C<u>H</u>), 3.46 (t, *J* = 6.7 Hz, 2H, $N_3C\underline{H}_2$), 2.11 (m, 1H, C<u>H</u>$_2$), 1.98 (m, 1H, C<u>H</u>$_2$), 1.46 (s, 9H, $C(C\underline{H}_3)_3$). ^{13}C (100 MHz, D_2O) δ: 28.2 ($(\underline{C}H_3)_3$), 31.4 (<u>C</u>β), 47.8 ($\underline{C}H_2N_3$), 51.3 (<u>C</u>α), 80.7 ($(CH_3)_3\underline{C}CONH$), 155.7 ($(CH_3)_3C\underline{C}ONH$), 175.8 (<u>C</u>OOH), LRMS (ESI) Calculated: [M-H] $C_9H_{16}N_4O_4$ (m/z) : 243, Obtained: $[M-H]^-$: 243.08, $[2M-H]^-$: 487.16, $[\alpha]^{20}_D$ = +16.1º (c = 1.0, $CHCl_3$), Lit. [54] $[\alpha]^{20}_D$ = +19.0º (c = 1.0, $CHCl_3$).

3. L-azidohomoalanine (**4**): The reaction was performed as reported in literature [53]. *N*-tert-butoxycarbonyl-(*S*)-azidohomoalanine (2.9 g, 11.88 mmol) was dissolved in 10 N HCl (30 mL), and the contents stirred at room temperature for 1 hour. The solution was then diluted with 170

mL water. 80 g of DOWEX50-WX8 resin was packed into a column, and conditioned by washing with 1N aq. ammonia (3 x 80 mL), water (600 mL), 1N HCl (250 mL), and water again until the eluate was at pH 7.0. The peptide solution was run through the column; the flow-through was collected and passed over the resin again. Aha was eluted from the column by washing with 500 mL of 1N aq. ammonia, and recovered as an off-white solid on evaporation. The compound was recrystallised from acetone/water to yield white crystals in 91% yield (1.54 g). IR (ν_{max}, film): 3368, 3036, 2978, 2926, 2103, 1677, 1589, 1489, 1351, 1333, 1265, 1204, 1101, 854, 771. ^{1}H NMR (400 MHz, D_2O) δ: 3.99 (t, J = 6.32, 1H, Cα<u>H</u>), 3.48 (T, J = 6.4, 2H, C<u>H</u>$_2$N$_3$), 2.07 (m, 2H, Cβ<u>H</u>$_2$). ^{13}C (100 MHz, D_2O) δ: 29.6 (Cβ), 47.6 (<u>C</u>H$_2$N$_3$), 57.8 (Cα), 171.9 (<u>C</u>OOH). LRMS (ESI) Calculated: [M+H] $C_4H_8N_4O_2$ (m/z): 144, Obtained: $[M+H]^+$: 145.08, $[M+HCl-H]^-$: 179.06, $[2M-H]^-$: 287.11, Melting pt. 180.6 ˚C, $[\alpha]^{20}_D$ = +6.60º (c = 0.50, H_2O), Lit. [54] $[\alpha]^{20}_D$ = +6.60º (c = 0.50, H_2O).

Half-lives

Figures 1B and **3B** include a simple exponential decay curve

$$i_t = (i_{max} - Plateau)e^{-\lambda t} + Plateau$$

fitted to the data (using 'Prism'; Graphpad) – but note that we do not know how many different kinetic populations there might be. The half-life ($t_{1/2}$) of tagged peptides can then be calculated using $t_{1/2} = 0.693/\lambda$ where λ is the rate constant of exponential decay, i_{max} is the maximum intensity (i.e., mean intensity after the pulse without a chase), and i_t is the intensity at time 't' (i.e., mean intensity at time 't' after initiating the chase). For labeling with Aha, heavy Met, and heavy Lys + Arg, *Plateau* values for the cytoplasm (and nucleus) were 03.5, 0.09, and 0.19 (and 3.3, 0.13, and 0.21) respectively, and half-lives were 33, 10, and 48 s (and 34, 10, and 32 s) respectively. R^2 values (the square of the Pearson product-moment correlation coefficient; high values indicate a good fit) for all exponential decay curves shown were >0.98, except for those

given by heavy Lys + Arg (which gave values of 0.92 and 0.96 for cytoplasm and nucleus, respectively).

Molecular cloning

Three vectors were used (**Fig. S4**); all encode an SV40 *ori* (to permit replication in Cos7 cells; [55]) and a modified P^{tight} 'Tet' promoter inducible with doxycycline. In the base vector, the P^{tight} promoter drives expression of a protein containing an N-terminal mitochondrial signal sequence (mt), the rat CD2 epitope recognized by the OX34 antibody [56-57] and encoded by exons 2 and 3 of rat *Cd2*, a shortened intron 2 of *Cd2* (with intact splice donor and acceptor sites), a ribosome pause sequence (rps), a C-terminal EGFP, and a poly-adenylation signal. Vector 1 is as the base vector except that the EGFP sequence contains the segment encoding the EGFP fluorochrome deleted (to free the green channel for other use). Vector 2 is as the base vector except that a haemagglutinin (HA) tag is inserted within the intron (in-frame with exon 2 and upstream of the first intronic stop codon). Vector 3 is as Vector 2 but also bears a PTC 67 nucleotides upstream of the exon2-intron2 junction (a position that should induce NMD; [58]). CD2 protein and the HA tag were expressed in the expected manner by these vectors (not shown); here, we only analyze RNA expression.

The base vector was created as follows. pLdPDH [59] encodes the mitochondrial import sequence of human *PDH1α* inserted between the *Sac*I and *Pst*I sites in pEGFP-N1 (Clontech). The P^{tight} promoter from pTRE-Tight (Clontech) was inserted between the *Xho*I and *Eco*RI sites in pLdPDH. Two fragments of rat *Cd2* were amplified from total rat DNA (Bioline) using oligo pair 1 (which target chromosome 2, between positions 196,332,778-196,333,452; [60]) and oligo pair 3 nested with oligo pair 2 (which target chromosome 2 between positions 196,336,082-196,338,119; [60]). These two *Cd2* fragments of 676 and 2,057 bp (with *Nhe*I- and *Xba*I-compatible cohesive ends, respectively) were simultaneously inserted between the *Pst*I and

*Bam*HI sites in pLdPDH (which now encodes P^{tight}). The base vector was formed by removing from the doubly-modified pLdPDH the fragments containing P^{tight}, the mitochondrial import sequence, and *Cd2*, and inserting the three between the *Xho*I and *Bam*HI sites in plasmid 0,*EGFP*,pA (a promoter-less form of pEGFP-N1; [30]). Vector 1 was constructed by deleting the EGFP chromophore (5′-CTGACCTACGGCGTGCAG-3′; 18 bp) from the base vector by site-directed mutagenesis ('QuikChange Lightning kit'; Stratagene) using oligo pair 4. Vector 2 was constructed by inserting a sequence encoding the HA epitope (5′-TACCCTTACGACGTTCCTGATTACGCT-3′; 27 bp) in the base vector by site-directed mutagenesis using oligo pair 5. Unexpectedly, the insert proved to be 81 bp (i.e., 5′-TACCCTTACGACGTTCCTGATTACGCTTAGCAAGTTGGGGGGACGGGTGCTATGTACCCTTACGACGTTCCTGATTACGCT-3′) in which two HA encoding sequences were interrupted by a 27-bp sequence (note that the HA tag is not used here, so that this modification has no impact on the experiments discussed here). Vector 3 was constructed (by site-directed mutagenesis using oligo pair 6) by introducing a point mutation (5′-TAT-3′ → 5′-TAA-3′) that creates an in-frame premature termination codon (PTC) in the *Cd2* exon in Vector 3. Correct construction of the vectors was confirmed by DNA sequencing after PCR with the following oligonucleotides. Presence of P^{tight} and *Cd2* fragments in the base vector was verified using oligos 7 and 8. Absence of the EGFP chromophore in Vector 1 was verified using oligo 9. Oligo 7 was also used to verify the presence of the HA-coding sequence in Vector 2 and the point mutation in Vector 3. Bacterial transformations were performed using One Shot TOP10 chemically-competent *Escherichia coli*. All PCRs (except for site-directed mutagenesis) were performed using PfuUltra II Fusion HS DNA Polymerase (Stratagene) at a re-annealing temperature of 58.6˚C (64˚C in case of oligo pair 3). All oligonucleotide DNA primers used (supplied by Sigma) are listed in **Table S1**.

References for Supplementary Information

51. Andruszkiewicz R, Rozkiewicz D (2004) An improved preparation of N 2-tert-butoxycarbonyl- and N 2-benzyloxy-carbonyl-(S)-2,4-diaminobutanoic acids. Synth Commun 34: 1049–1056. doi: 10.1081/scc-120028636.

52. Lundquist JT, Pelletier JC (2001) Improved solid-phase peptide synthesis method utilizing α-azide-protected amino acids. Org Lett 3: 781–783. doi: 10.1021/ol0155485.

53. Link AJ, Vink MK, Tirrell DA (2007) Preparation of the functionalizable methionine surrogate azidohomoalanine via copper-catalyzed diazo transfer. Nat Protoc 2: 1879–1883. doi: 10.1038/nprot.2007.268.

54. Fernandez-Gonzalez M, Boutureira O, Bernardes GJL, Chalker JM, Young MA, et al. (2010) Site-selective chemoenzymatic construction of synthetic glycoproteins using endoglycosidases. Chem Science 1: 709–715. doi: 10.1039/c0sc00265h.

55. Mellon P, Parker V, Gluzman Y, Maniatis T (1981) Identification of DNA sequences required for transcription of the human α1-globin gene in a new SV40 host-vector system. Cell 27: 279–288. doi: 10.1016/0092-8674(81)90411-6.

56. Williams AF, Barclay AN, Clark SJ, Paterson DJ, Willis AC (1987) Similarities in sequences and cellular expression between rat CD2 and CD4 antigens. J Exp Med 165: 368–380. doi: 10.1084/jem.165.2.368.

57. Davis SJ, Davies EA, van der Merwe PA (1995) Mutational analysis of the epitopes recognized by anti-(rat CD2) and anti-(rat CD48) monoclonal antibodies. Biochem Soc Trans 23: 188–194. doi: 10.1042/bst0230188.

58. Brogna S, Wen J (2009) Nonsense-mediated mRNA decay (NMD) mechanisms. Nat Struct Mol Biol 16: 107–113. doi: 10.1038/nsmb.1550.

59. Margineantu DH, Brown RM, Brown GK, Marcus AH, Capaldi RA (2002) Heterogeneous distribution of pyruvate dehydrogenase in the matrix of mitochondria. Mitochondrion 1: 327–338. doi: 10.1016/s1567-7249(01)00033-2.

60. Gibbs RA, Weinstock GM, Metzker ML, Muzny DM, Sodergren EJ, et al. (2004) Genome sequence of the Brown Norway rat yields insights into mammalian evolution. Nature 428: 493–521. doi: 10.1038/nature02426.

61. Pestka S (1971) Inhibitors of ribosome functions. Annu Rev Microbiol 25: 487–562. doi: 10.1146/annurev.mi.25.100171.002415.

62. Katunin VI, Muth GW, Strobel SA, Wintermeyer W, Rodnina MV (2002) Important contribution to catalysis of peptide bond formation by a single ionizing group within the ribosome. Mol Cell 10: 339-346. doi: 10.1016/s1097-2765(02)00566-x.

63. Dieterich DC, Link AJ, Graumann J, Tirrell DA, Schuman EM (2006) Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). Proc Nat Acad Sci USA 103: 9482–9487. doi: 10.1073/pnas.0601637103.

64. Williamson AR, Schweet R (1965) Role of the genetic message in polyribosome function. J Mol Biol 11: 358–372. doi: 10.1016/s0022-2836(65)80063-8.

65. Stevens A, Maupin MK (1989) 5,6-Dichloro-1-beta-D-ribofuranosylbenzimidazole inhibits a HeLa protein kinase that phosphorylates an RNA polymerase II-derived peptide. Biochem Biophys Res Commun 159: 508–515. doi: 10.1016/0006-291x(89)90022-3.

Legends to Supplementary Figures

Figure S1. Aha incorporation; some controls.

HeLa or HUVECs were grown ± Met for 15 min, pulsed ± 2 mM Aha for 2 min, and in some cases chased with 0.2 mM Met (without Aha) at 37˚C or 4˚C for the times indicated; after fixation, Alexa555 was 'clicked' on to incorporated Aha, DNA counterstained with DAPI, images collected of Alexa 555 fluorescence using a wide-field microscope (typical views are shown), and intensities in the cytoplasm (*cyto*) and nucleus (*nuc*) measured. Bars: 10 μm.

(A) Aha incorporation by unstarved cells (using conditions exactly as in Fig. 1A, except that cells were not initially grown without Met). Panels (iii) and (iv) show the only results presented where background seen in the absence of Aha in the nucleus or cytoplasm was not subtracted. Alexa 555 fluorescence in the cytoplasm and nucleus is expressed (± SD) relative to the value found in the Aha-pulsed cytoplasm. *: $P < 0.0001$ (Student's two-tailed t test, $n = 20$ cells).

(i) Omission of Aha yields background fluorescence.

(ii) Aha yields both cytoplasmic and nuclear signal (the latter is the strongest).

(iii) Signal is highest in nuclei, and nuclear signal is 3-fold higher than the (nuclear) background (– Aha).

(iv) Omission of Aha yields background fluorescence.

(v) Aha yields faint nuclear signal.

(vi) A faint – but nonetheless – significant signal (+ Aha) above background (– Aha) is seen in nuclei.

Some additional controls include:

(i) Anisomycin. In Figure 1A, 100 μg/ml anisomycin was added for 2 h to inhibit translation – conditions which are used routinely (e.g., [21]); we also use shorter times in critical experiments (e.g., 15 min in Fig. 1D).

(ii) Other inhibitors. Puromycin and cycloheximide give similar results to anisomycin, but neither inhibit Aha incorporation more quickly and/or to a greater extent than anisomycin, in accord with published data [61]. For example, after a 10-min pre-incubation in 2 mM puromycin, and a 20-s Aha pulse, the relative intensity of signal in nucleus and cytoplasm was reduced to 37 ± 2% and 28 ± 2%, respectively. It is unsurprising that inhibition is incomplete because the apparent K_m of puromycin for the bacterial ribosome is ~3 mM [62]. Note also that cycloheximide inhibits Aha incorporation into peptides [63], and its effects on Aha incorporation are discussed in the brochure provided by Invitrogen (who supplied some of the Aha and reagents used for 'click' chemistry).

(iii) Non-specific binding to RNA and DNA. In an experiment like that in Figure S1D (involving a 5-s Aha pulse), treatment with 5U RiboShredder (Epicentre) – an RNase cocktail – plus 3% BSA (Sigma) in PBS after fixation and immediately prior to 'clicking' slightly increases the relative intensity in nucleus and cytoplasm to 189 ± 26% and 139 ± 22%; similarly, treatment (30 min; 37°C) with 4.7 units DNase I (Worthington) and 0.5 mM $CaCl_2$ plus 3% BSA in PBS gave relative intensities of 177 ± 30% and 97 ± 17% (not shown). This is consistent with Aha not being attached inappropriately to RNA or DNA, and with the removal of nucleic acids increasing access of Alexa555 to Aha-labeled peptides during the 'click' reaction.

(B) Both nuclear and cytoplasmic signals disappear within minutes during a chase at 37°C. HeLa cells were starved of Met (15 min), pulsed for 2 min with 2 mM Aha, and chased for 0 or 5 min with 0.2 mM Met (without Aha) at 37˚C or 4˚C.

(i-iii) A 2-min Aha pulse labels both nucleus and cytoplasm, but essentially all signal disappears during a 5-min chase with Met at 37˚C – but not at 4˚C. Panel (ii) is the same as that in Figure 1Av.

(iv) After subtracting background, intensities (± SD) seen are expressed relative to the value found without a chase in cytoplasm. *: $P < 0.0001$ (Student's two-tailed *t* test, $n = 20$ cells).

(C) Some longer-lived peptides are labeled with longer Aha pulses. HUVECs were pulsed (without prior Met starvation) for 2-60 min with 2 mM Aha, and chased for 0 or 5 min with 0.2 mM Met (without Aha) at 37˚C.

(i) After a 2-min Aha pulse (without a chase), the nucleus appears the brightest (signal in both compartments is low, as there was no prior Met starvation).

(ii) With a 2-min pulse and 5-min chase, essentially all signal disappears.

(iii,iv) As pulse length increases (and chase length remains constant), nuclear signal increases.

(v) After subtracting background, intensities (± SD) seen are expressed relative to the value found without a chase in the cytoplasm. *: $P < 0.0004$ (Student's two-tailed *t* test, $n = 20$ cells). Some signal is seen in both nucleus and cytoplasm after the longest pulse, consistent with this signal marking the stable cellular proteome.

(D) During pulses of 5-15 s, Aha is incorporated into both nucleus and cytoplasm. HeLa cells were starved of Met for 30 min, and pulsed with 2 mM Aha for 0-15 s. After subtracting background, intensities (± SD) seen with Aha are expressed relative to the value found in the cytoplasm after a 15-s pulse. Signals in both compartments increase. *: $P < 0.0001$ (Student's two-tailed *t* test, $n = 20$ cells). Adding 100 µg/ml anisomycin 15 min before a 5-s pulse reduces the relative intensity in nuclei by 54%, and in the cytoplasm by 85% ($P < 0.0001$, Student's two-tailed *t* test, $n = 20$ cells).

Figure S1

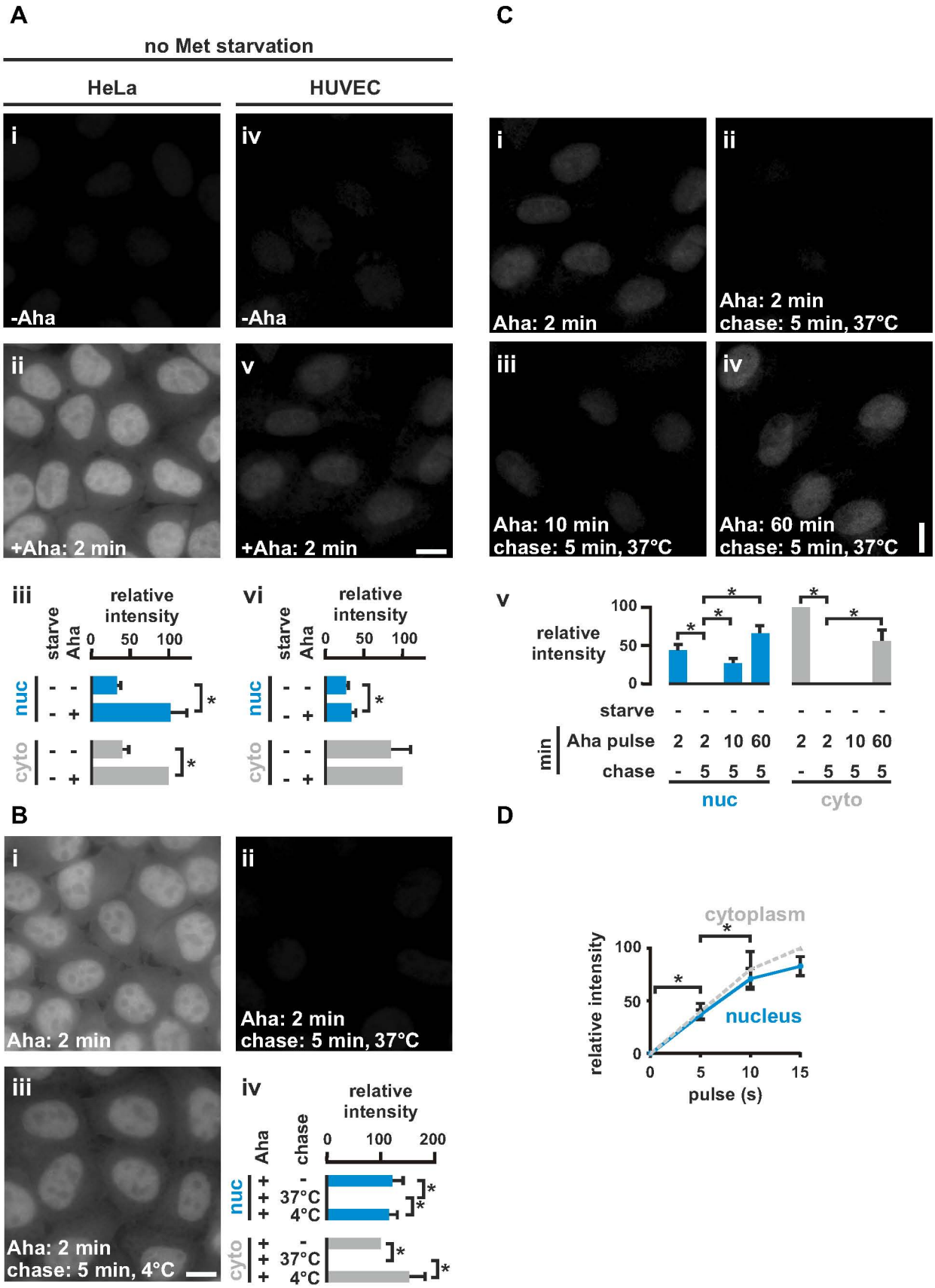

A
no Met starvation
HeLa
HUVEC
i
-Aha
iv
-Aha
ii
+Aha: 2 min
v
+Aha: 2 min
iii
relative intensity
starve
Aha
0 50 100
nuc
cyto
vi
relative intensity
starve
Aha
0 50 100
nuc
cyto
B
i
Aha: 2 min
ii
Aha: 2 min
chase: 5 min, 37°C
iii
Aha: 2 min
chase: 5 min, 4°C
iv
relative intensity
Aha
chase
0 100 200
nuc
37°C
4°C
cyto
37°C
4°C
C
i
Aha: 2 min
ii
Aha: 2 min
chase: 5 min, 37°C
iii
Aha: 10 min
chase: 5 min, 37°C
iv
Aha: 60 min
chase: 5 min, 37°C
v
relative intensity
100
50
0
starve
min
Aha pulse 2 2 10 60 2 2 10 60
chase - 5 5 5 - 5 5 5
nuc
cyto
D
cytoplasm
nucleus
relative intensity
100
50
0
0 5 10 15
pulse (s)

Figure S2. Comparison of images obtained using wide-field and confocal microscopes.

(A) Comparison of images (Alexa 555 fluorescence) obtained using wide-field and confocal microscopes. Bar: 10 μm.

(i,ii) Labeling as in Figure 1Dii. The single confocal section through the centre of the nucleus yields a punctate signal.

(iii,iv) Labeling as in Figure 2B; to allow comparison, panel iv is the same as that in Figure 2B. Again, the single confocal section yields a more punctate signal. Puromycin: puro.

Figure S2

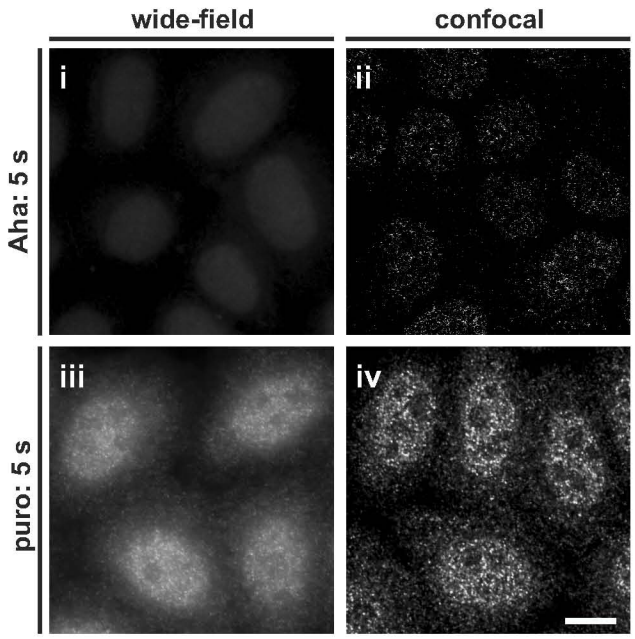


Figure S3. Puromycin incorporation: some controls. Bars: 10 μm.

Cells were pre-treated with 100 μg/ml cycloheximide (chx) for 15 min (to slow ribosomes), and pulsed ± 91 μM puromycin (puro) for various times; in some cases, cells were pretreated with 100 μg/ml anisomycin (aniso) for 2 h before fixation, and the pulse was followed by a chase at 37˚C or 4˚C. After fixation, puromycylated peptides were indirectly immuno-labeled with Cy3, DNA stained with DAPI, and images of Cy3 fluorescence collected using a confocal (all images shown are typical single confocal sections through the centre of nuclei) or wide-field microscope (used to determine relative intensities).

(A) The changing distributions of puromycylated peptides in HUVECs during pulse-chases.

(i,ii) After a 5-s pulse, puromycylated peptides are seen in foci in both cytoplasm and nucleus, with the latter being the brightest.

(iii) After a 5-s pulse followed by a 60-s chase at 37˚C, the peri-nuclear region now contains the brightest foci (presumably because puromycylated peptides made in both the nucleus and cytoplasm accumulate in the SER, and because some unincorporated puromycin remains within the cell during the chase to became incorporated).

(iv) After a 5-s pulse followed by a 60-s chase at 4˚C, the pattern remains more like that seen in (ii). In other words, the nuclear signal is brighter (which we attribute to the low temperature reducing translation and intra-cellular transport).

(v) After subtracting background, Cy3 intensities (± SD; $n = 20$ cells) are expressed relative to the value found in the cytoplasm after a 5-s pulse. Both nuclear and cytoplasmic signals increase after a 60-s chase at 37˚C; we attribute this to incomplete removal of all cellular puromycin on transfer from the puromycin-containing medium to the puromycin-free medium used for the chase. After a chase at 4˚C, the puromycin signal in both nucleus and cytoplasm increases only slightly. *: $P < 0.0001$ (Student's two-tailed t test).

(B) Puromycin incorporation by HeLa. David et al. [5] obtained a diffuse nuclear signal (with some nucleolar labeling) and brighter perinuclear signal after performing analogous experiments to those in Figure 2 with two differences; they permeabilized cells before adding puromycin, and pulsed for 5 min. Therefore, we repeated our experiment (where we only permeabilize after fixation) using a 5-min pulse.

(i) With no puromycin pulse, background levels of signal are seen.

(ii) A 5-min pulse of puromycin yields significant peri-nuclear labeling; this is like David et al. [5]. Unlike David et al. [5], there is little nucleolar labeling (which could be due to their permeabilization allowing some mRNAs to enter nucleoli to be translated by the high concentration of ribosomes there).

(iii) Pretreatment with anisomycin reduces both cytoplasmic and nuclear signal.

(iv) After subtracting background, Cy3 intensities (± SD) are expressed relative to the value found in the untreated (– anisomycin) cytoplasm; anisomycin reduces significantly the relative intensity in both nucleus and cytoplasm (*: $P < 0.0001$, Student's two-tailed t test, $n = 20$ cells).

(C) Effects of pre-treatment with cycloheximide (chx) on the distribution of puromycylated peptides in HeLa.

(i,ii) Without pre-treatment with cycloheximide, puromycylated peptides are seen in foci in both cytoplasm and nucleus, with the former yielding the brightest foci; with pre-treatment, the nucleus yields the brightest foci. This reversal is simply explained as follows. We assume a 5-s Aha pulse (where nuclear signal appears the brightest; Fig. 1Dii) gives the truest indication of the fraction of translation occurring in nuclei during a 5-s pulse. [This assumption seems appropriate as Aha-labeled peptides can only detach from the ribosome once the whole protein has been completed, whereas incomplete proteins that become end-labeled with puromycin are known to detach rapidly [64] and become concentrated at exit sites in the SER [5], [27].] In the absence of cycloheximide, some puromycylated peptides made in both the nucleus and cytoplasm become

concentrated in the SER; then, some labeled peptides will move from nucleus to cytoplasm to increase the cytoplasmic signal. However, in the presence of cycloheximide, ribosomes are slowed equally in both compartments, and more puromycylated peptides remain at the ribosome. As a result, the bias towards the cytoplasm is minimized (and the nucleus now appears brighter than the cytoplasm).

(iii) After subtracting background, Cy3 intensities in wide-field images are expressed relative to the value found in the cytoplasm after a 5-s puromycin pulse without pre-incubation in cycloheximide. The nuclear and cytoplasmic signal is reduced by cycloheximide (*: $P < 0.0001$, Student's two-tailed *t* test, $n = 20$ cells).

(D) Pre-treating HeLa cells with a transcriptional inhibitor slightly reduces puromycylation (5-s pulse) in the nucleus.

(i) Confocal sections through the center of nuclei. (i) Puromycylated peptides are seen in foci in both cytoplasm and nucleus; the latter yields the brightest foci. (ii) 100 μg/ml anisomycin added 30 min before puromycin marginally reduces nuclear signal. (iii) 100 μM DRB (5,6-dichloro-1-β-D-ribofuranosylbenzimidazole; [65]), added 30 min before puromycin also slightly reduces nuclear signal, without affecting the cytoplasmic one.

(iv) Quantitation using wide-field images. After subtracting background, intensities (± SD) seen with puromycin are expressed relative to the value found in the untreated (–anisomycin/–DRB) cytoplasm. Due to integration over the larger area, most signal is again seen in the cytoplasm (despite nuclear foci being the brightest). Anisomycin appears to have less effect on the cytoplasm; we attribute this to cycloheximide reducing both nuclear and cytoplasmic translation by 40-50% (Fig. S3Ciii), coupled with labeled peptides made in the cytoplasm remaining in the cytoplasm whereas those made in the nucleus are transported to the cytoplasm. Adding DRB slightly reduces nuclear signal (without effect on the cytoplasmic one), consistent with some close coupling of nuclear translation with transcription. *: $P < 0.0001$ (Student's two-tailed *t* test,

n = 20 cells). If translation occurs only in the cytoplasm, the reduction in nuclear signal can only be explained by splicing, mRNA export, cytoplasmic translation, and (an extraordinary) import of puromycylated peptides into nuclei all occurring in 5 s – which seems unlikely.

Figure S3

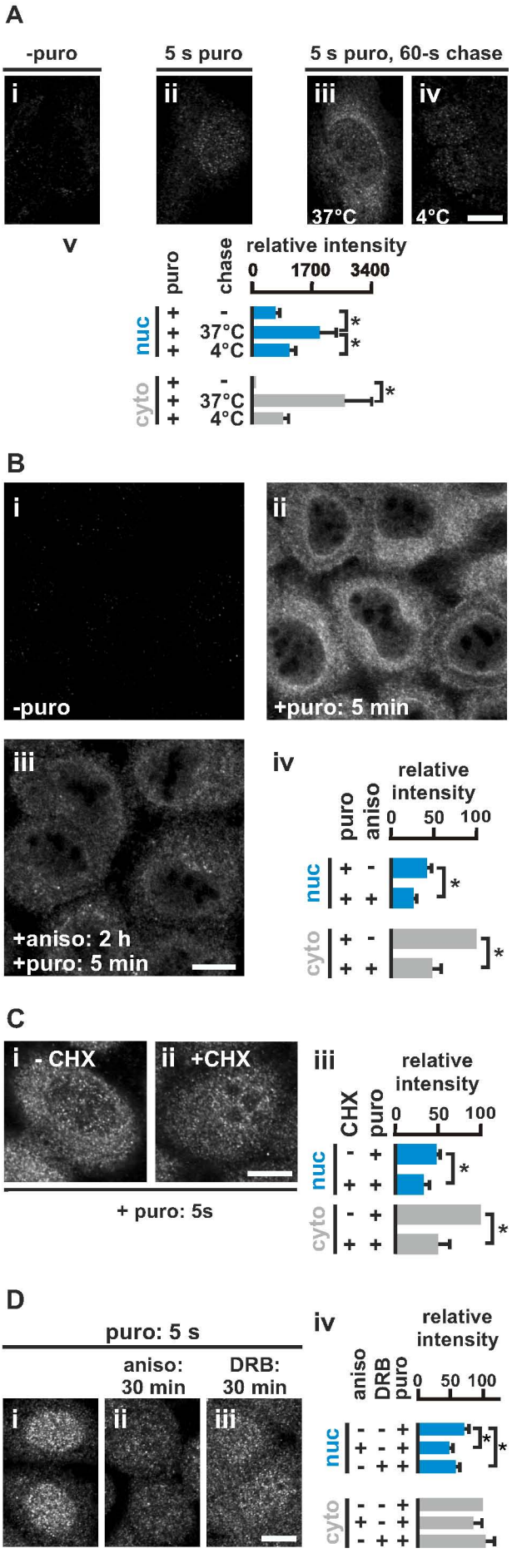

A
-puro
5 s puro
5 s puro, 60-s chase
i
ii
iii
iv
37°C
4°C
v
puro
chase
relative intensity
0 1700 3400
nuc
cyto
37°C
4°C
B
-puro
+puro: 5 min
+aniso: 2 h
+puro: 5 min
aniso
0 50 100
C
- CHX
+CHX
+ puro: 5s
CHX
D
puro: 5 s
aniso: 30 min
DRB: 30 min
DRB

Figure S4. CD2-EGFP expression constructs. Each construct encodes an SV40 *ori* (to permit replication) and a modified 'Tet' promoter inducible with doxycycline. In the base vector, the 'Tet' promoter drives expression of a protein containing an N-terminal mitochondrial signal sequence (mt), the rat CD2 epitope recognized by the OX34 antibody, a shortened intron 2 of *Cd2* (with intact splice donor and acceptor sites), a ribosome pause sequence (rps) which should slow a translating ribosome, and a C-terminal EGFP. Vector 1 encodes an EGFP sequence with the segment encoding the fluorochrome deleted (nfEGFP, to free the green channel for other use). Vector 2 encodes a haemagglutinin (HA) tag (not used in this study) inserted within the intron upstream of the first stop codon. Vector 3 also bears a PTC, 67 nucleotides upstream of the intron-exon junction (a position that should induce NMD; [58]).

Figure S4

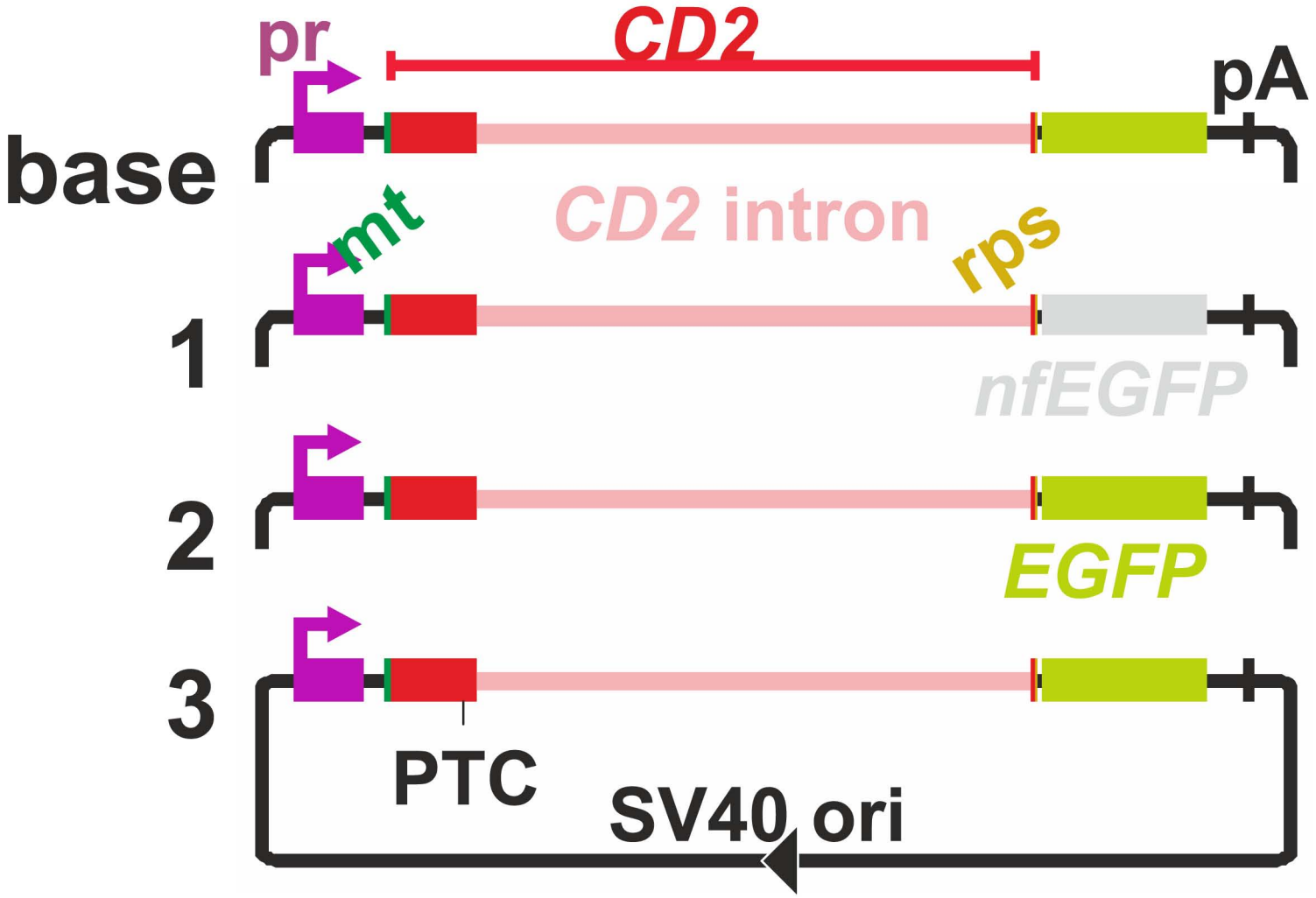


Table S1. Oligonucleotide DNA primers.

#	Name	Sequence (5′-3′)
1	Oligo pair 1	Fwd: GCAGA(CTGCAG)[b]AGAGACAGTGG Rev: CAGCATCTCCAGCCTTCTTC
2	Oligo pair 2	Fwd: ATCCACAGAAAGGTGCTCGT Rev: TA(GGATCC)[c]AT(TGTTCGTCGTCG)[d]CTTTGAGACCATCTCTAAAAG
3	Oligo pair 3	Fwd: AC(TCTAGA)[e]AAGACAGGCAGTCTCTG Rev: TA(GGATCC)[c]AT(TGTTCGTCGTCG)[d]CTTTGAGACCATCTCTAAAAG
4	Oligo pair 4[a]	Fwd: CACCCTCGTGACCACCTGCTTCAGCCGCTAC Rev: GTAGCGGCTGAAGCAGGTGGTCACGAGGGTG
5	Oligo pair 5[a]	Fwd: GGACGGGTGCTATG(TACCCTTACGACGTTCCTGATTACGCT)[f]TAGCAAGTTGGG Rev: CCCCAACTTGCTA(AGCGTAATCAGGAACGTCGTAAGGGTA)[f]CATAGCACCCGTCC
6	Oligo pair 6[a]	Fwd: GAGATGACAGTGGCACC(TAA)[g]AATGTAACGGTATACAGCAC Rev: GTGCTGTATACCGTTACATT(TTA)[g]GGTGCCACTGTCATCTC
7	Oligo 7	GCCATCATCCAACAGAGCAG
8	Oligo 8	CGTCGCCGTCCAGCTCGACCAG
9	Oligo 9	TGCCCTTCAGCTCGATGCGG

[a]: complimentary oligonucleotide primers for site-directed mutagenesis, HPLC purified

[b]: *PstI* recognition sequence

[c]: *BamHI* recognition sequence

[d]: ribosome pause sequence (rps)

[e]: *XbaI* recognition sequence

[f]: HA-epitope encoding sequence

[g]: premature termination codon (PTC)

Table S2. RNA FISH probes against rat *Cd2* intron-2[a].

#	**Sequence (5′-3′)**	#	**Sequence (5′-3′)**
1	GCACCCGTCCTAAAATGAAA	25	CTCCAAGGTTCTGAGTTCAT
2	TGTTACAGAAACGCTACTCC	26	TTACCAGAGTAAAGGCACCA
3	AACAAGAAACGACAGACAGG	27	AATGACTACATGAGAGGCCT
4	ACTGGAGTCTTCATTGTGAG	28	AGTGTAGGTGAAGTCAGTCA
5	AAGGAAAGGCAGACAGACTT	29	CACAAAGGATTCCCAAGTAC
6	AACCTGGGAGTCTTTACACA	30	ATGTCAGGCCAGAAGAAGAA
7	GCCGACCTGTTTTCTATCTT	31	CACCCGGGAATACAATTGTT
8	CAAGGATGTCCACCTTTATC	32	TACACACACACACACACACA
9	TGTCTACAACTTCATCAGCC	33	ATGTTGTAGATGCATGCGTG
10	TTAGCTCTCCAGACAAGAGA	34	TTTGCTGTTCTTCCAGAGGA
11	ATCTTGCCCTCTAACTCCTT	35	GTTGGCTGAACAGTTAAGAG
12	AGAGACTGCCTGTCTTTCTA	36	TAAGTACGACAAGTCAGAGC
13	ACTTGCCATCATCCAACAGA	37	AATGCCTAGCTTTGGGGTAA
14	GTCCTGAGTTTCTGTGTATC	38	CGTCCAAAGCTAATTGACCT
15	GCATGCAGAAATGCATTTCC	39	GTCTTCCCTTTTGCAGATGA
16	CATTTCTGCCTCCAACATAC	40	TCTAACAGCAAGCCTTCTGT
17	ATAGGCCTCCTTGGTGTAAT	41	TAAGCTGTGTCTCTACTGTG
18	TCAGATCACAGTGTCTTTCC	42	CAGGATGCGAGTAATATAGG
19	TTTTCCGATTTCCCCTCTCT	43	GACCTAAAGCATCTTGAGCT
20	TCAGTGGTCTCCATTCATCA	44	CCATCGAAGCTCTTTTGAAG
21	TTATTGAGATCAGGTCCAGG	45	GGACAAGAATCCTACCAACT
22	TCACACTGCAAATTCCACAC	46	AAACCCAGGTTTTCCTGCAT
23	ATTCCCTACCTGTCTCAAAC	47	TTTGCCAAGCTTCATGTGGT
24	GACTTGAAGGACCTCAACTT	48	TATTCTGACTCTCCCTCTAC

[a]: probes against the +ve strand.