Nuclear Translation

Sabyasachi Baboo

Lincoln College

The Sir William Dunn School of Pathology

Medical Science Division

University of Oxford

Thesis submitted for the degree of

Doctor of Philosophy

Michaelmas 2012

Abstract

Thesis title: **Nuclear Translation**

Candidate's name: **Sabyasachi Baboo**

Candidate's college: **Lincoln College**

Thesis submitted for the degree: **Doctor of Philosophy**

Term and year of submission: **Michaelmas 2012**

In bacteria, protein synthesis can occur tightly coupled to transcription. In eukaryotes, it is believed that translation occurs solely in the cytoplasm; I test whether some occurs in nuclei and find: (1) L-azidohomoalanine (Aha) – a methionine analogue (detected by microscopy after attaching a fluorescent tag using 'click' chemistry) – is incorporated within 5 s into nuclei in a process sensitive to the translation inhibitor, anisomycin. (2) Puromycin – another inhibitor that end-labels nascent peptides (detected by immuno-fluorescence) – is similarly incorporated in a manner sensitive to a transcriptional inhibitor. (3) CD2 – a non-nuclear protein – is found in nuclei close to the nascent RNA that encodes it (detected by combining indirect immuno-labelling with RNA fluorescence *in situ* hybridization using intronic probes); faulty (nascent) RNA is destroyed by a quality-control mechanism sensitive to translational inhibitors. I conclude that substantial translation occurs in the nucleus, with some being closely coupled to transcription and the associated proof-reading. Moreover, most peptides made in both the nucleus and cytoplasm are degraded soon after they are made with half-lives of about one minute.

I also collaborated on two additional projects: the purification of mega-complexes (transcription 'factories') containing RNA polymerases I, II, or III (I used immuno-fluorescence to confirm that each contained the expected constituents), and the

demonstration that some ‘factories’ specialize in transcribing genes responding to tumour necrosis factor α – a cytokine that signals through NFκB (I used RNA fluorescence *in situ* hybridization coupled with immuno-labelling to show active NFκB is found in factories transcribing responsive genes).

Acknowledgements

‘Acknowledgements’ is the most humbling part of any work, as it reminds the author of all the help, support and good will that has been offered to him/her, which is indispensable for completing the work – I feel as humbled when I write this.

I am grateful to Peter Cook for making me feel that I could not have been more fortunate than to pursue my research work under his guidance – he has been a source of inspiration, maverick ideas and criticism. He has worked equally hard to see this work to its conclusion and beyond.

I wish to express my gratitude to Argyris Papantonis and Joshua Larkin for the stimulating discussions on science and otherwise, for unreservedly sharing their expertise and suggestions, and responding to my unending queries. I thank Jon Bartlett for being a patient audience to my long discourses, a friend for discussing trivia, and making life simple and ‘socially active’ in the lab. I also extend my thanks to Binwei Deng, Svitlana Melnik, Kieran Finan, Anna Koeferle, James Halstead and Patrick Short for the good times we shared in the ‘Cook’ lab.

I am thankful to David Vaux in capacity of my ‘graduate advisor’ for critically discussing my work. I also thank Jordan Raff, Richard Jackson, Louis Mahadevan, Oreste Acuto and Steven Kelly for discussions. I owe a debt of gratitude to Haibo Jiang, Chris Grovenor and Errin Johnson for help with the experiments which could not be a part of this thesis. I appreciate the help of Alan Wainman and Mike Shaw for the up-keeping of microscopes.

It is my pleasure to thank Bhaskar Bhushan, Benjamin Davis, Philippe Pierre, Neil Barclay, Garry Brown, Minoru Yoshida, and Eric Cundliffe for generously offering various reagents for my experiments.

I appreciate the support extended by Lincoln College on the social front. My special thanks go to Lucinda Risius and Carmella Elan-Gaston for simplifying the formalities associated with my studies.

I was financially supported by the Felix Scholarship Trust and The Sir William Dunn School of Pathology through the course of my studies, and I sincerely acknowledge their kind support. The 'Dunn School', with its gregarious scientific community, has contributed immensely to my academic pursuit.

I am fortunate to have the company of Navneet, Prabhat, Pawan, Siddharth, Arpit, Chetak, Amit, Shweta, Masum, Anshuk, and Sahil, who maintained the diversity of 'Indian-ness' around me away from home.

I would rather not express my gratitude to my parents, as words can never do justice to it – to them I dedicate this work.

Contents

List of Tables

List of Figures

List of Abbreviations

Ab	antibody
Aha	L-azidohomoalanine
aniso	anisomycin
BODIPY	boron-dipyrromethene/4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene
BONCAT	bioorthogonal noncanonical amino acid tagging
BSA	bovine serum albumin
3C	chromatin conformation capture
CCF	cross-correlation function
CDF	cumulative distribution function
ChIA-PET	chromatin interaction analysis by paired-end tag sequencing
ChIP	chromatin immuno-precipitation
Chr.	chromosome
CHX	cycloheximide
CTD	C-terminal domain
CuAAC	Cu(I)-catalyzed azide-alkyne cyclo-addition
cyto	cytoplasm
DAPI	4′,6-Diamidino-2-phenylindole
DMEM	Dulbecco's modified Eagle medium
dox	doxycycline
DRB	5,6-dichlorobenzimidazole 1-β-D-ribofuranoside
DEPC	diethyl pyrrocarbonate
DMSO	dimethyl sulfoxide
DTT	DL-dithiothreitol
EDTA	ethylenediaminetetraacetic acid
EGFP	enhanced green fluorescent protein
EGM	endothelial growth medium
expt	experiment
FBS	fetal bovine serum
FISH	fluorescent *in situ* hybridization
freq	frequency
FUNCAT	fluorescent non-canonical amino-acid tagging
HA	haemagglutinin
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
Hpg	homopropargylglycine
HPLC	high-performance liquid chromatography

HUVEC	human umbilical vein endothelial cell
IgG	immunoglobulin G
K-S	Kolmogorov-Smirnov
MHC	major histocompatibility complex
NAS	nonsense-associated alternative splicing
nf	non-fluorescent
NLB	nuclear lysis buffer
NMD	nonsense-mediated mRNA decay
NN	nearest neighbour
nuc	nucleus
oligo	oligonucleotide
ORF	open reading frame
ori	origin of replication
pacta	pactamycin
PB	physiological buffer
PB^{+}	supplemented physiological buffer
PBS	phosphate buffered saline
PDF	probability density function
PLA	proximity ligation assay
PS	Penicillin/Streptomycin
PTC	premature termination codon (not to be confused with peptidyl transferase centre)
puro	puromycin
RCP	rolling circle product
RDP	rapidly degrading proteins
RER	rough endoplasmic reticulum
RNAP	RNA polymerase
SD	standard deviation
SDP	slowly degrading proteins
Ser	serine
SER	smooth endoplasmic reticulum
snRNP	small nuclear ribonucleoproteins
ss	single stranded
SSA	spliceostatin A
SSC	saline sodium citrate
S5^{P}	phosphorylated serine-5
SPAAC	strain promoted azide-alkyne cycloaddition

Tet	tetracycline
TNFα	tumor necrosis factor α
UTR	untranslated region

SI units and prefixes are used throughout, except the following cases.

min	minute
h	hour
S	Svedberg constant
Da	Dalton
bp	base pair
˚C	degree Celsius
ml or **µl**	millilitre or microlitre
Ω	ohm
M	molar

Chapter 1

Introduction

1.1 In the light of evolution

"*The numerous and fundamental differences between eukaryotic and prokaryotic organisms...have been fully recognized only in the past few years. In fact, this basic divergence in cellular structure which separates the bacteria and the blue green algae from all other cellular organisms, probably represents the greatest single evolutionary discontinuity to be found in the present day living world.*" **Stanier, Adelberg and Douderoff, 1963** [1].

"*Eukaryotes can be defined only by the ubiquity of nuclei...The argument that the nucleus was acquired as a symbiont makes no sense to me because, by itself, the nucleus falls far short of being an autopoietic system: it lacks any wherewithal for protein synthesis*" **Margulis, 1993** [2].

"*The origin of the eukaryotic nucleus marked a seminal evolutionary transition. We propose that the nuclear envelope's incipient function was to allow mRNA splicing, which is slow, to go to completion so that translation, which is fast, would occur only on mRNA with intact reading frames. The rapid, fortuitous spread of introns following the origin of mitochondria is adduced as the selective pressure that forged nucleus–cytosol compartmentalization.*" **Martin and Koonin, 2006** [3].

The above perspectives reflect the structural, functional and evolutionary significance of the nucleus in how we divide up the living world. Our classification of all cellular life-forms into 'prokaryotes' (from Gk. *pro* "before" + *karuon* "kernel, nucleus") or 'eukaryotes' (from Gk. *eu* "well" formed + *karuon* = "kernel, nucleus") depends on the presence or absence of a well-defined nucleus. There have been debates about whether such a classification has any phylogenetic basis, and a recent one even questions the concept of a 'prokaryote':

> "*There is not a single (monophyletic) phylogenetic group upon which to hang the tag prokaryote. The major eukaryotic organelles, mitochondria and chloroplasts, are definitely bacterial in origin, but the nucleus is not. The nuclear line of descent is as ancient as the archaeal line and not derived from either archaea or bacteria. Thus, the prokaryote/eukaryote model for biological diversity and evolution is invalid.*" **Pace, 2006** [4].

> "*Prokaryotes are cells with co-transcriptional translation on their main chromosomes; they translate nascent messenger RNAs into protein. The presence of this character distinguishes them from cells that possess a nucleus and do not translate nascent transcripts on their main chromosomes.*" **Martin and Koonin, 2006** [5].

It is evident from the above statements [2,3,5], that there exists a well-accepted perspective claiming transcription and translation are coupled in all eubacteria (including cyanobacteria) and archaea, and that this is the single-most important determinant of 'prokaryotic' character. In contrast, a well-defined nuclear membrane segregates the 'translationally-active' cytoplasm from the 'translationally-inactive' nucleus. The reason

cited for the lack of translating ribosomes in the nucleus is the massive invasion by 'introns' into the eukaryotic genome, which interrupt protein-coding DNA sequences – to divide them into 'coding' exons and 'non-coding' introns. An RNA transcribed from such a gene will contain introns with multiple stop codons, and this would be translated into a truncated protein that might be hazardous to the cell. Thus, Martin and Koonin [3] suggested that during the evolution of eukaryotes from prokaryotes (eukaryosis), the α-proteobacterial endosymbiont (which is supposed to have evolved into mitochondria [6]) contributed to the invading of group II introns from its genome to the host-cell genome, thus creating selective pressure on the host to generate the nuclear envelope – to protect the 'intron-ridden' host genome from translating ribosomes. Here, Martin and Koonin reflect the ideas of Cavalier-Smith [7] and Doolittle [8], and cite the work of Audibert *et al*. [9]; they conclude that the sluggish rate of splicing relative to the fast pace of translation creates the possibility of intron translation, and hence production of potentially unwanted 'truncated proteins'.

However, since the overall splicing rate includes the rate of transcription through an intron, the rate of intron removal from pre-mRNA by splicing, and the rate of degradation of the spliced-out intronic-RNA, Audibert *et al*. [9] calculated the rate of RNA degradation and excluded it from the overall rate of splicing. Thus, their definition of 'rate of splicing' includes the time needed to transcribe through introns, and so depends on intron lengths [10]. But it has been reported that the overall rate of splicing is independent of intron length [11], that splicing is co-transcriptional (i.e., splicing is completed before the entire gene is transcribed) [12], and splicing (*sensu stricto* intron removal from pre-mRNA) takes only 15-30 s to complete [13]. Considering the above arguments, I may assume that splicing in eukaryotes is a quick process, and may be fast enough to keep up with transcribing RNA polymerases and translating ribosomes.

Further, it is known that self-splicing introns also interrupt protein-coding genes in prokaryotes [14], and that ribosomes regulate the splicing of these introns by translating into them [15,16]. Again, translating ribosomes regulate transcription in bacteria [17], and one of the mechanisms known is dependent on folding of the transcript [18,19] in a way similar to the regulatory mechanism for splicing [15,16]. This coupling of splicing with transcription and translation in prokaryotes questions Martin and Koonin's hypothesis of 'evolutionary nucleus-cytosol compartmentalization during eukaryosis to segregate splicing and translation'. Rather, the invasion of group II introns from the endosymbiont α-proteobacteria into the 'already-formed' nucleus of an eukaryotic host might be responsible for the evolution of the 'more efficient' modern 'spliceosomal' splicing [20].

If the invasion of introns did not stimulate a prokaryote to evolve into a eukaryote, then what did? The endosymbiotic theory of eukaryosis proposed by Margulis [21,22], which has been refined over the years [23,24], claims that the spirochetes (eubacteria) formed a symbiotic relation with *Thermoplasma*-like (archaebacteria) hosts to increase the survivability of both symbionts. After incorporating the work of Cleveland [25] and Kirby [26] on extant protists, she proposed that the 'karyomastigont' [27] structure would have been the 'primitive nucleus' resulting from a eubacteria-archaebacteria symbiosis. Alternatively, Thomas Cavalier-Smith proposes the 'phagotrophic origin of [a] nucleus' [28], which posits that phagocytosis in uni-membrane 'Posibacteria' led to the formation of endomembranes (with attached DNA), and hence a nuclear envelope. There have been other theories for the evolution of eukaryotic nuclei [29,30] (or evolution of 80S ribosomes which could be associated with evolution of nuclei [31,32]) but there is yet to be a consensus on this debate.

Also, in spite of a debate on the endosymbiotic origin of the nucleus [33,34], there are strong arguments to support the idea that the nucleus did not originate as an endosymbiont [35,36,28]. Again, Koonin agrees that the origin of eukaryotes can be explained independently of his 'intron invasion' theory [37]. Further, it has been reported that *trans*-splicing (splicing of two discrete transcripts, as compared to *cis*-splicing which refers to the splicing of two parts of the same transcript) occurs in both prokaryotes and eukaryotes [38], even in the absence of introns (e.g., in eukaryotes such as Trypanosomes [39] which have a well-formed nucleus but hardly any introns [40]). This brings in to question whether splicing of newly-gained intra-genic introns could be a trigger to evolve a nucleus even when Trypanosome nuclei harbor almost no introns (however, lack of introns in the genome could be attributed to evolutionary loss of introns due to genome compaction, though there is not much evidence for this [41]). Thus it seems most likely that eukaryosis was not triggered to protect the 'intron-ridden' archaeal host-genome from translating ribosomes. The discovery of a nucleus-like structure (a double membrane enclosing the bacterial 'nucleoid') in bacteria from phyla like Planctomycetes, Verrucomicrobia, and Chlamydiae (most distinct in *Gemmata obscuriglobus*), and the presence of ribosome-like structures in this nucleus, strengthen this notion [42,43].

Rewinding back to the debate on usage of the term 'prokaryote', against the usual belief, the original use of the classes 'prokaryotic' (Fr. *procariotíque*) and 'eukaryotic' (Fr. *eucariotíque*) by Edouard Chatton in 1925 was based on the evolution of 'flagella-centriole-mitotic spindle' structure, rather than signifying presence or absence of the nucleus [44,45,24].

All of the above evidence strongly supports the notion that coupling of transcription, splicing, and translation in prokaryotes is arguably discrete from the reason

for the ‘origin of the eukaryotic nucleus’. This coupling exists in prokaryotes, and there seems no apparent molecular, functional, or evolutionary reason for it not to occur in eukaryotes. Thus, the absence of ‘active translation’ in a ‘eukaryotic nucleus’ is an assumption, and should be experimentally verified.

However, addressing whether some ‘nuclear translation’ occurs is difficult if only because an inability to detect translation in a nucleus is not a proof of its absence. The problem is compounded by the structural continuity of the nuclear membrane and endoplasmic reticulum (a prominent site of active translation by cytoplasmic ribosomes) which makes it difficult to be sure that an isolated nucleus is free of contaminating cytoplasmic ribosomes.

1.2 Historical perspectives

The first (indirect) report indicating that some protein synthesis might occur in the nucleus was made by Alfred Mirsky’s group (1952) [46]; labelled amino acids were incorporated into nuclear proteins. This was made at a time when it was yet to be confirmed that (i) DNA is the genetic material, (ii) an intermediate molecule – RNA is involved in expression of inherited traits, and (iii) there are specific organelles – ribosomes – which act as machines incorporating amino acids to produce proteins. In short, the ‘central dogma of molecular biology’ was yet to be proposed by Francis Crick (1958) [47]. In 1953, Mirsky’s group [48] reported that it is the microsomal fraction (later discovered to be the ribosome-rich fraction associated with the endoplasmic reticulum) – obtained by ‘differential centrifugation’ [49] of the cytoplasm – that is the ‘protein-synthesis factory’ of the cell. This was just after the very first experiments (in 1950) which demonstrated that amino acids are incorporated by microsomes *in vivo* [50,51], and well before mRNA was discovered as the substrate for protein synthesis (in 1961) [52].

In 1954, Vincent Allfrey made the first claim of nuclear protein synthesis when he reported incorporation of labelled amino acids in isolated nuclei from the thymus [53]. The period from 1954-1974 was replete with publications discussing evidence mostly 'for' and some 'against' protein synthesis in nuclei – most of them cited in reviews by Lester Goldstein (1970) [54] and LeRoy Kuehl (1974) [55] (for additional ones, see [56-61]). Most of these reports included experiments evaluating: (i) the incorporation of radio-labelled amino acids by isolated nuclei and/or into nuclear proteins, (ii) the incorporation of radio-labelled amino acids in whole tissues/cells followed by autoradiography and the counting of silver grains lying over nuclei, and (iii) the isolation of ribosomes from nuclei. This culminated in a debate between Goidl and Allen in 1978 [62], where both agreed that nuclear protein synthesis was a possibility, and that the central issue to be resolved was the possible contamination of nuclear fractions or isolated nuclei by cytoplasmic ribosomes – and this continues to be the central issue. Subsequently, reports of the synthesis of nuclear proteins in the cytoplasm followed by transport into nuclei [63,64] fuelled the idea that all translation might occur in the cytoplasm. Then, the discovery of 'introns' in 1977 [65] was probably responsible for ending the debate for twenty years; it raised the apprehension that – if nuclear ribosomes were to translate introns with their many termination codons – too many unwanted 'truncated' peptides would be formed. However, during these twenty years, independent studies on the effects of incorrectly-placed termination codons pointed to some translation occurring in the nucleus.

1.3 Nuclear NMD: strong circumstantial evidence for nuclear translation

In 1979 and 1984, it was discovered that β^0-thalassemia was caused by a mutation that changed a sense codon to a non-sense (stop/termination) codon in the β-globin gene, and this further led to a decrease in the number of β-globin mRNAs (without affecting the stability of the mature mRNA) [66-69]. This was the discovery of 'nonsense-mediated mRNA decay (NMD)' [70] – a phenomenon which causes selective decay of those immature transcripts that possess an unusual stop-codon or 'premature termination codon' (PTC) in their coding sequence (CDS). Later, in 1987 and 1993, it was discovered that constitutive exons could be skipped if they possessed a nonsense mutation – a phenomenon which came to be known as 'nonsense-associated alternative splicing' (NAS) [71,72].

The end of any open reading frame (ORF) is defined by a stop codon. As multiple stop codons are found in introns, mis-splicing will often introduce a PTC into an ORF [73,74], and this would lead to the production of truncated (and faulty) peptides that might be toxic to the cell. The process of NMD is a proof-reading mechanism that weeds out faulty transcripts bearing such PTCs [75,76]. It is experimentally recognised by expressing genes encoding messages with and without PTCs: levels of mRNA with PTCs are lower, presumably due to the action of NMD [77,78]. This raises an important question: where does NMD occur – in the nucleus or cytoplasm? The relative levels of transcripts bearing PTCs prove to be lower in both nuclear and cytoplasmic fractions, with the largest reduction being seen in the nucleus; this implies that NMD is a nuclear process [79,80]. Then, what mechanism recognises the PTC? As the only machinery known to recognise any termination codon is a translating ribosome, how is nuclear NMD

possible in the absence of translation in the nucleus? A similar kind of argument concerns NAS [81,82,83]. Here a PTC affects another nuclear event – alternate splicing. Once again, how does the system recognise a PTC in the nucleus? These questions are easily answered if an active nuclear ribosome recognises a PTC at the transcription site to trigger NMD and NAS [84]. The dependency of NMD [78,85] (and NAS [86,87]) on active translation has been shown by demonstrating that the levels of PTC-bearing transcripts stabilize when translation is inhibited (or exon-skipping is avoided in case of NAS).

Current models for NMD can be classified into four types (the first three assume that no translation can occur in the nucleus):

1. A PTC is recognised during cytoplasmic translation, and this generates a signal that is transmitted to the nucleus, where it induces degradation of homologous transcripts [88]. There is no experimental evidence to support this argument, and no one has proposed a plausible mechanism to explain how this signal might specifically target homologous transcripts in the nucleus.
2. The translating (cytoplasmic) ribosome that detects the PTC lies close to the nuclear pore, so that the resulting NMD appears to be nuclear because that ribosome (and associated message) co-purifies with the nucleus [80,89]. In other words, the 'pioneer round of translation [90]' occurs in the cytoplasm as the message exits the nuclear pore. However, introducing a PTC into a message reduces nuclear transcript levels to a third (or less) [91,92], so we would then expect to find (at least) two out of three nuclear transcripts close to pores; however, fluorescence *in situ* hybridization (FISH) reveals no such concentration of polyadenylated RNA or transcripts at the nuclear periphery [93,94].

3. Some unknown mechanism scans nascent transcripts for PTCs [95]. However, there is again no suggestion as to what this mechanism might be.
4. Nuclear ribosomes translate nascent transcripts, and – if they contain a PTC – the associated NMD machinery degrades the faulty transcripts. We believe this simple model is consistent with most of the data discussed here (i.e., **section 1.3**).

Further, if it is assumed that proofreading of nascent transcripts occurs in the nucleus, then sites of the 'pioneer round of translation' would most probably be a transcription factory [96-98], where transcription, splicing, and translation might be coupled. The idea that transcription might occur in factories where fixed RNA polymerases reel in chromatin has been controversial for the past few decades [99]. One major criticism against the existence of such factories has been the inability to isolate them biochemically [98]. We have recently succeeded in partially purifying these huge complexes (~10 MDa) – separating three kinds based on the three different eukaryotic RNA polymerases they contain [100].

1.4 Recent discussions on nuclear translation (1998-2012)

In 1998, it was reported that aminoacylation of tRNA occurs in *Xenopus* oocyte nuclei, and such charging of the tRNA increases its export efficiency [101]; this report also speculated that such aminoacylation may indicate monitoring of transcripts in nuclei by NMD. In 1999, it was suggested that translation must be coupled with transcription in *Dictyostelium discoideum* nuclei [102], where ribosomes isolated from the nuclear fraction were shown to be associated with immature transcripts (lacking a 3' UTR or untranslated region) as against cytoplasmic ribosomes which were associated with mature transcripts (containing a 3' UTR); it was further shown that transcripts containing PTCs

were retained in the nucleus and were associated only with nuclear ribosomes, thus implicating nuclear translation in NMD.

The coupling of translation to transcription provided the best evidence for nuclear translation [94,103]. [The methodology for detecting this coupling has been published in detail [104].] HeLa cells were permeabilised in a 'physiological' buffer, and allowed to extend nascent polypeptides by ≤15 residues in the presence of tagged precursors (e.g., biotin-lysine-tRNA or BODIPY-lysine-tRNA). Then, essentially all label is incorporated into 'nascent' polypeptide. [I will apply the term 'nascent' to peptides that are being synthesised at the ribosome.] Although most label was found in the cytoplasm, 9-15% was also seen in discrete nuclear foci (factories) that were the sites of transcription by RNA polymerase II. Functional coupling of transcription with translation was shown by activating (or repressing) transcription by adding the necessary NTPs (or inhibitors of transcription like α-amanitin and 3′ dATP). Moreover, components involved in translation (e.g., initiation/release factors, ribosomal subunits) co-localised (to within a few nanometres) and co-purified with nascent RNA and the active form of the polymerase.

Despite these seemingly convincing demonstrations supporting the idea that some translation occurred in nuclei, an influential paper claimed to have repeated these experiments but failed to obtain the same results [105]. However a close examination of the experimental conditions used revealed that they used a different buffer which supported substantially less transcription, and this could obviously reduce the amount of nuclear translation if that translation were coupled to transcription [106]. They also did not include protease inhibitors in their buffers; again this would be expected to reduce the amount of signal due to nuclear translation if faulty peptides were rapidly degraded by the proteasome [94,106-109]. This failure to use exactly the same conditions undermines the credibility of this evidence [105].

Other criticisms [88,110-112] of the evidence for nuclear translation include the following:

1. Contaminating cytoplasmic ribosomes are responsible for the nuclear translation signal seen. If this were so, the amount of nuclear translation should reflect the degree of contamination. However, a >20 fold variation in the number of contaminating cytoplasmic ribosomes does not vary the nuclear signal [94,106]. It is also possible that cytoplasmic ribosomes enter the nucleus during permeabilisation. However, after peremeabilising the cell, the nuclei were still intact enough to exclude dextrans smaller than ribosomes [94,106], and – even if cytoplasmic ribosomes did enter the nucleus – why should they concentrate at the transcription factories?
2. The nuclear membrane evolved to separate transcription and splicing from translation, and that no translation can be allowed to happen within nuclei. As we have seen (see **Section 1.1**), the argument here is that if these processes were not spatially separated then translation of intron-containing RNAs by nuclear ribosomes would generate faulty peptides [113]. However, these processes could still be spatially separated if the translating ribosome were positioned downstream of the spliceosome at the transcription site in the nucleus. Further, as mentioned in **section 1.1**, ribosomes have been shown to regulate splicing by translating into introns in prokaryotes [10-12].
3. The presence of nascent peptides in the nuclear fraction does not necessarily mean that the translating ribosomes lie within the nucleoplasm; they could be at the outer nuclear membrane so that they co-purify with nuclei. However, little label in nascent peptides was observed at the nuclear periphery [93,94].

4. The concentration of translation machinery in the nucleus is too low. Immuno-labelling shows that nuclei do not contain critical components of the translation machinery. However, we should not attach too much weight to negative results of this kind (as the machinery may be present below the level of detection). There is also a practical point: chromatin is present at a high concentration in the nucleus, and this could prevent access of immuno-labelling antibodies. Indeed, significant concentrations of the critical components can be seen after using milder fixation conditions [94,103].

There is also scepticism about the possibility of nuclear translation based on the failure to find functional 80S ribosomes within nuclei. It is suggested that the 40S ribosomal-subunit must possess 18S rRNA, and the 60S ribosomal-subunit must possess 5.8S rRNA, before the two can come together to form an active ribosome. As the maturation of 18S rRNA from 20S rRNA [114], and 5.8S rRNA from 6S rRNA [115], is reported to happen in the cytoplasm of *Saccharomyces cerevisiae*, it is argued that an 80S ribosome has to mature in the cytoplasm to be functional. But again, the failure to detect a process cannot be taken as proof that the process does not occur (because it may lie below the level of detection). Moreover, ribosomal subunits might mature in the cytoplasm to be imported into the nucleus. Further, it has even been reported that translating ribosomes may not necessarily require 20S to 18S rRNA maturation in *Saccharomyces cerevisiae* [116], and that immature rRNA may even be essential for the formation of active ribosomes [117].

Another argument against a role for an active nuclear ribosome runs as follows. We might presume that if nuclear translation scans for transcripts that contain PTCs (arising, perhaps, from faulty splicing), then the rate or order of intron removal might also be affected; however, no effects of PTCs on splicing were detected [118]. But we would

not expect any effects if the active nuclear ribosome scans the maturing transcript after it has been spliced. Even if one considers that translation is faster than splicing, and may cause ribosomes to venture into introns, then this would be applicable to all transcripts (whether bearing PTCs or not), as every transcript is expected to be scanned by nuclear ribosomes. So, presence of a nuclear scanning mechanism should not affect the splicing rate, irrespective of transcripts bearing PTCs.

I suggest that none of this evidence against nuclear translation is decisive. Moreover, in the years since the reports of coupled transcription and translation in eukaryotic nuclei [94,102] a wealth of data consistent with the idea that some translation occurs in nuclei has been reported. These findings have been further reviewed/evaluated against the background of NMD [119-122]. For example, ribosomes are assembled in nucleoli (which are the site of RNA polymerase I transcription), and so are naturally found in nuclei; however, many ribosomal components and translation factors are also found at RNA polymerase II transcription sites [94,103,123-125], sometimes associated with the NMD machinery [103] and/or proteasomes [107,108]. Significantly, translation initiation factors are found in the proteome of isolates containing large fragments of polymerase II factories, but not in the equivalent fragments of polymerase I and III factories [100]. In addition, precursor RNAs – if inefficiently spliced [126] or harboring nonsense codons [127,128] – accumulate near the site of transcription, and the initiation factors involved in translation are both nuclear and found at sites of splicing [94,103,129,130]. There was also a report of translation in the nucleus induced by an RNA editing protein [131]. However, due to the undue suspicions raised [105] on the reproducibility of data generated by Iborra *et. al.* [94], by about 2004 the debate on nuclear translation had again subsided without decisive evidence for or against its existence.

As a result of the scepticism concerning nuclear translation, there have been many instances in the past where results that are most simply interpreted in terms of nuclear translation are (i) ignored as artefacts [132-134], (ii) explained by cytoplasm-to-nucleus transport of highly-abundant and newly-formed nucleolar or nuclear proteins [135-138], or (iii) mentioned in the context of 'speculation' [139]. In contrast, Jonathan Yewdell and Jack Bennink's group has worked on 'rapidly turning-over' peptides presented by major histocompatibility complex class I (MHC-I) molecules for immuno-surveillance [140], and they suspect that such peptides are translated in the nucleus [141]. The extension of this hypothesis led them recently to use puromycylation of ribosome-bound nascent peptides to visualize nuclear translation [142], and this revived the debate on nuclear translation [143,144]. In this report, David *et. al.* [142] arrest translation by pre-incubating mammalian cells with cycloheximide or emetine, before allowing extension of the nascent peptide in puromycin; the incorporated puromycin in these polysome-arrested peptides is then indirectly immuno-detected *via* immuno-fluorescence.

1.5 Scope of this thesis

In this thesis, I have examined whether nuclear translation occurs. I use an analogue of methionine – L-azidohomoalanine (Aha) – to label nascent peptides. [This precursor is widely used to label proteins for mass spectrometry, and kits to facilitate such labelling are available commercially from Invitrogen.] I find that Aha is incorporated into the nucleus of living mammalian cells after incubations of only 5 s. I further see that this incorporation is inhibited by anisomycin (a translation inhibitor), and that Aha-tagged nascent nuclear peptides have short half-lives dependent on proteasomal activity. I also use puromycin (a translation inhibitor), which is effectively an amino acid analogue, and observe its incorporation into the nucleus of living mammalian cells after incubations of

only 5 s. I go on to find that translation of these nascent peptides is closely coupled to transcription. Also, using a 'multi-copy' and 'inducible' expression-system, I express a specific non-nuclear protein in large amounts for short periods, and find it in the nucleus, close to sites containing the nascent transcripts that encode it: introducing a PTC also reduces the amount of nascent RNA. Combining all these results, I believe that some translation occurs in the nucleus, coupled to transcription and splicing at transcription factories, where a 'pioneer round of translation' proofreads nascent transcripts – the first step in RNA quality control.

Further, we have used biochemical methods to partially purify transcription factories [100]. **Chapter 5** explains this novel purification method. I contributed to this work by validating the purification method using immuno-detection assays. Also, we find there is a division of labour among transcription factories, with some factories specializing in transcribing selected sets of genes [97,145,146]. Such a set of genes which are transcribed in a specialized transcription factory may share (i) a common RNA polymerase which transcribes them, (ii) a common transcription factor essential for their transcription, or (iii) a common cell-signalling network. We recently provided evidence for this hypothesis, by showing that transcription dependent on NFκB-mediated signalling occurs in special 'NFκB-factories' [147]. **Chapter 6** elaborates on this finding. I contributed to this work by combining RNA *in situ* fluorescence hybridization (RNA FISH) and immuno-fluorescence to show that active NFκB (in particular, its phosphorylated p65 subnit) is present in factories transcribing selected genes.

1.6 References

1. **Stanier RY, Adelberg E, Douderoff M.** *The microbial world*. 1963; Prentice-Hall, Englewood Cliffs, N.J.
2. **Margulis L.** *Symbiosis in cell evolution: Microbial communities in the archean and proterozoic eons*. 1993; W. H. Freeman, New York.:217-218.
3. **Martin W, Koonin EV.** Introns and the origin of nucleus–cytosol compartmentalization. *Nature*. 2006;**440**(7080):41-45.
4. **Pace NR.** Time for a change. *Nature*. 2006;**441**(7091):289.
5. **Martin W, Koonin EV.** A positive definition of prokaryotes. *Nature*. 2006;**442**(7105):868.
6. **Falah M, Gupta RS.** Cloning of the hsp70 (dnaK) genes from *Rhizobium meliloti* and *Pseudomonas cepacia*: phylogenetic analyses of mitochondrial origin based on a highly conserved protein sequence. *J Bacteriol*. 1994;**176**(24):7748-7753.
7. **Cavalier-Smith T.** Intron phylogeny: A new hypothesis. *Trends Genet.* 1991;**7**(5):145-148.
8. **Doolittle WF.** The origin of introns. *Curr. Biol.* 1991;**1**(3):145-146.
9. **Audibert A, Weil D, Dautry F.** *In vivo* kinetics of mRNA splicing and transport in mammalian cells. *Mol. Cell. Biol.* 2002;**22**(19):6706-6718.
10. **Swinburne IA, Silver PA.** Intron delays and transcriptional timing during development. *Dev Cell.* 2008;**14**(3):324-330.
11. **Singh J, Padgett RA.** Rates of in situ transcription and splicing in large human genes. *Nat Struct Mol Biol.* 2009;**16**(11):1128-1133.
12. **Tilgner H, Knowles DG, Johnson R, Davis CA, Chakrabortty S, Djebali S, Curado J, Snyder M, Gingeras TR, Guigó R.** Deep sequencing of subcellular

RNA fractions shows splicing to be predominantly co-transcriptional in the human genome but inefficient for lncRNAs. *Genome Res*. 2012;**22**(9):1616-1625.

13. **Huranová M, Ivani I, Benda A, Poser I, Brody Y, Hof M, Shav-Tal Y, Neugebauer KM, Stanek D.** The differential interaction of snRNPs with pre-mRNA reveals splicing kinetics in living cells. *J Cell Biol.* 2010;**191**(1):75-86.
14. **Woodson SA.** Ironing out the kinks: splicing and translation in bacteria. *Genes Dev.* 1998;**12**(9):1243-1247.
15. **Semrad K, Schroeder R.** A ribosomal function is necessary for efficient splicing of the T4 phage thymidylate synthase intron *in vivo*. *Genes Dev*. 1998;**12**(9):1327-1337.
16. **Sandegren L, Sjöberg BM.** Self-splicing of the bacteriophage T4 group I introns requires efficient translation of the pre-mRNA in vivo and correlates with the growth state of the infected bacterium. *J Bacteriol.* 2007;**189**(3):980-990.
17. **Proshkin S, Rahmouni AR, Mironov A, Nudler E.** Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science.* 2010;**328**(5977):504-508.
18. **Yanofsky C.** Attenuation in the control of expression of bacterial operons. *Nature.* 1981;**289**(5800):751-758.
19. **Richardson JP.** Preventing the synthesis of unused transcripts by Rho factor. *Cell.* 1991;**64**(6):1047-1049.
20. **Lambowitz AM, Zimmerly S.** Group II introns: mobile ribozymes that invade DNA. *Cold Spring Harb Perspect Biol.* 2011;**3**(8):a003616.
21. **Sagan L.** On the origin of mitosing cells. *J Theor Biol.* 1967;**14**(3):255-274.
22. **Margulis L.** *Origin of eukaryotic cells*. 1970; Yale University Press, New

Haven.

23. **Margulis L, Dolan MF, Whiteside JH.** “Imperfections and Oddities” in the origin of the nucleus. *Paleobiology.* 2005;**31**(2):175-186.

24. **Margulis L, Chapman M, Guerrero R, Hall J.** The last eukaryotic common ancestor (LECA): Acquisition of cytoskeletal motility from aerotolerant spirochetes in the proterozoic eon. *Proc Natl Acad Sci U S A.* 2006;**103**(35):13080-13085.

25. **Cleveland LR, Grimstone AV.** The fine structure of the flagellate *Mixotricha paradoxa* and its associated micro-organisms. *Proc. R. Soc. Lond. B.* 1964;**159**(977):668-686

26. **Kirby H, Margulis L.** Harold Kirby's symbionts of termites: karyomastigont reproduction and calonymphid taxonomy. *Symbiosis.* 1994;**16**(1):7-63.

27. **Janicki C.** Untersuchungen an parasitischen flagellaten. *Zeitschrift fur wissenschftliche zoologie*. 1915; **112**:573-691.

28. **Cavalier-Smith T.** Origin of the cell nucleus, mitosis and sex: roles of intracellular coevolution. *Biol Direct.* 2010;**5**(7).

29. **Pennisi E.** Evolutionary biology. The birth of the nucleus. *Science.* 2004;**305**(5685):766-768.

30. **Vesteg M, Krajčovič J.** Origin of eukaryotic cells as a symbiosis of parasitic α-proteobacteria in the periplasm of two-membrane-bounded sexual pre-karyotes. *Commun Integr Biol.* 2008;**1**(1):104-113.

31. **Jékely G.** Origin of the nucleus and Ran-dependent transport to safeguard ribosome biogenesis in a chimeric cell. *Biol Direct.* 2008;**3**(31).

32. **Barbieri M.** The ribotype theory on the origin of life. *J Theor Biol.* 1981;**91**(4):545-601.

33. **Lake JA, Rivera MC.** Was the nucleus the first endosymbiont? *Proc Natl Acad Sci U S A.* 1994;**91**(8):2880-2881.

34. **López-García P, Moreira D.** Selective forces for the origin of the eukaryotic nucleus. *Bioessays.* 2006;**28**(5):525-533.

35. **Poole AM.** Did group II intron proliferation in an endosymbiont-bearing archaeon create eukaryotes? *Biol Direct.* 2006;**1**(36).

36. **Poole AM, Neumann N.** Reconciling an archaeal origin of eukaryotes with engulfment: A biologically plausible update of the Eocyte hypothesis. *Res Microbiol.* 2011;**162**(1):71-76.

37. **Koonin EV.** Intron-dominated genomes of early ancestors of eukaryotes. *J Hered.* 2009;**100**(5):618-623.

38. **Fuerst JA, Webb RI.** Membrane-bounded nucleoid in the eubacterium *Gemmata obscuriglobus*. *Proc Natl Acad Sci U S A.* 1991;**88**(18):8184-8188.

39. **Fuerst JA, Sagulenko E.** Beyond the bacterium: planctomycetes challenge our concepts of microbial structure and function. *Nat Rev Microbiol.* 2011;**9**(6):403-413.

40. **Lasda EL, Blumenthal T.** Trans-splicing. *Wiley Interdiscip Rev RNA.* 2011;**2**(3):417-434.

41. **Perelman D, Boothroyd JC.** Lack of introns in the ribosomal protein gene S14 of trypanosomes. *Mol Cell Biol.* 1990;**10**(6):3284-3288.

42. **Mair G, Shi H, Li H, Djikeng A, Aviles HO, Bishop JR, Falcone FH, Gavrilescu C, Montgomery JL, Santori MI, Stern LS, Wang Z, Ullu E, Tschudi C.** A new twist in trypanosome RNA metabolism: cis-splicing of pre-mRNA. *RNA.* 2000;**6**(2):163-169.

43. **Penny D, Hoeppner MP, Poole AM, Jeffares DC.** An overview of the introns-

first theory. *J Mol Evol.* 2009;**69**(5):527-540.

44. **Chatton E.** *Pansporella perplexa,* réflexions sur la biologie et la phylogénie des Protozoaires. *Ann Sci Nat Biol Anim Zool.* 1925;10e série,**VII**:6-84.

45. **Sapp J.** The prokaryote-eukaryote dichotomy: Meanings and mythology. *Microbiol Mol Biol Rev.* 2005;**69**(2):292-305.

46. **Daly MM, Allfrey VG, Mirsky AE.** Uptake of glycine-N^{15} by components of cell nuclei. *J Gen Physiol.* 1952;**36**(2):173-179.

47. **Crick F.** Central dogma of molecular biology. *Nature.* 1970;**227**(5258):561-563.

48. **Allfrey V, Daly MM, Mirsky AE.** Synthesis of protein in the pancreas. II. The role of ribonucleoprotein in protein synthesis. *J Gen Physiol.* 1953;**37**(2):157-175.

49. **Claude A.** Fractionation of mammalian liver cells by differential centrifugation: II. Experimental procedures and results. *J Exp Med.* 1946;**84**(1):61-89.

50. **Hultin T.** Incorporation in vivo of ^{15}N-labeled glycine into liver fractions of newly hatched chicks. *Exp Cell Res.* 1950;**1**(3):376-381.

51. **Borsook H, Deasy CL, Haagensmit AJ, Keighley G, Lowy PH.** Metabolism of C^{14} labeled glycine, L-histidine, L-leucine, and L-lysine. *J Biol Chem.* 1950;**187**(2):839-848.

52. **Brenner S, Jacob F, Meselson M.** An unstable intermediate carrying information from genes to ribosomes for protein synthesis. *Nature.* 1961;**190**(4776)576-581.

53. **Allfrey VG.** Amino acid incorporation by isolated thymus nuclei. I. The role of desoxyribonucleic acid in protein synthesis. *Proc Natl Acad Sci U S A.* 1954;**40**(10):881-885.

54. **Goldstein L.** On the question of protein synthesis by cell nuclei. *Adv. Cell Biol.*

1970;**1**:187-210.

55. **Kuehl L.** Nuclear protein synthesis. 1974; In: Busch H. (ed) *The cell nucleus.* Academic Press, New York pp 345-375.

56. **Allfrey VG, Mirsky AE.** On the supposed contamination of thymus nuclear fractions by whole cells. *Science.* 1955;**121**(3155):879-880.

57. **Allfrey VG, Hopkins JW, Frenster JH, Mirsky AE.** Reactions governing incorporation of amino acids into the proteins of the isolated cell nucleus. *Ann N Y Acad Sci.* 1960;**88**:722-740.

58. **Chatterjee NK, Weissbach H.** Translation in vitro of total nuclear RNA from HeLa cell nuclei infected with adenovirus 2. *Proc Natl Acad Sci U S A.* 1974;**71**(8):3129-3133.

59. **Wang TY.** The isolation, properties, and possible functions of chromatin acidic proteins. *J Biol Chem.* 1967;**242**(6):1220-1226.

60. **Mirsky AE, Osawa S, Allfrey VG.** The nucleus as a site of biochemical activity. *Cold Spring Harb Symp Quant Biol.* 1956;**21**:49-73.

61. **Allfrey VG, Mirsky AE.** Mechanisms of synthesis and control of protein and ribonucleic acid synthesis in the cell nucleus. *Cold Spring Harb Symp Quant Biol.* 1963;**28**:247-262.

62. **Goidl JA, Allen WR.** Does protein synthesis occurs within the nucleus? *Trends Biochem Sci.* 1978;**3**:N225-N228.

63. **Feldherr CM.** The nuclear annuli as pathways for nucleocytoplasmic exchanges. *J Cell Biol.* 1962;**14**(1):65-72.

64. **Wu RS, Warner JR.** Cytoplasmic synthesis of nuclear proteins. Kinetics of accumulation of radioactive proteins in various cell fractions after brief pulses. *J Cell Biol.* 1971;**51**(3):643-652.

65. **Gilbert W.** Why genes in pieces? *Nature.* 1978;**271**(5645):501.

66. **Chang JC, Kan YW.** β^0 thalassemia, a nonsense mutation in man. *Proc Natl Acad Sci U S A.* 1979;**76**(6):2886-2889.

67. **Losson R, Lacroute F.** Interference of nonsense mutations with eukaryotic messenger RNA stability. *Proc Natl Acad Sci U S A.* 1979;**76**(10):5134-5137.

68. **Humphries RK, Ley TJ, Anagnou NP, Baur AW, Nienhuis AW.** β^0-39 thalassemia gene: A premature termination codon causes β-mRNA deficiency without affecting cytoplasmic β-mRNA stability. *Blood.* 1984;**64**(1):23-32.

69. **Takeshita K, Forget BG, Scarpa A, Edward J. Benz Jr.** Intranuclear defect in β-globin mRNA accumulation due to a premature translation termination codon. *Blood.* 1984;**64**(1):13-22.

70. **Schoenberg DR, Maquat LE.** Regulation of cytoplasmic mRNA decay. *Nat Rev Genet.* 2012;**13**(4):246-259.

71. **Ricketts MH, Simons MJ, Parma J, Mercken L, Dong Q, Vassart G.** A nonsense mutation causes hereditary goitre in the Afrikander cattle and unmasks alternative splicing of thyroglobulin transcripts. *Proc Natl Acad Sci U S A.* 1987;**84**(10):3181-3184.

72. **Dietz HC, Valle D, Francomano CA, Kendzior RJ Jr, Pyeritz RE, Cutting GR.** The skipping of constitutive exons in vivo induced by nonsense mutations. *Science.* 1993;**259**(5095):680-683.

73. **Jaillon O, Bouhouche K, Gout JF, Aury JM, Noel B, Saudemont B, Nowacki M, Serrano V, Porcel BM, Ségurens B, Le Mouël A, Lepère G, Schächter V, Bétermier M, Cohen J, Wincker P, Sperling L, Duret L, Meyer E.** Translational control of intron splicing in eukaryotes. *Nature.* 2008;**451**(7176):359-362.

74. **Miriami E, Sperling R, Sperling J, Motro U.** Regulation of splicing: the importance of being translatable. *RNA.* 2004;**10**(1):1-4.

75. **Chang Y-F, Imam JS, Wilkinson MF.** The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem.* 2007;**76**:51-74.

76. **Brogna S, Wen J.** Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat Struct Mol Biol.* 2009;**16**(2):107-113.

77. **Daar IO, Maquat LE.** Premature translation termination mediates triosephosphate isomerase mRNA degradation. *Mol Cell Biol.* 1988;**8**(2):802-813.

78. **Carter MS, Doscow J, Morris P, Li S, Nhim RP, Sandstedt S, Wilkinson MF.** 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.* 1995;**270**(48):28995-29003.

79. **Carter MS, Li S, Wilkinson MF.** A splicing dependent regulatory mechanism that detects translation signals. *EMBO J.* 1996;**15**(21):5965-5975.

80. **Maquat LE.** Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat Rev Mol Cell Biol.* 2004;**5**(2):89-99.

81. **Mendell JT, ap Rhys CMJ, Dietz HC.** Separable roles for rent1/hUpf1 in altered splicing and decay of nonsense transcripts. *Science.* 2002;**298**(5592):419-422.

82. **Lareau LF, Brooks AN, Soergel DA, Meng Q, Brenner SE.** The coupling of alternative splicing and nonsense-mediated mRNA decay. *Adv Exp Med Biol.* 2007;**623**:190-211.

83. **McGlincy NJ, Smith CWJ.** Alternative splicing resulting in nonsense-mediated mRNA decay: what is the meaning of nonsense? *Trends Biochem Sci.*

2008;**33**(8):385-393.

84. **Maquat LE.** 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.* 2002;**16**(14):1743-1753.

85. **Stalder L, Mühlemann O.** Transcriptional silencing of nonsense codon-containing immunoglobulin micro genes requires translation of its mRNA. *J Biol Chem.* 2007;**282**(22):16079-16085.

86. **Wang J, Hamilton JI, Carter MS, Li S, Wilkinson MF.** Alternatively spliced TCR mRNA induced by disruption of reading frame. *Science.* 2002;**297**(5578):108-110.

87. **Chang YF, Chan WK, Imam JS, Wilkinson MF.** Alternatively spliced T-cell receptor transcripts are up-regulated in response to disruption of either splicing elements or reading frame. *J Biol Chem.* 2007;**282**(41):29738-29747.

88. **Dahlberg JE, Lund E, Goodwin EB.** Nuclear translation: What is the evidence? *RNA.* 2003;**9**(1):1-8.

89. **Ishigaki Y, Li X, Serin G, Maquat LE.** Evidence for a pioneer round of mRNA translation: mRNAs subject to nonsense-mediated decay in mammalian cells are bound by CBP80 and CBP20. *Cell.* 2001;**106**(5):607-617.

90. **Maquat LE, Tarn WY, Isken O.** The pioneer round of translation: features and functions. *Cell.* 2010;**142**(3):368-374.

91. **Urlaub G, Mitchell PJ, Ciudad CJ, Chasin LA.** Nonsense mutations in the dihydrofolate reductase gene affect RNA processing. *Mol Cell Biol.* 1989,**9**(7):2868-2880.

92. **Cheng J, Maquat LE.** Nonsense codons can reduce the abundance of nuclear mRNA or the half-life of cytoplasmic mRNA. *Mol Cell Biol.* 1993;**13**(3):1892-

1902.

93. **Carter KC, Taneja KL, Lawrence JB.** Discrete nuclear domains of poly(A) RNA and their relationship to the functional organization of the nucleus. *J Cell Biol.* 1991;**115**(5):1191-1202.

94. **Iborra FJ, Jackson DA, Cook PR.** Coupled transcription and translation within nuclei of mammalian cells. *Science.* 2001;**293**(5532):1139-1142.

95. **Wachtel C, Li B, Sperling J, Sperling R.** Stop codon-mediated suppression of splicing is a novel nuclear scanning mechanism not affected by elements of protein synthesis and NMD. *RNA.* 2004;**10**(11):1740-1750.

96. **Iborra FJ, Pombo A, Jackson DA, Cook PR.** Active RNA polymerases are localized within discrete transcription 'factories' in human nuclie. *J Cell Sci.* 1996;**109**(6):1427-1436.

97. **Cook PR.** A model for all genomes: the role of transcription factories. *J Mol Biol.* 2010;**395**(1):1-10.

98. **Edelman LB, Fraser P.** Transcription factories: genetic programming in three dimensions. *Curr Opin Genet Dev.* 2012;**22**(2):110-114.

99. **Papantonis A, Cook PR.** Fixing the model for transcription: the DNA moves, not the polymerase. *Transcription.* 2011;**2**(1):41-44.

100. **Melnik S, Deng B, Papantonis A, Baboo S, Carr IM, Cook PR.** The proteomes of transcription factories containing RNA polymerases I, II or III. *Nat Methods.* 2011;**8**(11):963-968.

101. **Lund E, Dahlberg JE.** Proofreading and aminoacylation of tRNAs before export from the nucleus. *Science.* 1998;**282**(5396):2082-2085.

102. **Mangiarotti G.** Coupling of transcription and translation in *Dictyostelium discoideum* nuclei. *Biochemistry.* 1999;**38**(13):3996-4000.

103. **Iborra FJ, Escargueil AE, Kwek KY, Akoulitchev A, Cook PR.** Molecular cross-talk between the transcription, translation and nonsense-mediated decay machineries. *J Cell Sci.* 2004;**117**(6):899-906.

104. **Iborra FJ, Jackson DA, Cook PR.** Approaches for monitoring nuclear translation. *Methods Mol Biol.* 2004;**257**:103-114.

105. **Nathanson L, Xia T, Deutscher MP.** Nuclear protein synthesis: A re-evaluation. *RNA.* 2003;**9**(1):9-13.

106. **Iborra FJ, Jackson DA, Cook PR.** The case for nuclear translation. *J Cell Sci.* 2004;**117**(24):5713-5720.

107. **Rockel TD, Stuhlmann D, von Mikecz A.** Proteasomes degrade proteins in focal subdomains of human cell nucleus. *J Cell Sci.* 2005;**118**(22):5231-5242.

108. **Gillette TG, Gonzales F, Delahodde A, Johnston SA, Kodadek T.** Physical and functional association of RNA polymerase II and the proteasome. *Proc Natl Acad Sci U S A.* 2004;**101**(16):5904-5909.

109. **Keppler BR, Archer TK, Kinyamu HK.** Emerging roles of the 26S proteasome in nuclear hormone receptor-regulated transcription. *Biochim Biophys Acta.* 2011;**1809**(2):109-118.

110. **Bohnsack MT, Regener K, Schwappach B, Saffrich R, Paraskeva E, Hartmann E, Gorlich D.** Exp5 exports eEF1A via tRNA from nuclei and synergizes with other transport pathways to confine translation to the cytoplasm. *EMBO J.* 2002;21(22):6205-6215.

111. **Cosson B, Philippe M.** Looking for nuclear translation using *Xenopus* oocytes. *Biol Cell.* 2003;**95**(5):321-325.

112. **Dahlberg JE, Lund E.** Does protein synthesis occur in the nucleus? *Curr Opin Cell Biol.* 2004;**16**(3):335-338.

113. **Holbrook JA, Neu-Yilik G, Hentze MW, Kulozik AE.** Nonsense-mediated decay approaches the clinic. *Nat Genet.* 2004;**36**(8):801-808.

114. **Udem SA, Warner JR.** The cytoplasmic maturation of a ribosomal precursor ribonucleic acid in yeast. *J Biol Chem.* 1973;**248**(4):1412-1416.

115. **Thomson E, Tollervey D.** The final step in 5.8S rRNA processing is cytoplasmic in *Saccharomyces cerevisiae*. *Mol Cell Biol.* 2010;**30**(4):976-984.

116. **Soudet J, Gélugne JP, Belhabich-Baumas K, Caizergues-Ferrer M, Mougin A.** Immature small ribosomal subunits can engage in translation initiation in *Saccharomyces cerevisiae*. *EMBO J.* 2010;**29**(1):80-92.

117. **Mangiarotti G, Chiaberge S, Bulfone S.** rRNA maturation as a "quality" control step in ribosomal subunit assembly in *Dictyostelium discoideum*. *J Biol Chem.* 1997;**272**(44):27818-27822.

118. **Lytle JR, Steitz JA.** Premature termination codons do not affect the rate of splicing of neighboring introns. *RNA.* 2004;**10**(4):657-668.

119. **Custódio N, Carmo-Fonseca M.** Quality control of gene expression in the nucleus. *J Cell Mol Med.* 2001;**5**(3):267-75.

120. **Wilkinson MF, Shyu AB.** Multifunctional regulatory proteins that control gene expression in both the nucleus and the cytoplasm. *Bioessays.* 2001;**23**(9):775-787.

121. **Wilkinson MF, Shyu AB.** RNA surveillance by nuclear scanning? *Nat Cell Biol.* 2002;**4**(6):E144-E147.

122. **Bühler M, Wilkinson MF, Mühlemann O.** Intranuclear degradation of nonsense codon-containing mRNA. *EMBO Rep.* 2002;3(7):646-651.

123. **Brogna S, Sato TA, Rosbash M.** Ribosome components are associated with sites of transcription. *Mol Cell.* 2002;**10**(1):93-104.

124. **De S, Varsally W, Falciani F, Brogna S.** Ribosomal proteins' association with transcription sites peaks at tRNA genes in *Schizosaccharomyces pombe*. *RNA.* 2011;17(9):1713-1726.

125. **De S, Brogna S.** Are ribosomal proteins present at transcription sites on or off ribosomal subunits? *Biochem Soc Trans.* 2010;**38**(6):1543-1547.

126. **Custódio N, Carmo-Fonseca M, Geraghty F, Pereira HS, Grosveld F, Antoniou M.** Inefficient processing impairs release of RNA from the site of transcription. *EMBO J.* 1999;**18**(10):2855-2866.

127. **Mühlemann O, Mock-Casagrande CS, Wang J, Li S, Custodio N, Carmo-Fonseca M, Wilkinson MF, Moore MJ.** Precursor RNAs harboring nonsense codons accumulate near the site of transcription. *Mol Cell.* 2001;**8**(1):33-44.

128. **de Turris V, Nicholson P, Orozco RZ, Singer RH, Mühlemann O.** Cotranscriptional effect of a premature termination codon revealed by live-cell imaging. *RNA.* 2011;**17**(12):2094-2107.

129. **Strudwick S, Borden KL.** The emerging roles of translation factor eIF4E in the nucleus. *Differentiation.* 2002;**70**(1):10-22.

130. **Ferraiuolo MA, Lee CS, Ler LW, Hsu JL, Costa-Mattioli M, Luo MJ, Reed R, Sonenberg N.** A nuclear translation-like factor eIF4AIII is recruited to the mRNA during splicing and functions in nonsense-mediated decay. *Proc Natl Acad Sci U S A.* 2004;**101**(12):4118-4123.

131. **Herbert A, Wagner S, Nickerson JA.** Induction of protein translation by ADAR1 within living cell nuclei is not dependent on RNA editing. *Mol Cell.* 2002;**10**(5):1235-1246.

132. **Rodriguez AJ, Shenoy SM, Singer RH, Condeelis J.** Visualization of mRNA translation in living cells. *J Cell Biol.* 2006;**175**(1):67-76.

133. **Dieterich DC, Hodas JJ, Gouzer G, Shadrin IY, Ngo JT, Triller A, Tirrell DA, Schuman EM.** In situ visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nat Neurosci.* 2010;**13**(7):897-905.

134. **Schmidt EK, Clavarino G, Ceppi M, Pierre P.** SUnSET, a nonradioactive method to monitor protein synthesis. *Nat Methods.* 2009;**6**(4):275-277.

135. **Liu J, Xu Y, Stoleru D, Salic A.** Imaging protein synthesis in cells and tissues with an alkyne analog of puromycin. *Proc Natl Acad Sci U S A.* 2012;**109**(2):413-418.

136. **Dieterich DC, Link AJ, Graumann J, Tirrell DA, Schuman EM.** Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc Natl Acad Sci U S A.* 2006;**103**(25):9482-9487.

137. **Beatty KE, Liu JC, Xie F, Dieterich DC, Schuman EM, Wang Q, Tirrell DA.** Fluorescence visualization of newly synthesized proteins in mammalian cells. *Angew Chem Int Ed Engl.* 2006;**45**(44):7364-7367.

138. **Beatty KE, Tirrell DA.** Two-color labeling of temporally defined protein populations in mammalian cells. *Bioorg Med Chem Lett.* 2008;**18**(22):5995-5999.

139. **Bhalla AD, Gudikote JP, Wang J, Chan WK, Chang YF, Olivas OR, Wilkinson MF.** Nonsense codons trigger an RNA partitioning shift. *J Biol Chem.* 2009;**284**(7):4062-4072.

140. **Dolan BP, Bennink JR, Yewdell JW.** Translating DRiPs: progress in understanding viral and cellular sources of MHC class I peptide ligands. *Cell Mol Life Sci.* 2011;**68**(9):1481-1489.

141. **Dolan BP, Knowlton JJ, David A, Bennink JR, Yewdell JW.** RNA

polymerase II inhibitors dissociate antigenic peptide generation from normal viral protein synthesis: a role for nuclear translation in defective ribosomal product synthesis? *J Immunol.* 2010;**185**(11):6728-6733.

142. **David A, Dolan BP, Hickman HD, Knowlton JJ, Clavarino G, Pierre P, Bennink JR, Yewdell JW.** Nuclear translation visualized by ribosome-bound nascent chain puromycylation. *J Cell Biol.* 2012;**197**(1):45-57.
143. **Reid DW, Nicchitta CV.** The enduring enigma of nuclear translation. *J Cell Biol.* 2012;**197**(1):7-9.
144. **Dahlberg J, Lund E.** Nuclear translation or nuclear peptidyl transferase? *Nucleus*. 2012;**3**(4):1-2.
145. **Bartlett J, Blagojevic J, Carter D, Eskiw C, Fromaget M, Job C, Shamsher M, Trindade IF, Xu M, Cook PR.** Specialized transcription factories. *Biochem Soc Symp*. 2006;**73**:67-75.
146. **Chakalova L,** Fraser **P.** Organization of transcription. *Cold Spring Harb Perspect Biol*. 2010;**2**(9):a000729.
147. **Papantonis A, Kohro T, Baboo S, Larkin JD, Deng B, Short P, Tsutsumi S, Taylor S, Kobayashi M, Li G, Poh MH, Ruan X, Aburatani H, Ruan Y, Kodama T, Wada Y, Cook PR.** TNFα signals through specialized factories where responsive coding and miRNA genes are transcribed. *EMBO J*. 2012; **31**(23): 4404-4414.

Chapter 2

Materials and methods

The following methods were used in this thesis.

1. Molecular cloning
2. Cell culture and transfection
3. Microscope imaging of fluorescently tagged protein and RNA
    a. Fluorescent detection of amino acid analogues by click chemistry
    b. Immuno-fluorescence and its variants
    c. '*In cell*' transcription 'run-on' assay with BrUTP
    d. RNA fluorescent *in situ* hybridization (RNA FISH)
    e. Simultaneous immuno-labelling with RNA FISH
    f. 'Colocalization' or 'proximity' analysis of features detected by fluorescent imaging

2.1 Molecular cloning

pLdPDH [148] encodes the mitochondrial import sequence of human *PDH1α* inserted between the *Sac*I and *Pst*I sites in pEGFP-N1. The P^{tight} promoter from pTRE-Tight was inserted between the *Xho*I and *Eco*RI sites in pLdPDH. Two fragments of rat *Cd2* were amplified from total rat DNA using oligo pair 1 (which target chromosome 2, between positions 196,332,778-196,333,452 [149]) and oligo pair 3 nested with oligo pair 2 (which target chromosome 2 between positions 196,336,082-196,338,119 [149]). 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 encodes P^{tight}).

1. Vector 1 was constructed by removing from pLdPDH the fragments containing P^{tight}, the mitochondrial import sequence, and *Cd2*, and inserting the three fragments between the *Xho*I and *Bam*HI sites in plasmid 0,*EGFP*,pA (which is a promoter-less form of pEGFP-N1) [150].
2. Vector 2 was constructed by deleting the EGFP chromophore (5′-CTGACCTACGGCGTGCAG-3′; 18 bp) from vector 1 by site-directed mutagenesis using oligo pair 4.
3. Vector 3 was constructed by inserting a sequence encoding the HA epitope (5′-TACCCTTACGACGTTCCTGATTACGCT-3′; 27 bp) in Vector 1 by site-directed mutagenesis using oligo pair 5. Unexpectedly, the insert proved to be 81 bp – 5′-TACCCTTACGACGTTCCTGATTACGCTTAGCAAGTTGGGGGGAC GGGTGCTATGTACCCTTACGACGTTCCTGATTACGCT-3′ – in which two HA encoding sequences were interrupted by a 27-bp sequence. As an in-frame stop codon is found after the first HA-coding sequence, only one HA epitope should be expressed.
4. Vector 4 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.

All plasmids used are listed in **Table 1**.

Table 1: Plasmids

#	Plasmids	Source	Catalogue
1	pEGFP-N1 (4.7 kbp)	Clontech (discontinued)	6085-1
2	pTRE-Tight (2.6 kbp)	Clontech	631059
3	pTet-On Advanced (7.1 kbp)	Clontech	630930
4	pLdPDH (4.7 kbp)	Gift from Dr Garry Brown [148]	-
5	0,*EGFP*,pA (4.2 kbp)	**Ref** [150]	-
6	Vectors 1-4 (7.4 kbp)	This work, diagrams in **Fig. 3.10**	-

All the above-mentioned plasmid modifications were confirmed by DNA sequencing. Presence of P^{tight} and *Cd2* fragments in Vector 1 was verified using oligos 7 and 8. Absence of EGFP chromophore in Vector 2 was verified using oligo 9. Oligo 7 was also used to verify the presence of the HA-coding sequence in Vector 3 and the point mutation in Vector 4. Site-directed mutagenesis was performed using the QuikChange Lightning kit, and all other necessary bacterial transformations were performed using One Shot TOP10 chemically-competent *Escherichia coli*. All the polymerase chain reactions (PCRs) except for site-directed mutagenesis were performed using PfuUltra II Fusion HS DNA Polymerase 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 2**.

Table 2: 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); discussed in **Section 3.4**

e: *XbaI* recognition sequence

f: HA-epitope encoding sequence

g: premature termination codon (PTC); discussed in **Section 3.4**

2.2 Cell culture and transfection

Three different cell lines were used for various experiments performed in this work, and are listed in **Table 3**.

Table 3: Mammalian cell lines

#	Name	Origin	ATCC[a] number	Culture conditions (at 37˚C, 5% CO_2)
1	HeLa[b]	epithelial adenocarcinoma cells from cervix of Henrietta Lacks [151]	CCL-2	DMEM[c] + 5% FBS[d] + 1% PS[e]
2	Cos 7[b]	created by transforming CV-1 cell line [152] (fibroblast-like cells from kidney of African green monkey – *Cercopithecus aethiops*; ATCC[a] number: CCL-70) with SV40 (Simian vacuolating) virus [153-155] to endogenously express SV40 large T-antigen	CRL-1651	DMEM[c] + 10% FBS[d] + 1% PS[e]
3	HUVEC	human umbilical vein endothelial cells [156], pooled from multiple donors (Lonza)	-	EGM-2 MV[f]

a: American type culture collection

b: from 'cell bank' of the Sir William Dunn School of Pathology, University of Oxford

c: Dulbecco's modified Eagle medium

d: Fetal bovine serum

e: 'Penicillin + Streptomycin'

f: Endothelial growth medium-2 MV

For culturing cells on glass coverslips (22 x 22 mm; #1.5; 0.17 mm thick) were etched by sonication in 1% HF (hydrofluoric acid) for 5 min at room temperature, then washed 10 times with distilled water (henceforth annotated as dH_2O; 18.2 MΩ/cm at 25˚C; Milli-Q from Millipore), followed by rinsing in absolute ethanol (such cleansed coverslips were often stored in coplin jars, immersed in absolute ethanol, until further use) and flaming to eliminate ethanol before seeding cells.

For transfection, Cos 7 were grown in DMEM + 10% FBS (Tetracycline free). Confluent cells (80-90% occupancy) were transfected using 8 µl FuGENE HD with 2 µg DNA (0.2 µg of Vector 1-4 + 0.2 µg pTetOn Advanced + 1.6 µg sheared salmon sperm DNA) in 6-well plates. After 12 h, cells were trypsinized, washed, and replated on coverslips in 6-well plates; replating reduces background caused by input DNA. Cells were then grown for 12 h, and fixed. [For **Section 3.4.4**, cells were transfected with 1 µg Vector 3 and 1 µg pTetOn Advanced for 50 h.] For staining mitochondria, 200 nM MitoTracker Deep Red was added to live cells 15 min before fixation.

Inducers of transcription and various inhibitors (**Table 4**) were added to live cells, as indicated in Figure legends.

Table 4: Inducers and inhibitors

#	Name	Mode of action	Concentration	Solvent
1	doxycycline (dox) [157]	antibiotic from tetracycline family; induces firing from P^{tight} promoter by converting rtTA-Advanced (constitutively expressed from pTet-On Advanced under a CMV-promoter [a]) from a repressor to activator [158,159]	10 µg/ml	dH_2O
2	TNF-α [160]	tumor necrosis factor-α; phosphorylates IκBα (when unphosphorylated, sequesters NF-κB in the cytoplasm; when phosphorylated, subjected to proteasomal degradation), hence releasing NF-κB to translocate into the nucleus, inducing transcription from NF-κB dependent promoters [161]	10 ng/ml	dH_2O
3	puromycin[b] (puro) [162]	mimics amino acyl tRNA; inhibits translation by incorporating into nascent peptides and terminating them prematurely; used in this work as a marker of nascent peptides [163-165]	91 µM	dH_2O
4	cycloheximide[c] (CHX) [166]	attaches to the E-site of the 80S ribosome, thus inhibiting translation by blocking translocation of amino acyl tRNA from A-site to P-site [167]	100 µg/ml	dH_2O
5	anisomycin[c] (aniso) [168]	mimics amino acyl tRNA; inhibits translation by competitively binding to peptidyl transferase centre of 80S ribosome [169,170]	100 µg/ml	10% DMSO[d] in dH_2O
6	pactamycin[b] (pacta) [171]	inhibits translation initiation and elongation by causing structural distortions in ribosome [172,173]	100 µg/ml	DMSO[d]
7	MG132 [174]	proteasome inhibitor; peptidyl aldehyde: Z-leu-leu-leu-CHO	100 µg/ml	DMSO[d]
8	BAY 11-7082 (BAY) [175]	specifically blocks TNF-α dependent phosphorylation of IκBα, thus checks release of NF-κB into nucleus, and hence inhibits NF-κB dependent transcription	10 µM	DMSO[d]
9	DRB[e] [176]	a nucleoside analogue; inhibits casein kinase II, CDK[f] 7, CDK[f] 8, CDK[f] 9, thus inhibiting necessary phosphorylations on CTD[g] of RNA polymerase II, and hence inhibits transcription elongation [177]	100 µM	absolute ethanol
10	spliceostatin A (SSA) [178]	binds to SF3b[h] of U2 snRNP[i] in spliceosomes to inhibit splicing and accumulate pre-mRNA	100 nM	methanol

a: cytomegalovirus immediate-early (IE) promoter [179]

b: disintegrates polysome as a secondary effect [180]

c: stabilizes polysome as a secondary effect [180]

d: dimethyl sulfoxide (organic solvent)

e: 5,6-dichloro-1-β-D-ribofuranosylbenzimidazole

f: cyclin-dependent kinase [181]

g: C-terminal domain in the largest catalytic subunit (RPB1) of RNA polymerase II, which contains heptad repeats that can be phosphorylated at their serine residues, and hence regulate various catalytic states of RNA polymerse II [182-186]

h: splicing factor 3b [187]

i: small nuclear ribonucleoprotein [188-190]

2.3 Microscope imaging of fluorescently-tagged protein and RNA

Cultured and treated cells were fixed, permeabilized, and tagged with fluorescent probes to indirectly display the intra-cellular localization of proteins and RNA, and imaged by fluorescence microscopy. All buffers used for RNA detection were treated with DEPC (RNase inhibitor), and incubations done in presence of RNase inhibitor.

1. Fixation: Cells on glass coverslip were immersed in 4% paraformaldehyde in 250 mM HEPES (pH adjusted to 7.6) for 20 min at 4˚C or 20˚C. For fixation after 5-60 s pulses with puromycin (**Section 3.3.1**) – which involved dipping a glass coverslip populated with adherent HeLa cells in DMEM + 5% FBS + 1% PS ± cycloheximide ± puromycin in a 37˚C room – the glass coverslip was plunged into ice-cold fixative.
2. Permeabilization: Cells on glass coverslip were immersed in 0.5% Triton X-100 and 0.5% saponin in PBS for 20 min at 20°C. For click chemistry, saponin was excluded.
3. Blocking against non-specific binding by antibodies: Cells on glass coverslip were inverted onto a 50-µl drop (on Parafilm) of 3% BSA and 0.2% cold water fish skin gelatin in PBS for 20 min at 20°C.

4. Microscopes used for this work can be broadly divided into two classes:
    a. Wide-field fluorescence microscopes, Carl Zeiss MicroImaging, GmbH; either upright – Axioplan 2e microscope, or inverted – Axioplan 2 microscope, equipped with a 175 W Xenon arc lamp (Perkin Elmer). Images were acquired with 63x/100x Zeiss Plan-APOCHROMAT oil-immersion objectives (numerical aperture 1.4), optical filters (Chroma; mentioned in **Table 5**), and a CoolSNAP$_{HQ}$ camera (Photometrics) running under MetaMorph software (Molecular Devices), and analysed using ImageJ.
    b. Confocal microscope (FV1000, Olympus): Images were acquired with Olympus IX81, equipped with 100x Olympus UPlanSApo oil-immersion objective (numerical aperture 1.4), optical filters (Olympus; mentioned in **Table 5**); confocal aperture 175 μm (350 μm: **Fig. 5.2**; 117 μm: **Fig. 5.3**), scanning at 10 μs/pixel (8 μs/pixel: **Fig. 5.3**), 405 and 559 nm 'diode-pumped solid-state' plus argon (488 nm) lasers, and FLUOVIEW v2.1b software. The images were retrieved through MetaMorph software, and analysed using ImageJ.

Table 5: Optical filters on microscopes

	Wide-field			**Confocal**		
Fluor	**Excitation**	**Dichroic**	**Emission**	**Excitation**	**Dichroic**	**Emission**
DAPI	S360/40x	84100bs	S457/50m	BP330-385	BA420	DM400
Alexa488	S490/20x	84100bs	S528/38m	BP470-490	BA520IF	DM500
Cy3/Alexa555	S555/28x	84100bs	S617/73m	BP510-550	BA590	DM570
MitoTracker Deep Red	S635/20x	84100bs	S685/40m			
GFP	HQ470/40x	495DCLP	HQ525/50m			

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. [For **Figure 3.12H**, **3.13C**, **3.14C**, **5.4B** and **6.4** the mean intensity within a randomly-selected 30 x 30 pixel area (1292 pixel area for **Fig. 6.4**) over the nucleoplasm (or cytoplasm in **Fig. 3.14C**) was measured]. Such values from each experiment were exported to Excel for calculating mean and standard deviation (SD). *P* values (two-tailed) from unpaired Student's *t*-test were calculated using GraphPad (http://www.graphpad.com/quickcalcs/). The images were processed using Photoshop 7.0 and were presented with graphs using CorelDRAW X3.

2.3.1 Fluorescent detection of amino acid analogues by click chemistry

L-azidohomoalanine (Aha; synthesised by Mr Bhaskar Bhushan), an analogue of the essential amino acid – methionine, was dissolved in dH_2O at a stock concentration of 200 mM, and stored at -20˚C.

HeLa cells were grown in the absence of methionine ($DMEM^-$ – DMEM without methionine, cystine and L-glutamine) for 15-30 min at 37˚C (to deplete endogenous pools of methionine), pulsed with 2 mM Aha (from 5 s to 2 min) at 37˚C, and plunged into ice-cold fixative to stop translation quickly. [The 5-15 s pulses were done by dipping a glass coverslip populated with adherent HeLa cells in $DMEM^-$ + 2 mM Aha, in a 37˚C room. When a pulse was followed by chase, $DMEM^-$ + 2 mM Aha was replaced by DMEM + 5% FBS + 1% PS (which contains ~200 µM methionine).] Cells were washed in PBS (5

min; 20°C), permeabilized, washed in 3% BSA in PBS (5 min; 20°C), an alkyne-conjugated fluor (i.e., Alexa555; 100 nM) 'clicked' on to incorporated Aha using the Click-iT Cell Reaction Buffer kit as per the supplier's instructions (30 min; 20°C; vigorous shaking), washed in 3% BSA in PBS (5 min; 20°C), and rinsed in PBS. After DAPI/Vectashield counterstaining/mounting, images were acquired images using an Axioplan 2e microscope.

For **Fig. 3.3E**, R^2 values were computed by fitting an exponential decay curve ($i_t = i_{max}e^{-\lambda t}$) in Excel, and half-life ($t_{½}$) of Aha-bearing peptides was calculated.

$$t_{½} = \frac{0.693}{\lambda}$$

here,

λ = rate constant of exponential decay,

i_{max}= maximum intensity i.e., mean intensity after Aha pulse (without chase), and

i_t= intensity at time 't' i.e., mean intensity at time 't' since initiating chase, post Aha pulse.

2.3.2 Immuno-fluorescence and its variants

Conventional immuno-fluorescence [191]

It is used to localize one or more candidate proteins, and is also the most commonly-used technique (immuno-fluorescence colocalization) to determine the proximity of two candidate proteins; the two targets are indirectly immuno-labelled with spectrally-distinct fluors, and – if the fluorescent signals colocalize – it is concluded that the two targets lie together [192,193]. However, this approach is limited by the poor resolution of the light microscope; when different targets lying several hundreds of

nanometers apart in the *x* and *y* axes are labelled with red and green fluors, the two may still yield a yellow focus and so appear to colocalize. Higher-resolution approaches include the proximity-ligation assay, and antibody blocking, though each is associated with its own set of difficulties (Discussed in **Chapter 5**).

Fixed cells were washed with PBS (5 min; 20°C), permeabilized, washed with 0.05% Tween 20 in PBS (10 min; 20°C), and blocked. Antigens were indirectly immuno-labelled by incubating cells (1 h; 20°C) with a primary antibody (or two antibodies) targeting the epitope(s) of interest, washed 4x with 0.05% Tween 20 in PBS (1 min each; 20°C), then labelled further (30 min; 20°C) with secondary antibodies targeting the primary antibody (or respective primary antibodies), 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 DAPI/Vectashield counter-staining/mounting, images were acquired with the confocal microscope and/or widefield microscopes.

In situ proximity ligation assay (PLA) [194]

For conventional indirect immuno-labelling, the secondary antibodies targeting the two specific primary antibodies carry spectrally-distinct fluors. For *in situ* PLA (**Fig. 2.1**), the secondaries carry distinct (covalently-attached) single-stranded DNA oligonucleotides. Each oligonucleotide is complementary to a sequence at one end of each of two single-stranded 'padlock probes'. Then, if the two target antigens lie less than ~40 nm apart, binding of the primary and secondary antibodies brings the two oligonucleotides close together, so that they can hybridize with the ends of the padlock probes. Then, addition of a ligase circularizes the padlock probes. Now, *in situ* rolling-circle amplification (RCA) by the φ29 DNA polymerase of the (now-circular) padlock probes creates tandem repeats of the padlock sequence. Finally, hybridization of

fluorescent oligonucleotides complementary to the sequence in the centre of the padlock probe allows detection indicative of proximity between the two antigens.

Fixed cells were washed with PBS (PAA; 5 min; 20°C), permeablized, washed with 0.05% Tween 20 in PBS (10 min; 20°C), and blocked. Antigens were indirectly immuno-labelled by incubating cells (1 h; 20°C) with various primary antibodies. Next, following the manufacturer's instructions (Duolink II kits), the secondary antibodies covalently attached to oligonucleotides were bound to their targets, and then the tethered oligonucleotides were detected using 'Duolink II Detection Reagents Orange'. After DAPI/Vectashield counterstaining/mounting, images were acquired with the confocal microscope. Multiple images were acquired through *z*-axis (known as a Z-stack) for each field of view, Z-projections for each Z-stack were made, and the number of fluorescent foci/nucleus counted.

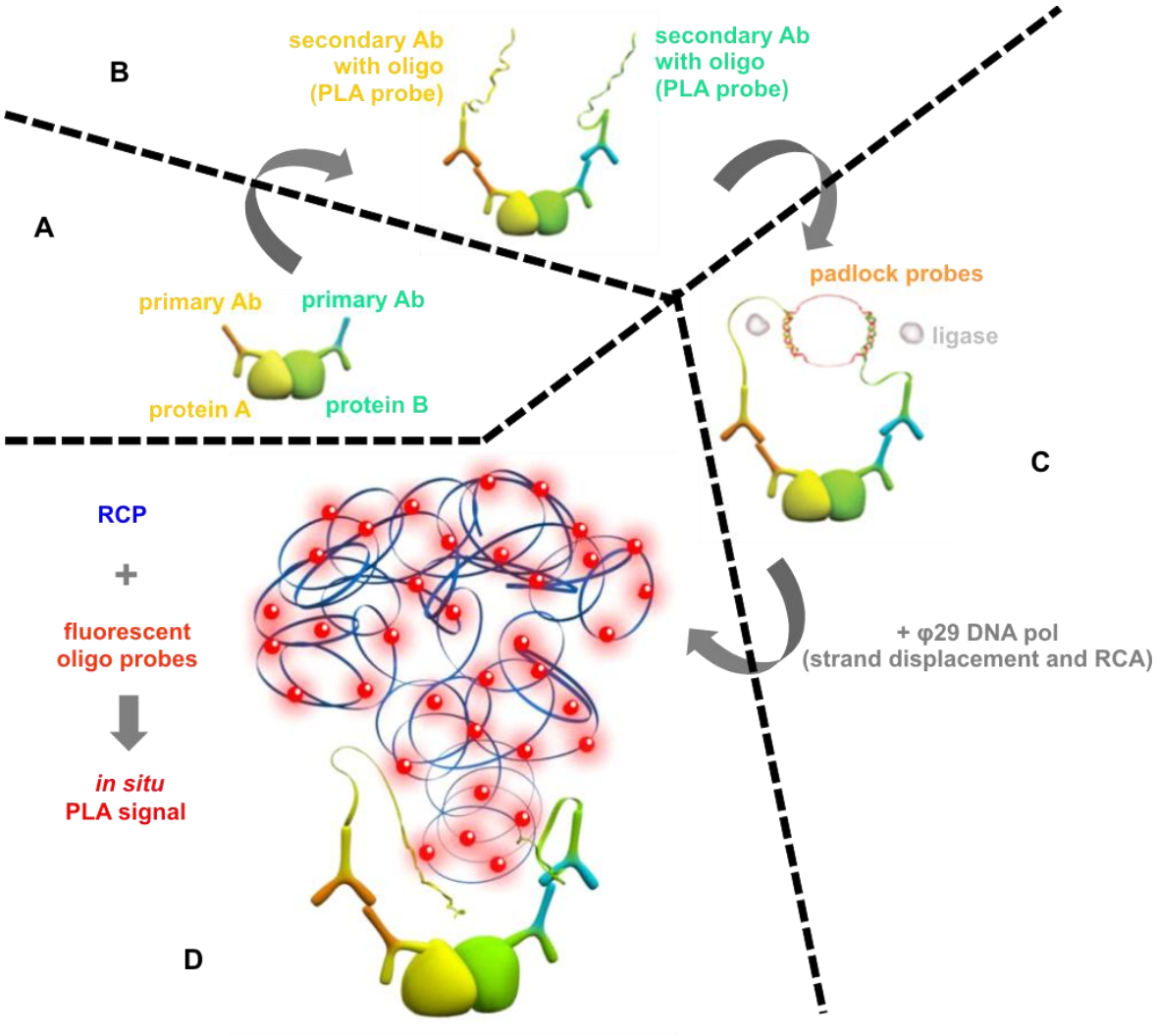


Fig. 2.1. Principle of *in situ* proximity ligation assay (PLA). *In situ* PLA is used to evaluate proximity between two proteins in cells, if both proteins have antibody-binding epitopes. Adapted by permission from Macmillan Publishers Ltd: [Nature Methods] (Gullberg *et al.* [195]), © (2011).

A. primary antibodies specifically detect proteins A and B;

B. secondary antibodies, with covalently attached oligonucleotides, subsequently detect the primary antibodies;

C. if proteins A and B lie close enough (within 40 nm), padlock probes – complimentary to the oligonucleotides attached to the secondary antibodies – hybridize to these oligonucleotides and are ligated to form a circular DNA;

D. φ29 DNA polymerase initiates a rolling circle amplification (RCA) on this circular DNA using one oligonucleotide (attached to the secondary antibody) as a primer, to give a rolling circle product (RCP)

which can be detected using fluorescently-tagged oligonucleotide probes. The signal from such an *in situ* PLA experiment are easily detected under a fluorescence microscope.

Antibody blocking assay [196,197]

This provides an even higher-resolution approach; it relies on the large size of an antibody (**Fig. 2.2)**. If two target proteins, A and B, lie within ~10 nm of each other (roughly the dimensions of one antibody), binding of one primary antibody to A will block access of a second primary antibody to B. In other words, anti-A blocks access of anti-B only if A and B lie close together. If bound anti-B is detected using a fluorescent secondary, the presence of (non-fluorescent) anti-A reduces the signal. Here, anti-A is called the 'blocking antibody', and anti-B the 'detection antibody'.

Fixed cells were washed with PBS (5 min; 20°C), permeabilized, washed with 0.05% Tween 20 in PBS (10 min; 20°C), and blocked. Antigens were first immuno-blocked by incubating cells (16 h; 4°C) with various blocking antibodies, then washed 2x in 0.05% Tween 20 in PBS (1 min each; 20°C), fixed again with 4% paraformaldehyde in PBS (5 min; 20°C), and washed with PBS (10 min; 20°C). Next, cells were incubated (1 h; 20°C) with a detection antibody, and washed 4x with 0.05% Tween 20 in PBS (1 min each; 20°C). Now, cells were incubated (30 min; 20°C) with the appropriate antibody conjugated with a fluor, 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 DAPI/Vectashield counterstaining/mounting, images were acquired using an Axioplan 2e microscope. The mean fluorescent intensities of 5 nuclei from each experiment were expressed (as percentages) relative to those given by an 'irrelevant' blocking antibody (i.e., goat IgG).

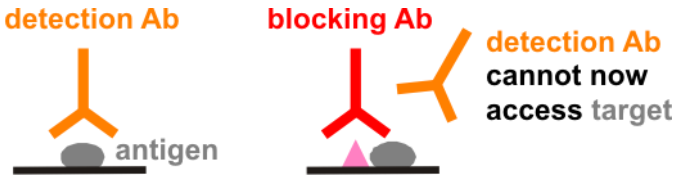


Fig. 2.2. Principle of 'antibody blocking' assay. This exploits the ability of one non-fluorescent antibody ('blocking Ab') to prevent access of another fluorescent antibody ('detection Ab') to its target (the two targets must lie within ~10 nm, the dimensions of an antibody). In this assay, only the detection antibody is tagged with a fluor. It should be noted that such antibody blocking inevitably reduces the degree of colocalization obtained when using the approaches such as the traditional 'immuno-fluorescence colocalization assay' and the *in situ* PLA.

2.3.3 '*In cell*' transcription 'run-on' assay with BrUTP

Nascent RNA (i.e., RNA that is not yet released from the elongating polymerase) is an excellent marker for the active transcription complex. Nascent RNA can be immuno-localized after gently permeabilizing cells in a 'physiological' buffer (PB^+), and allowing the still-engaged polymerases to extend their transcripts [197] by ~500 nucleotides [198]; the resulting BrRNA cannot be spliced and does not leave transcription sites [199,200]. PB (stored at 4°C) contains: 100 mM CH_3COOK, 30 mM KCl, 10 mM Na_2HPO_4, 1 mM $MgCl_2$. PB^+ (made up immediately before the experiment, and stored at 4°C) is PB supplemented with 10 mM NaF (phosphatase inhibitor), 10 mM β-glycerophosphate (energy source, phosphatase inhibitor), 200 µM Na_3VO_4 (phosphatase inhibitor) 1 mM Na_2ATP (energy source), 1 mM DTT (reducing agent), 1/1000 protease inhibitor cocktail, 1/2000 RNase inhibitor; pH 7.6.

For 'run-ons' in BrUTP, cells on coverslips were washed with ice-cold PB^+ for 1 min, mildly permeabilized with 170 µg/ml saponin in ice-cold PB^+ (5 min; 4°C), washed 3x with ice-cold PB^+ (5 min each; 4°C), transferred to transcription buffer, pre-incubated for 5 min at 33˚C, and transcription started by the addition of BrUTP (to a final

concentration of 100 µM); transcription continued for 15 min at 33°C. Transcription buffer is PB^+ supplemented with 100 µM each of ATP, GTP and CTP (but no UTP), + 400 µM $MgCl_2$ + 1/100 RNase inhibitor. Transcription was stopped by adding EDTA (pH 8.0) to a final concentration of 2.5 mM, and cells washed 2x with ice-cold PB^+ and fixed. In experiments requiring cycloheximide pretreatment, cycloheximide was maintained throughout the PB^+ washes, permeabilization and run-on (until fixation). Nascent BrRNA in fixed cells was detected by indirect immuno-fluorescence as described in **Section 2.3.2**.

2.3.4 RNA fluorescent *in situ* hybridization (RNA FISH)

Definite stretches of nucleic acids (DNA and RNA) can be spatially localized in the cells by hybridizing fluorescently-labelled complementary single-stranded (ss) nucleic acids – a technique known as fluorescent *in situ* hybridization (FISH). FISH techniques are well established [201], and now sensitive enough to detect a single molecule of RNA [202] – albeit with some difficulty. I use either (i) four/five fluorescently-tagged ssDNA ~50-mer probes (each 50-mer containing an average of 3.5 to 4.5 fluorophores) [203] to target a single nascent RNA molecule, or (ii) forty-eight fluorescently-tagged ssDNA 20-mer probes [204], and a protocol modified from van de Corput *et al*. [205,206,150].

Designing and labelling RNA FISH probes

1. Sets of four/five ~50-mer probes were synthesized (Gene Design, Japan) [206]; one set targeted *SAMD4A* region *1s* (~1.5 kbp into intron 1), another *EXT1* (~2 kbp into intron 1) – both close to the transcription start site ('*tss*'), and a third *EDN1* (intron 2). In each 50-mer, roughly every tenth thymine residue was substituted by an amino-modified C6-dT. The amino group was subsequently tagged with Alexa488 fluor (for *EDN1*) or Alexa555 fluor (for *SAMD4A* and

EXT1) reactive dye according to the manufacturers' instructions. Probes were then purified using G-50 columns, ethanol precipitated twice, and concentrated using a Microcon YM-30 column. Labelling efficiency was found to be between 7-9 fluors per 100 nucleotides. **Table 6** lists probe sequences.

2. A set of forty-eight 20-mer probes (designed using Probe Designer; **Table 11**; #9) were synthesized (Biosearch Technologies, USA) targeting the intronic region (~1.6 kbp) of rat *Cd2* DNA present in Vectors 1-4 (**Section 2.1** and **Fig. 3.10**). At the 3′-end of each 20-mer, an mdC(TEG-Amino) modification was added. The amino group was subsequently tagged with Alexa488 fluor reactive dye according to the manufacturers' instructions. Probes were then purified using G-25 columns, ethanol precipitated twice, and concentrated using a Microcon YM-3 column. Labelling efficiency was found to be ~5 fluors per 100 nucleotides (i.e., ~100% efficient). **Table 7** lists probe sequences.

Labelling efficiencies for all probes were estimated using an online 'Base:Dye ratio calculator' (**Table 11**; #10) which calculates nucleotide-base:fluorescent-dye ratios (after measuring absorbance of fluors using a spectrophotometer).

RNA FISH

Fixed cells were washed with PBS (5 min; 20°C). When using 50-mer probes, they were stored in 70% ethanol (48 h; 4°C), and transferred to PBS (5 min; 20°C) prior to use. Cells were permeabilized as above, rinsed in water, post-fixed with 4% paraformaldehyde in PBS (5 min; 20°C), washed in PBS (10 min; 20°C), gradually dehydrated in ethanol (70%, 90% and 100%; 3 min each; 20°C), and allowed to hybridize (16 h; 37°C) with the relevant probes in hybridization mix. The hybridization mix consists of 25% (10% when using 20-mers) deionized formamide, 2X SSC, 200 ng/µl

sheared salmon sperm DNA, 5X Denhardt's, 50 mM phosphate buffer, 1 mM EDTA (pH 8.0), and 0.5 µl (per coverslip) RNase inhibitor. Next day, coverslips were washed 3x in 2xSSC (10 min each; 37°C), followed by a wash with PBS (10 min; 20°C) [When Vectors 3 and 4 (which express EGFP in green channel) were used, FISH probes labelled with Alexa488 fluor (which also fluoresces in the green channel) were immuno-labelled using primary antibody against Alexa488 fluor and Cy3-conjugated secondary antibody (as described in **Section 2.3.5**) to avoid mixing of EGFP and Alexa488 fluorescence in the RNA FISH experiment for **Figure 3.13**.]. After DAPI/Vectashield counterstaining/mounting, images were acquired using an Axioplan 2e microscope.

Table 6: Sequences of ~50-mer RNA FISH probes

Gene ***H. sapiens***	**#**	**length (nucleotides)**	**Sequence (5′-3′)** **T → AminoC6 dT**
SAMD4A region *1s* intron 1 (Chr. 14) +ve strand	1	54	GCGGTCACTTCCACCCTATGATTCTATATATGGAAGCATCC CATTCTAGTGTTT
	2	50	TCCATACTTGGCCATGCAGGATCTTTCGGGTGCAGTCCTGT CTGGTCAGG
	3	53	GTGCTGCTCCAACATTTGAATAACCTAAAGATGAGTTTGAA CAGTCCCTACCA
	4	57	TGATTCAATCATTTCTTGGAATCTCCAATGACCATGAGCCT GAGGCAAAGGATGAAC
EXT1 intron 1 (Chr. 8) -ve strand	1	50	AGGCTCACTTTAGGGTCCTTCACAATGCGGATAACTTGGGG TTCCTAGAA
	2	52	TCTGTCTCTCCAACTATGGCCAAACGATTTAGCAGTTTGTTC AGAGCTCAAA
	3	52	GGACTGCAAAGCACTTAGTGCACAGTAAAGTCCTACATAA GCATTTGTGGCG
	4	51	TACATCACCCTTCCTCAATCAAATCATCAGGACCCCTCTCT GAGGTTCCAT
EDN1 intron 3 (Chr. 6) +ve strand	1	52	CTCATAGCCAGAGGGCTCTCCAATAATTTCCAATGTGTCCA TTTTCATTCTC
	2	49	CTCCTTGTGGTTTTTGGTGTGGTTTGATATTGTTTGGATTTT GGTCCTC
	3	51	AATGTGTGCTTGTTTATGACTGCTTCCGACAGATGATGAAC TAGTGTCCAG
	4	55	CTGCTACAAACTCACTCCTGCACAAATGGCTTCAACACTTT GAGCCTAGGTTTTT
	5	55	ATTCTAACCCTCTATATCATCACTTCTGCCTCTCAGTCTCCA CCCTCCCATGAGA

Table 7: Sequences of 20-mer 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

2.3.5 Simultaneous immuno-labelling and RNA FISH

There are several protocols available that combine immuno-fluorescence with RNA FISH [203,205,207], and almost all of them target either steady-state mRNA molecules or multiple copies of RNA. They yield relatively robust RNA FISH signals, in spite of the strong detergent washes necessary to obtain low backgrounds. When detecting a single nascent RNA molecule at the transcription site, the RNA FISH probes are targeted against intronic RNA which is known to exist only transiently, as it is degraded quickly [208]. This results in low signal-to-noise ratio, because the RNA FISH signal is very low as compared to the background noise due to non-specific binding. To improve the signal-to-noise ratio, cells/tissues are often treated with low amounts of pepsin in acidic pH to improve probe penetration [205], and this also reduces background binding. However, in the case where RNA FISH is combined with immuno-fluorescence, the pepsin will destroy the protein targeted. Therefore, to strike a compromise between the RNA FISH and immuno-fluorescence protocols, cells are sometimes exposed to a mild but prolonged deproteinization in 70% (cold) ethanol (instead of a short pepsin treatment), and washed with BSA (instead of Tween 20). Thus, conditions have to be optimized in such a way that there had to be little detergent washing (to retain the RNA FISH signal), but enough to remove as much background signal as possible.

After performing RNA FISH as described in **Section 2.3.4**, 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. The protein of interest was detected by incubating the cells with primary antibody (1 h; 20°C), washed 4x with 0.05% Tween 20 in PBS (1 min each; 20°C), then labelled further by incubating these cells with secondary antibody (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). [For experiments with HUVECs (**Section 6.2.3**), all washings

after 2xSSC wash were done with 0.2% BSA in PBS (20°C), for a maximum of 3 min.] After DAPI/Vectashield counter-staining and mounting, images were collected using an Axioplan 2 inverted microscope.

2.3.6 'Colocalization' or 'proximity' analysis of features detected by fluorescent imaging

Spectrally distinct fluors may be used to indirectly localize two different species of sub-cellular molecules – for example, specific proteins and definite stretches of DNA or RNA (using immuno-fluorescence and FISH, respectively) – so that any co-localization/proximity between the two can be analyzed. [Colocalization usually implies the complete or partial overlap of fluorescent signal from two spectrally-distinct fluors targeting two distinct moieties. Proximity between two such distinct moieties is represented by a definite separating distance (which can be measured with a defined precision) between the two fluors.]

Sub-cellular localization patterns of many biological molecules are complex – as is the case in certain experiments in this work (**Figs 3.8**, **3.9**, and **3.12**). To test proximity (rather than colocalization) detected in such patterns, I (with the help of Dr Joshua Larkin) applied a general and automatic 'super-resolution' method to determine whether the red and green foci were closer to each other than expected by chance [209]. The method involves the following steps (**Fig. 2.3**):

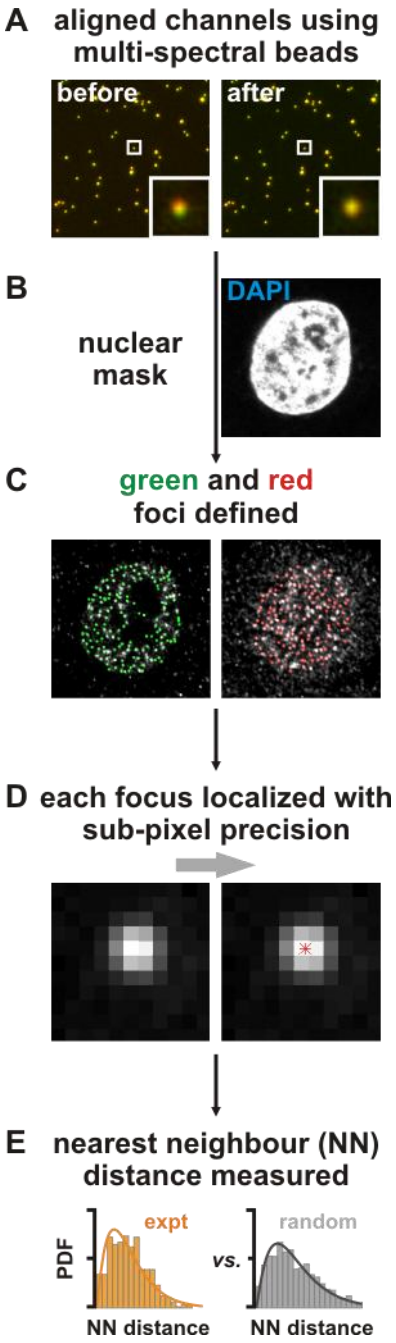


Fig. 2.3. Super-resolution proximity analysis. The method. Adapted with permission of [Journal of Cell Science], from (Larkin *et al*. [209], © 2013); permission conveyed through Copyright Clearance Center, Inc.

A. The red channel was aligned relative to the green channel using multi-spectral beads. [The translational and rotational misalignment between channels was measured using 0.1 μm TetraSpeck beads. Misalignment was corrected using a 2D spatial transformation (*i.e.*, bi-linear interpolation following a local weighted mean of a minimum of 12 fiduciary points throughout the image).]

B. Only those features in the red and green channels were selected which lay within a nuclear mask – created using DAPI-staining of nuclear DNA (DAPI is 4′,6-diamidino-2-phenylindole [210]).

C. Foci/spots were separately and automatically selected in red and green channels by fitting a Gaussian shape to the features (independent of intensity), then removing any feature of <4 pixels (which was considered an artefact as it was too small to be a diffraction-limited spot), followed by choosing only such features which were above a minimum intensity (for each channel this was the mean fluorescence under the nuclear mask + ½SD), and finally (from these chosen features) – screening for features which have a minimum local contrast (ratio of intensity of the brightest pixel in a feature, to the mean local intensity surrounding the foci; where a focus is defined as a 5x5 pixel area with the brightest pixel at the centre, and the local surrounding area is a 2 pixel-wide area around this foci).

D. The 'Joint distribution' algorithm [211] was now applied to such selected spots (in each channel, independently) to estimate the location of fluors (that produce these spots) with sub-pixel precision (~15 nm).

E. For each red spot, the distance to its nearest neighbouring (NN) green spot was estimated, and *vice-versa*, with a precision of 10-50 nm. The distribution of such NN distances from the experiments was plotted as a 'probability density function' (PDF) or 'cumulative distribution function' (CDF). The PDF and CDF were now compared to the PDF and CDF generated from red and green spots randomly distributed in a 250x250 pixel area with the same density as in the experiment. All custom software routines were implemented in MATLAB, and are available upon request.

2.4 Antibodies, general chemicals and reagents

Lists of the suppliers of general chemicals/reagents, enzymes, and software are included in **Tables 8-11**.

Table 8: Antibodies

#	Epitope (clone)	Type (Fluor)	Concentration	Source	Catalogue
1	puromycin (12D10)	Mouse IgG_{2a} monoclonal	1/100	Gift from Dr Philippe Pierre [212]	-
2	RNAPII CTD $Ser5^P$	Rabbit IgG polyclonal	250 ng/ml	Abcam	ab5131
3	BrdU/BrU	Rabbit IgG polyclonal	730 ng/ml	Rockland	600-401-C29
4	Rat CD2 (OX34)	Mouse IgG_{2a} monoclonal	1/2 hybridoma supernatant	Gift from Prof Neil Barclay [213]	-
5	HA (C29F4)	Rabbit IgG monoclonal	1/100	Cell Signaling Technology	3724S
6	RNAPII CTD $Ser2^P$	Rabbit IgG polyclonal	10 ng/ml	Abcam	ab5095
7	RNAPIII (C32)	Mouse IgG_1 monoclonal	20 µg/ml	Santa Cruz biotechnology	sc-21754
8	CTCF (E-14)	Goat IgG polyclonal	2^a µg/ml	Santa Cruz biotechnology	sc-15913
9	EXOSC6 (G-13)	Goat IgG polyclonal	200 µg/ml	Santa Cruz biotechnology	sc-107534
10	DDX1 (N-13)	Goat IgG polyclonal	2 µg/ml	Santa Cruz biotechnology	sc-49817
11	hnRNP A2/B1 (F-16)	Goat IgG polyclonal	2 µg/ml	Santa Cruz biotechnology	sc-10035
12	Lupus La/ SSB (N-20)	Goat IgG polyclonal	2 µg/ml	Santa Cruz biotechnology	sc-21391
13	U2AF65 (N-14)	Goat IgG polyclonal	2 µg/ml	Santa Cruz biotechnology	sc-19958
14	Sp3 (F-7)	Mouse IgG_{2a} monoclonal	2 µg/ml	Santa Cruz biotechnology	sc-28305
15	ATRX (D-5)	Mouse IgG_{2a} monoclonal	2 µg/ml	Santa Cruz biotechnology	sc-55584
16	-	normal Mouse IgG	2.5 µg/ml	Upstate Cell Signalling Solutions	12-371
17	-	normal Goat IgG	2 µg/ml	Santa Cruz biotechnology	sc-2028

Table 8 (contd.): Antibodies

#	Epitope (Clone)	Type (Fluor)	Concentration	Source	Catalogue
18	p65 Ser536^{P} (93H1)	Rabbit IgG monoclonal	1/1000	Cell Signaling Technology	3033S
19	Alexa488	Rabbit IgG polyclonal	1 μg/ml	Invitrogen	A11094
20	Mouse	Donkey IgG (Cy3)	1/200	Jackson ImmunoResearch	715-165-150
21	Rabbit	Donkey IgG (Alexa488)	500 ng/ml	Invitrogen	A21206
22	Rabbit	Donkey IgG (Cy3)	1/2000	Jackson ImmunoResearch	711-165-152
23	Goat	Chicken IgG (Alexa488)	1/200	laboratory stock[b]	-
24	Mouse	Donkey IgG (PLUS[c])	1/5	Olink Bioscience	92001-0030
25	Rabbit	Donkey IgG (PLUS[c])	1/5	Olink Bioscience	92002-0030
26	Goat	Donkey IgG (MINUS[c])	1/5	Olink Bioscience	92006-0030

a: 20 μg/ml for conventional immuno-fluorescence; **Figure 5.2**

b: from unknown commercial vendor

c: an oligonucleotide attached to the antibody instead of a fluor

Table 9: Chemicals, biochemicals, kits and other accessories

#	Name	Source	Catalogue
1	Acrodisc 32 mm syringe filter (0.2 µm)	Pall corporation	4652
2	UltraPure Agarose	Invitrogen	16500-500
3	ampicillin sodium salt	Sigma	A9518
4	anisomycin from *Stroptomyces gresiolus*	Sigma	A9789
5	Alexa Fluor 555 alkyne, triethylammonium salt	Invitrogen	A20013
6	ARES Alexa Fluor 488 DNA labeling kit	Invitrogen	A-21665
7	ARES Alexa Fluor 555 DNA labeling kit	Invitrogen	A-21677
8	ATP disodium salt (Na_2ATP) hydrate	Sigma	A3377
9	L-azidohomoalanine (Aha)	gift[a]/Invitrogen	C10102
10	BAY 11-7082	Santa Cruz biotechnology	sc-200615
11	bovine serum albumin (BSA)	Sigma	A7030
12	5-bromouridine 5′-triphosphate (BrUTP) sodium salt	Sigma	B7166
13	25 cm^2 cell culture flask	Corning	3056
14	75 cm^2 cell culture flask	Corning	430641
15	Click-iT Cell reaction buffer kit	Invitrogen	C10269
16	cold water fish skin gelatin	Sigma	G7041
17	22x22 mm, No. 1.5, cover glass (coverslips)	VWR international	631-0125
18	cycloheximide	Sigma	C1988
19	50x Denhardt's solution	Invitrogen	750018
20	deoxy-nucleotide triphosphate (dNTP) 100 mM set	Invitrogen	10297-018
21	4′,6-Diamidino-2-phenylindole (DAPI)	Sigma	D9542
22	5,6-dichlorobenzimidazole 1-β-D-ribofuranoside (DRB)	Sigma	D1916
23	diethyl pyrocarbonate (DEPC)	Sigma	D5758
24	dimethyl sulfoxide (DMSO) hybrimax	Sigma	D2650
25	disodium hydrogen orthophosphate (Na_2HPO_4) dihydrate	BDH AnalaR	103834G
26	DL-dithiothreitol (DTT)	Sigma	D5545
27	DMEM high glucose	PAA	E15-843
28	DMEM (without methionine, cystine, L-glutamine)	Sigma	D0422
29	DNA Engine, Peltier thermal cycler	MJ Research	PTC-200
30	1 kb DNA ladder	Promega	G5711

Table 9 (contd.): Chemicals, biochemicals, kits and other accessories

#	Name	Source	Catalogue
31	doxycycline hydrochloride	Calbiochem	324385
32	Duolink II detection reagents orange	Olink Bioscience	92007-0030
33	Duolink II wash buffers for fluorescence	Olink Bioscience	82049-0004
34	EDTA disodium salt	BDH AnalaR	20302.260
35	E-gel *96* High Range DNA marker	Invitrogen	12352-019
36	EGM-2 MV	Lonza	CC-3202
37	ethanol (C_2H_5OH) absolute	BDH AnalaR	20821.330
38	Fetal bovine serum (FBS)	Biosera	S1650-500
39	Fetal bovine serum (FBS)	PAA	A15-304
40	formamide (deionized)	Fluka	BP228-100
41	FuGENE HD transfection reagent	Promega	E2311
42	FuGENE HD transfection reagent	Roche	04709705001
43	Gel loading dye blue (6X) for DNA	New England BioLabs	B7021S
44	β-glycerophosphate	Calbiochem	35675
45	1M HEPES buffer solution	PAA	S11-001
46	40% Hydrofluoric acid (HF)	Merck	100335
47	Illustra Microspin G-25 columns	GE Healthcare	27-5325-01
48	Illustra Microspin G-50 columns	GE Healthcare	27-5330-01
49	isopropanol (propan-2-ol)	Fisher Scientific	P/7500/17
50	kanamycin sulphate from *Streptomyces kanamyceticus*	Sigma	K1377
51	magnesium chloride ($MgCl_2$) hexahydrate	Fluka	63068
52	Microcon Ultracel YM-3	Millipore	42403
53	Microcon Ultracel YM-30	Millipore	42409
54	MitoTracker Deep Red	Invitrogen	M22426
55	MG-132	Enzo	PI-102
56	Murine RNase inhibitor	New England BioLabs	M0314L
57	NanoDrop 1000 spectrophotometer	Thermo Scientific	ND-1000
58	One Shot TOP10 Chemically competent *E. Coli*	Invitrogen	C4040-06
59	pactamycin from *Streptomyces pactum*	gift[b]	-
60	16% paraformaldehyde (methanol free)	Electron Microscopy Sc.	15710-S

Table 9 (contd.): Chemicals, biochemicals, kits and other accessories

#	Name	Source	Catalogue
61	Parafilm "M"	Pechiney	PM-999
62	1x PBS (phosphate buffered saline)	PAA	H15-002
63	Penicillin/Streptomycin 100x (PS)	PAA	P11-010
64	round base 2.5 ml polystyrene tubes (for transfection)	Gosselin	TL55-04
65	potassium acetate (CH_3COOK) trihydrate	BDH AnalaR	27652.260
66	potassium chloride (KCl)	BDH AnalaR	101984L
67	Protease inhibitor cocktail (PIC), EDTA-free	Sigma	S8830
68	puromycin dihydrochloride *Stroptomyces alboniger*	Sigma	P7255
69	QIAfilter plasmid maxi kit	Qiagen	12262
70	QIAprep spin miniprep kit	Qiagen	27104
71	QIAquick gel extraction kit	Qiagen	28704
72	Quick-Load 100 bp DNA ladder	New England BioLabs	N0467S
73	QuikChange Lightning site-directed mutagenesis kit	Stratagene	210518
74	Rat (*Rattus norvegicus*) genomic DNA	Bioline	BIO-35026
75	rCTP, rATP, rUTP, rGTP 100 mM	Promega	E6000
76	RiboLock RNase inhibitor	Thermo Scientific	EO0381
77	RNaseOUT recombinant ribonuclease inhibitor	Invitrogen	10777-019
78	RNaseZAP	Sigma	R2020
79	SafeView nucleic acid stain	NBS Biologicals	NBS-SV1
80	Salmon sperm DNA (sheared)	Invitrogen	AM9680
81	saponin from *Quillaja* bark	Sigma	S4521
82	sodium chloride (NaCl)	BDH AnalaR	27810.295
83	sodium fluoride (NaF)	Riedel-deHaën	01148
84	sodium hydroxide (NaOH)	Sigma	S8045
85	sodium orthovandate (Na_3VO_4)	Sigma	S6508
86	spliceostatin A (SSA)	gift[c]	-
87	20x SSC (saline sodium citrate) buffer	Sigma	S6639
88	SuperFrost microscope slides	VWR international	631-0909
89	syringe, without needle	Terumo	BS-50ES
90	Tet system approved FBS (tetracycline free)	Clontech	631106

Table 9 (contd.): Chemicals, biochemicals, kits and other accessories

#	Name	Source	Catalogue
91	TetraSpeck fluorescent microspheres sampler kit	Invitrogen	T7284
92	6 well 'tissue culture'-treated plates	Corning	3516
93	recombinant human TNF-α	PeproTech	300-01A
94	Triton X-100	Sigma	T8787
95	1x TrypLE Express	Gibco	12604-013
96	tryptone	Sigma	T9410
97	Tween 20, for molecular biology	Sigma	P9416
98	Vectashield mounting medium for fluorescence	Vector Laboratories	H1000
99	yeast extract	Sigma	Y1625

a: from Prof Benjamin Davis, Department of Chemistry, University of Oxford, UK

b: from Prof Eric Cundliffe, Department of Biochemistry, University of Leicester, UK

c: from Prof Minoru Yoshida, RIKEN Institute, Japan

Table 10: Enzymes

#	Name	Source	Catalogue
1	PfuUltra II Fusion HS DNA polymerase	Stratagene	600670
2	T4 DNA Ligase	Roche	10716359001
3	T4 DNA Ligase (high concentration)	New England BioLabs	M0202T
4	*XhoI*	Roche	10703770001
5	*EcoRI*	Roche	11175084001
6	*NheI*	New England BioLabs	R0131S
7	*XbaI*	Roche	10674265001
8	*PstI*	New England BioLabs	R0140S
9	*BamHI*	New England BioLabs	R0136S
10	*BamHI*	Roche	10567604001

Table 11: Softwares and online calculators

#	Name	Source
1	ImageJ	**Rasband WS.** ImageJ, U.S. National Institutes of Health, Bethesda, Maryland, USA (1997-2007); http://rsb.info.nih.gov/ij/
2	Metamorph 7.7.2.0	Molecular devices
3	FLUOVIEW v2.1b	Olympus
4	Office 2010	Microsoft
5	CorelDrawX3	Corel
6	Photoshop 7.0	Adobe
7	MATLAB	MathWorks
8	GraphPad	http://www.graphpad.com/quickcalcs/
9	Probe Designer	http://singlemoleculefish.com/
10	Base:Dye ratio calculator	Invitrogen; http://probes.invitrogen.com/resources/calc/basedyeratio.html

2.5 References

148. **Margineantu DH, Brown RM, Brown GK, Marcus AH, Capaldi RA.** Heterogeneous distribution of pyruvate dehydrogenase in the matrix of mitochondria. *Mitochondrion*. 2002;**1**(4):327-338.

149. **Rat Genome Sequencing Project Consortium.** Genome sequence of the Brown Norway rat yields insights into mammalian evolution. *Nature*. 2004;**428**(6982):493-521.

150. **Xu M, Cook PR.** Similar active genes cluster in specialized transcription factories. *J Cell Biol*. 2008;**181**(4):615-623.

151. **Lucey BP, Nelson-Rees WA, Hutchins GM.** Henrietta Lacks, HeLa cells, and cell culture contamination. *Arch Pathol Lab Med*. 2009;**133**(9):1463-1467.

152. **Jensen FC, Girardi AJ, Gilden RV, Koprowski H.** Infection of human and simian tissue cultures with Rous sarcoma virus. *Proc Natl Acad Sci U S A*. 1964;**52**(1):53-59.

153. **Myers RM, Rio DC, Robbins AK, Tjian R.** SV40 gene expression is modulated by the cooperative binding of T antigen to DNA. *Cell*. 1981;**25**(2):373-384.

154. **Gluzman Y.** SV40-transformed simian cells support the replication of early SV40 mutants. *Cell*. 1981;**23**(1):175-182.

155. **Mellon P, Parker V, Gluzman Y, Maniatis T.** Identification of DNA sequences required for transcription of the human alpha 1-globin gene in a new SV40 host-vector system. *Cell*. 1981;**27**(2):279-288.

156. **Jaffe EA, Nachman RL, Becker CG, Minick CR.** Culture of human endothelial cells derived from umbilical veins. Identification by morphologic and immunologic criteria. *J Clin Invest*. 1973;**52**(11):2745-2756.

157. **Joshi N, Miller DQ.** Doxycycline revisited. *Arch Intern Med*. 1997;**157**(13):1421-1428.

158. **Gossen M, Freundlieb S, Bender G, Müller G, Hillen W, Bujard H.** Transcriptional activation by tetracyclines in mammalian cells. *Science*. 1995;**268**(5218):1766-1769.

159. **Urlinger S, Baron U, Thellmann M, Hasan MT, Bujard H, Hillen W.** Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proc Natl Acad Sci U S A*. 2000;**97**(14):7963-7968.

160. **Pennica D, Nedwin GE, Hayflick JS, Seeburg PH, Derynck R, Palladino MA, Kohr WJ, Aggarwal BB, Goeddel DV.** Human tumour necrosis factor: precursor structure, expression and homology to lymphotoxin. *Nature*. 1984;**312**(5996):724-729.

161. **Beg AA, Finco TS, Nantermet PV, Baldwin AS Jr.** Tumor necrosis factor and interleukin-1 lead to phosphorylation and loss of IκBα: a mechanism for NF-κB activation. *Mol Cell Biol*. 1993;**13**(6):3301-10.

162. **Yarmolinsky MB, Haba GL.** Inhibition by puromycin of amino acid incorporation into protein. *Proc Natl Acad Sci U S A*. 1959;**45**(12):1721-1729.

163. **Nathans D.** Puromycin inhibition of protein synthesis: incorporation of puromycin into peptide chains. *Proc Natl Acad Sci U S A*. 1964;**51**(4):585-592.

164. **Williamson AR, Schweet R.** Role of the genetic message in polyribosome function. *J Mol Biol*. 1965;**11**(2):358-372.

165. **David A, Dolan BP, Hickman HD, Knowlton JJ, Clavarino G, Pierre P, Bennink JR, Yewdell JW.** Nuclear translation visualized by ribosome-bound nascent chain puromycylation. *J Cell Biol*. 2012;**197**(1):45-57.

166. **Obrig TG, Culp WJ, McKeehan WL, Hardesty B.** The mechanism by which cycloheximide and related glutarimide antibiotics inhibit peptide synthesis on reticulocyte ribosomes. *J Biol Chem*. 1971;**246**(1):174-181.

167. **Schneider-Poetsch T, Ju J, Eyler DE, Dang Y, Bhat S, Merrick WC, Green R, Shen B, Liu JO.** Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat Chem Biol*. 2010;**6**(3):209-217.

168. **Grollman AP.** Inhibitors of protein biosynthesis. II. Mode of action of anisomycin. *J Biol Chem*. 1967;**242**(13):3226-3233.

169. **Hansen JL, Moore PB, Steitz TA.** Structures of five antibiotics bound at the peptidyl transferase center of the large ribosomal subunit. *J Mol Biol*. 2003;**330**(5):1061-1075.

170. **Carter MS, Doscow J, Morris P, Li S, Nhim RP, Sandstedt S, Wilkinson MF.** 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.* 1995;**270**(48):28995-29003.

171. **Colombo B, Felicetti L, Baglioni C.** Inhibition of protein synthesis in reticulocytes by antibiotics. I. Effects on polysomes. *Biochim Biophys Acta*. 1966;**119**(1):109-119.

172. **Kappen LS, Goldberg IH.** Analysis of the two steps in polypeptide chain initiation inhibited by pactamycin. *Biochemistry*. 1976;**15**(4):811-818.

173. **Brodersen DE, Clemons WM Jr, Carter AP, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V.** The structural basis for the action of the

antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell*. 2000;**103**(7):1143-1154.

174. **Palombella VJ, Rando OJ, Goldberg AL, Maniatis T.** The ubiquitin-proteasome pathway is required for processing the NF-kappa B1 precursor protein and the activation of NF-kappa B. *Cell*. 1994;**78**(5):773-785.

175. **Pierce JW, Schoenleber R, Jesmok G, Best J, Moore SA, Collins T, Gerritsen ME.** Novel inhibitors of cytokine-induced IkBα phosphorylation and endothelial cell adhesion molecule expression show anti-inflammatory effects *in vivo*. *J Biol Chem*. 1997;**272**(34):21096-103.

176. **Allfrey VG, Mirsky AE, Osawa S.** Protein synthesis in isolated cell nuclei. *J Gen Physiol*. 1957;**40**(3):451-490.

177. **Stevens A, Maupin MK.** 5,6-Dichloro-1-beta-D-ribofuranosylbenzimidazole inhibits a HeLa protein kinase that phosphorylates an RNA polymerase II-derived peptide. *Biochem Biophys Res Commun*. 1989;**159**(2):508-515.

178. **Kaida D, Motoyoshi H, Tashiro E, Nojima T, Hagiwara M, Ishigami K, Watanabe H, Kitahara T, Yoshida T, Nakajima H, Tani T, Horinouchi S, Yoshida M.** Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat Chem Biol*. 2007;**3**(9):576-583.

179. **Stinski MF, Roehr TJ.** Activation of the major immediate early gene of human cytomegalovirus by cis-acting elements in the promoter-regulatory sequence and by virus-specific trans-acting components. *J Virol*. 1985;**55**(2):431-441.

180. **Pestka S.** Inhibitors of ribosome functions. *Annu Rev Microbiol*. 1971;**25**:487-562.

181. **Morgan DO.** Cyclin-dependent kinases: engines, clocks, and microprocessors. *Annu Rev Cell Dev Biol*. 1997;**13**:261-291.

182. **Corden JL, Cadena DL, Ahearn JM Jr, Dahmus ME.** A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc Natl Acad Sci U S A*. 1985;**82**(23):7934-7938.

183. **Cramer P, Bushnell DA, Kornberg RD.** Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science*. 2001;**292**(5523):1863-1876.

184. **Cadena DL, Dahmus ME.** Messenger RNA synthesis in mammalian cells is catalyzed by the phosphorylated form of RNA polymerase II. *J Biol Chem*. 1987;**262**(26):12468-12474.

185. **Komarnitsky P, Cho EJ, Buratowski S.** Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev*. 2000;**14**(19):2452-2460.

186. **Egloff S, Dienstbier M, Murphy S.** Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends Genet*. 2012;**28**(7):333-341.

187. **Golas MM, Sander B, Will CL, Lührmann R, Stark H.** Molecular architecture of the multiprotein splicing factor SF3b. *Science*. 2003;**300**(5621):980-984.

188. **Lerner MR, Steitz JA.** Antibodies to small nuclear RNAs complexed with proteins are produced by patients with systemic lupus erythematosus. *Proc Natl Acad Sci U S A*. 1979;**76**(11):5495-5499.

189. **Lerner MR, Boyle JA, Mount SM, Wolin SL, Steitz JA.** Are snRNPs involved in splicing? *Nature*. 1980;**283**(5743):220-224.

190. **Nelson KK, Green MR.** Mammalian U2 snRNP has a sequence-specific RNA-binding activity. *Genes Dev*. 1989;**3**(10):1562-1571.

191. **Coons AH.** The beginnings of immunofluorescence. *J Immunol*. 1961;**87**:499-503.

192. **Brandtzaeg P.** The increasing power of immunohistochemistry and immunocytochemistry. *J Immunol Methods*. 1998;**216**(1-2):49-67.

193. **Bolte S, Cordelières FP.** A guided tour into subcellular colocalization analysis in light microscopy. *J Microsc*. 2006;**224**(3):213-32.

194. **Söderberg O, Leuchowius KJ, Gullberg M, Jarvius M, Weibrecht I, Larsson LG, Landegren U.** Characterizing proteins and their interactions in cells and tissues using the *in situ* proximity ligation assay. *Methods*. 2008;**45**(3):227-232.

195. **Gullberg M, Göransson C, Fredriksson S.** Duolink – "In-cell Co-IP" for visualization of protein interactions *in situ*. *Nat Methods*. 2011;**8**(11): Application note.

196. **Mason DW, Williams RF.** Kinetics of antibody reactions and the analysis of cell surface antigens. 1986; In: Weir DM *et al.* (eds) *Handbook of experimental immunology*, vol 1. Immunochemistry. Blackwell Scientific Publications, Oxford pp **38**:1-34.

197. **Pombo A, Jackson DA, Hollinshead M, Wang Z, Roeder RG, Cook PR.** Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. *EMBO J*. 1999;**18**(8):2241-2253.

198. **Iborra FJ, Pombo A, Jackson DA, Cook PR.** Active RNA polymerases are localized within discrete transcription 'factories' in human nuclei. *J Cell Sci*. 1996;**109**(6):1427-1436.

199. **Wansink DG, Schul W, van der Kraan I, van Steensel B, van Driel R, de Jong L.** Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *J Cell Biol*. 1993;**122**(2):283-293.

200. **Wansink DG, Nelissen RL, de Jong L.** In vitro splicing of pre-mRNA containing bromouridine. *Mol Biol Rep*. 1994;**19**(2):109-113.

201. **Lawrence JB, Singer RH.** Quantitative analysis of *in situ* hybridization methods for the detection of actin gene expression. *Nucleic Acids Res*. 1985;**13**(5):1777-1799.

202. **Zenklusen D, Singer RH.** Analyzing mRNA expression using single mRNA resolution fluorescent *in situ* hybridization. *Methods Enzymol*. 2010;**470**:641-659.

203. **Femino AM, Fay FS, Fogarty K, Singer RH.** Visualization of single RNA transcripts *in situ*. *Science*. 1998;**280**(5363):585-590.

204. **Raj A, van den Bogaard P, Rifkin SA, van Oudenaarden A, Tyagi S.** Imaging individual mRNA molecules using multiple singly labeled probes. *Nat Methods*. 2008;**5**(10):877-879.

205. **van de Corput MP, Grosveld FG.** Fluorescence *in situ* hybridization analysis of transcript dynamics in cells. *Methods*. 2001 Sep;**25**(1):111-118.

206. **Wada Y, Ohta Y, Xu M, Tsutsumi S, Minami T, Inoue K, Komura D, Kitakami J, Oshida N, Papantonis A, Izumi A, Kobayashi M, Meguro H, Kanki Y, Mimura I, Yamamoto K, Mataki C, Hamakubo T, Shirahige K, Aburatani H, Kimura H, Kodama T, Cook PR, Ihara S.** A wave of nascent transcription on activated human genes. *Proc Natl Acad Sci U S A*. 2009;**106**(43):18357-18361.

207. **Chaumeil J, Augui S, Chow JC, Heard E.** Combined immunofluorescence, RNA fluorescent *in situ* hybridization, and DNA fluorescent *in situ* hybridization to study chromatin changes, transcriptional activity, nuclear organization, and X-chromosome inactivation. *Methods Mol Biol*. 2008;**463**:297-308.

208. **Moore MJ.** Nuclear RNA turnover. *Cell*. 2002;**108**(4):431-434.

209. **Larkin JD, Papantonis A, Cook PR.** Promoter type influences transcriptional topography by targeting genes to distinct nucleoplasmic sites. *J Cell Sci*. 2013;**126**(9):2052-2059.

210. **Kapuscinski J.** DAPI: a DNA-specific fluorescent probe. *Biotech Histochem*. 1995;**70**(5):220-233.

211. **Larkin JD, Cook PR.** Maximum precision closed-form solution for localizing diffraction-limited spots in noisy images. *Opt Express*. 2012;**20**(16):18478-18493.

212. **Schmidt EK, Clavarino G, Ceppi M, Pierre P.** SUnSET, a nonradioactive method to monitor protein synthesis. *Nat Methods*. 2009;**6**(4):275-277.

213. **Williams AF, Barclay AN, Clark SJ, Paterson DJ, Willis AC.** Similarities in sequences and cellular expression between rat CD2 and CD4 antigens. *J Exp Med*. 1987;**165**(2):368-380.

Chapter 3

Results

3.1 Introduction

As discussed in **Chapter 1**, the best evidence supporting the idea that some translation occurs in nuclei came from the use of isolated nuclei and/or permeablized cells, but this can be criticized on the grounds that the nuclear interior became contaminated with cytoplasmic ribosomes during isolation and/or permeabilization. To overcome this criticism, I used three different approaches to localize sites of protein synthesis in living cells:

1. Cells were pulse-labelled by growth in L-azidohomoalanine (Aha), an analogue of the essential amino acid, methionine; this analogue is incorporated by the ribosome into the growing peptide chain. After fixing cells, 'click' chemistry is used to conjugate fluors on to the (nascent) Aha-tagged peptides so that they can be localized by fluorescence microscopy.
2. Cells were pulse-labelled by growth in puromycin, a translational inhibitor; this inhibitor is a structural mimic of aminoacyl-tRNA and it is incorporated by the ribosome into nascent peptides to terminate synthesis. After fixation, the now puromycylated peptides are immuno-localized using an anti-puromycin antibody.
3. Cells were transfected with a multi-copy plasmid encoding the non-nuclear protein – CD2 – under the control of an inducible promoter. On induction, newly-made CD2 was immuno-localized using anti-CD2 antibody.

Use of the three different approaches indicated that significant amounts of nascent and/or newly-made peptides are found in nuclei, consistent with the idea that some translation occurs in nuclei.

3.2 Approach 1: Aha incorporation

Aha is an analogue of methionine, the essential amino acid (**Fig. 3.1**). When cells are starved of methionine and Aha added, the analogue is rapidly incorporated into nascent peptides instead of methionine. As this analogue possesses a reactive azide group, 'click' chemistry can then be used to conjugate alkyne-containing molecules on to the Aha-containing peptides rapidly and efficiently. The product is a 1,4-disubstituted 1,2,3-triazole, and the reaction is usually carried out in the presence of Cuprous/Cu(I) ions at ambient temperatures (i.e., 20-40˚C). ['Click' chemistry includes various Cu(I)-catalyzed azide-alkyne cyclo-addition (CuAAC) reactions that have been optimized for high yields in the presence of multi-dentate ligands for Cu(I) ions [214].]

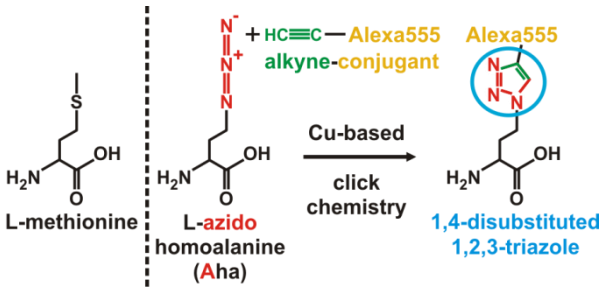


Fig. 3.1. Click chemistry. L-azidohomoalanine (Aha) is an analogue of the essential amino acid, methionine. Aha has an azide group (red) instead of the usual sulpho-methyl group of methionine. After growth in the absence of methionine to reduce internal pools, Aha is rapidly incorporated into nascent peptides. After fixation, the azido group in Aha-containing peptides can then react (in the presence of Cu(I) ions at room temperature) with an added alkyne group to form a 1,4-disubstituted 1,2,3-triazole. This Cu(I)-catalyzed azide-alkyne cyclo-addition (CuAAC) reaction provides one example of 'click chemistry'. If the alkyne is tagged with a fluorophore (in this case, Alexa555), then Aha-containing peptides can be localized by fluorescence microscopy.

This general approach is currently being used to label newly-made proteins (and – using other precursors carrying an azide – nucleic acids, lipids, and sugars) [215]. The synthesis of newly-made proteins is being explored in two common ways [216]. The first involves the downstream analysis of proteomes using a biotin-tagged alkyne that can be 'clicked' on to the newly-made azide-bearing peptides; then the biotin-tagged peptides can be purified rapidly and efficiently by affinity purification, and analysed by mass spectrometry (a technique known as BONCAT – biorthogonal non-canonical amino-acid tagging) [217]. The second involves 'clicking' on alkyne-carrying fluor, so that Aha-bearing peptides can be localized by fluorescence (a technique known as FUNCAT – fluorescent non-canonical amino-acid tagging) [218-220]. Both approaches are in routine use. Thus, Aha is incorporated without bias into all kinds of protein – whether endogenous [217] or over-expressed [221] – by bacteria [222,223], viruses [224,225], various cultured and primary cells [218,220,226], and whole organisms [227]. Aha-substitution does not affect cell viability [217], virus replication [225], and transcription in *Drosophila* S2 cells [226]. The tagged proteins are not specifically targeted to ubiquitination-mediated decay [217], they are not prone to changes in secondary structure or overall folding patterns [228], they localize to their usual cellular locations (in the case of a bacterial surface protein) [222], their expression responds normally to biological stimuli [227,229], and they are functionally active (e.g., viral particles containing tagged proteins retain their infectivity) [224]. In brief, all reports indicate that Aha-bearing proteins behave like their methionine-containing counterparts.

Here, I use FUNCAT to localize nascent peptides in the nucleus and cytoplasm. I give methionine-starved cells a short-pulse of Aha, fix, 'click' an alkyne covalently linked to a fluor like Alexa555 on to any Aha-containing (nascent) peptides, image cells

in a wide-field or ‘confocal’ microscope, and determine the amount of fluorescence in the nucleus and cytoplasm.

3.2.1 Aha is incorporated into nuclei

If cells are grown in the absence of methionine for 15 min at 37˚C (to deplete endogenous pools of methionine), pulsed with Aha for 2 min at 37˚C, plunged into ice-cold fixative to stop translation quickly, and a fluor (i.e., Alexa555) ‘clicked’ on to incorporated Aha, fluorescence can be seen in both cytoplasm and nucleus (**Fig. 3.2A**,**B**). Pre-incubation for 2 h in anisomycin – an inhibitor of translation that acts on the 80S ribosome at the peptidyl-transferase center [230,231] – reduces Aha incorporation (**Fig. 3.2B**-**D**); this is consistent with Aha being incorporated by the ribosome into newly-made peptides in both cytoplasm and nucleus. [The anisomycin concentration and pre-incubation time applied here are routinely used to test whether the RNA-quality control mechanism is tightly coupled to translation (e.g., see **Ref** [232]); as we will see, shorter pre-incubation times (i.e., 15-30 min) will be used in critical experiments.]

The approach described above involved methionine starvation for 15 min. As imposing such a stress on cells might have unforeseen results, the experiment was repeated without prior starvation, and 0.2 mM methionine (the concentration normally found in the culture medium) was present both before and during the Aha pulse; incorporated Aha was again found in nuclei (**Fig. 3.2E**). This indicates the nuclear signal does not result from any stress induced by methionine starvation.

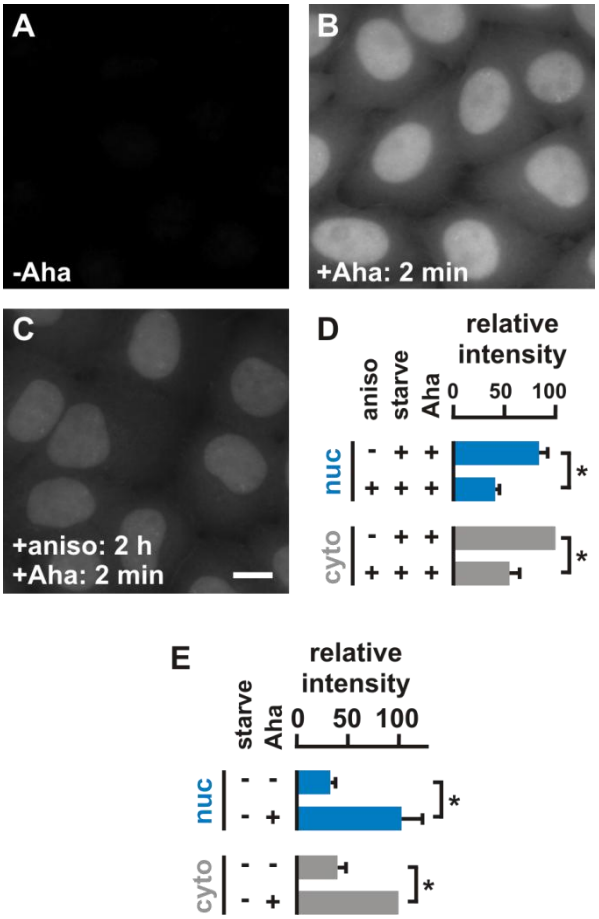


Fig. 3.2. The methionine analogue, Aha, is incorporated into nuclei. HeLa cells were generally starved of methionine for 15 min, and then pulsed ± 2 mM Aha for 2 min; after fixation, incorporated Aha was labelled with Alexa555 using 'click' chemistry, DNA counterstained with DAPI, images collected using a wide-field microscope (typical ones are shown in A-C), and the intensities of Alexa 555 fluorescence in the cytoplasm (cyto) and nucleus (nuc) measured.

A. Omission of Aha yields background fluorescence.

B. Inclusion of Aha yields both cytoplasmic and nuclear signal.

C. Adding 100 µg/ml anisomycin for 2 h before adding Aha reduces both cytoplasmic and nuclear signal (bar: 10 µm).

D. After subtracting nuclear or cytoplasmic background seen in the absence of Aha, intensities (± SD) seen with Aha are expressed relative to the value found in cytoplasm untreated with anisomycin (*: $P < 0.0001$, Student's two-tailed *t* test, $n = 20$ cells). Anisomycin reduces both cytoplasmic and nuclear signal.

E. Aha incorporation does not depend on methionine starvation. Intensities (± SD) seen with Aha are expressed relative to the value found in the Aha-pulsed cytoplasm; this is the only case shown in all Figures

where nuclear and cytoplasmic background seen in the absence of Aha were not subtracted. (*: $P < 0.0001$, Student's two-tailed t test, $n = 20$ cells). Nuclear signal is 3 times higher than the (nuclear) background (-Aha).

3.2.2 Aha-tagged peptides turn over rapidly

Although no unforeseen effects have been reported with the commercially-available Aha-tagging kit used here, it remained formally possible there might be some (e.g., Aha might bind preferentially to some unknown nuclear site, or the 'click' chemistry might conjugate the alkyne-tagged flour on to some unknown nuclear ligand instead of the expected azide group in Aha-containing peptides, or the rapidly-degraded polypeptide [233] products from newly-synthesized proteins may have higher affinity towards chromatin [234]). Therefore, a pulse-chase experiment was performed to test whether the nuclear signal resulted from a genuine metabolic incorporation of Aha into nascent peptides, and if these peptides preferably attached to any unknown nuclear ligands. Cells were starved of methionine, pulsed for 2 min with Aha, the medium replaced with normal growth medium containing methionine, and cells regrown for up to 5 min at 37˚C or 4˚C. After 'clicking' on the fluor, nuclei exposed to just the Aha pulse gave the now-expected nuclear and cytoplasmic signals (**Fig. 3.3A**,**D**). Both signals disappeared after a chase at 37˚C (**Fig. 3.3B**,**D**), but not at 4˚C (**Fig. 3.3C**,**D**); both fell during the chase with similar half-lives (**Fig. 3.3E**). These results are consistent with Aha being incorporated into nascent peptides that are turned over during a chase at 37˚C by normal cellular processes (which are slowed at 4˚C). They confirm earlier work showing that a considerable fraction of newly-made peptides in both the cytoplasm and nucleus are rapidly degraded [233,235,236]. Moreover, if the 'click' chemistry attaches the fluor to some unexpected nuclear target, that target must be degraded to a degree that depends on

the temperature previously used during the chase, and thus should behave differently from any cytoplasmic target. However, the disappearance of Aha-containing peptides in the nucleus during the chase at 37˚C and 4˚C occurs with the same kinetics as seen with those in the cytoplasm; therefore, it seems unlikely that the 'click' chemistry attaches the fluor to unforeseen targets found uniquely in nuclei. If Aha-containing rapidly-degraded peptides bind to such nuclear targets (chromatin can protect against degradation by amino-peptidases [234]), then nuclei should lose less signal than the cytoplasm during the chase at 37˚C; however, there is no significant difference in the rate of loss of signal in the nucleus and cytoplasm.

The proteasome is responsible for degrading many proteins in both nucleus and cytoplasm [237,238], and proteasomal subunits are closely associated with sites of transcription and have been implicated in most steps of transcription [239,240], through both their non-proteolytic [241] and proteolytic [242] functions. Therefore, I next examined what effect inhibiting the proteasome might have on nuclear and cytoplasmic signals; cells were treated with or without the proteasomal inhibitor – MG132 [243], cells starved of methionine, pulsed with Aha, and nuclear and cytoplasmic signal compared after 'clicking' on the fluor. Pre-treatment with MG132 increases both nuclear and cytoplasmic signal (**Fig. 3.3F**). This increase is consistent with the following scenario. A significant fraction of nuclear signal results from nuclear ribosomes proof-reading the nascent transcript; if mis-splicing generates a premature stop codon (PTC), then the proof-reading ribosome will detect the PTC and trigger the RNA-quality surveillance mechanism – NMD (nonsense-mediated mRNA decay) [244,245], with the consequence that the truncated peptides are degraded by nearby proteasomes [246,247]. Then, inhibiting the proteasome should prevent this degradation and increase nuclear signal. Inhibition of proteasomal degradation would also explain the increase in the cytoplasmic

signal (**Fig. 3.3F**). These results do not exclude the possibility that non-proteasomal pathways may also be involved.

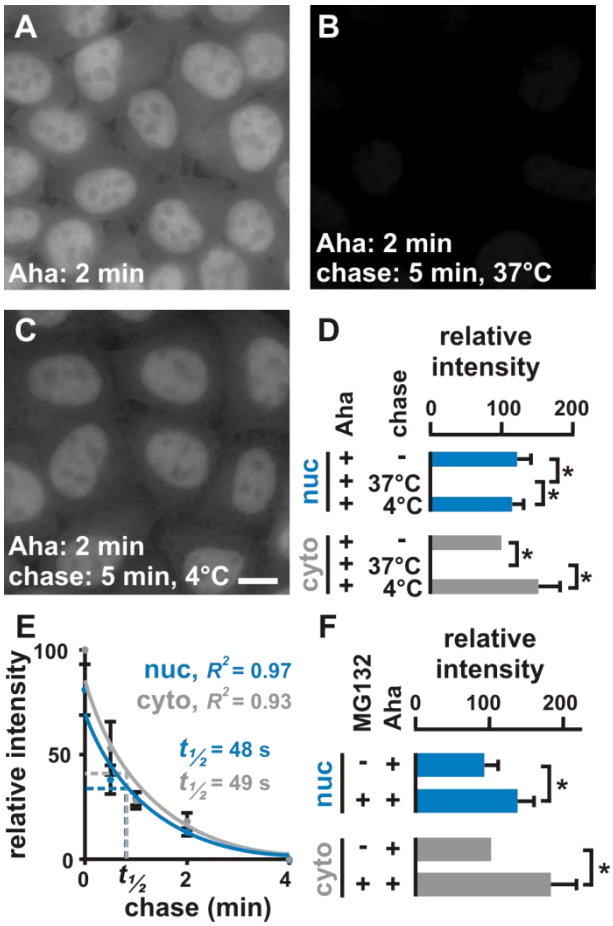


Fig. 3.3. Aha-tagged nuclear peptides are degraded quickly (half-life 37 s) by the proteasome. HeLa cells were starved of methionine for 15 min, pulsed for 2 min with 2 mM Aha, and chased for 0-5 min with 0.2 mM methionine (without Aha) at 37˚C or 4˚C. After fixation, incorporated Aha was labelled with Alexa555 using 'click' chemistry, DNA counterstained with DAPI, images collected using a wide-field microscope (typical ones are shown in A-C), and the intensities of Alexa555 in the cytoplasm (cyto) and nucleus (nuc) measured.

A. A 2-min Aha pulse yields both nuclear and cytoplasmic signal.

B. A 5-min chase with methionine at 37˚C then removes essentially all signal.

C. In contrast, significant signal remains after a 5-min chase with methionine at 4˚C (bar: 10 µm).

D. 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).

E. After chasing with methionine at 37˚C, relative intensities were determined (as in D); the half-lives of cytoplasmic and nuclear signal are indicated (exponential curves were fitted to the data; R^2 is the square of the Pearson product-moment correlation coefficient, and the high values indicate a good fit).

F. In some cases 100 μg/ml MG132 (a proteasome inhibitor) was added for 2 h before the pulse, and the chase was omitted; the inhibitor increases both nuclear and cytoplasmic signal (*: $P < 0.0001$, Student's two-tailed *t* test, $n = 20$ cells).

3.2.3 Nuclear Aha-bearing peptides are unlikely to be imported from the cytoplasm

In the living human cell, a ribosome polymerizes ~5 amino acids per second [248], and so completes a typical human protein with ~340-415 amino acid residues (in *Homo sapiens*, median protein length is ~340-415 amino acid residues [249,250], with a mean of ~475-550 [249,251]) in ~65-83 seconds. Therefore, it is possible – although unlikely – that the nuclear signal seen in **Figure 3.2** resulted from cytoplasmic synthesis followed by rapid import of newly-made peptides into nuclei during the 120-s pulse. Therefore, methionine-starved cells were pulsed for only 5 s with Aha – by dipping the coverslip on which they were growing into Aha-containing medium for 5 s, and then immediately plunging the coverslip into ice-cold fixative. After 'clicking' on the fluor, faint signal is seen in the nucleus and fainter signal in the cytoplasm (**Fig. 3.4B**,**C**,**D**). Both signals are again sensitive to anisomycin (**Fig. 3.4C**). Assuming that a typical protein contains ~2% methionine [249,252], only 25 amino acids will be incorporated into a typical protein during the 5-s pulse, of which only 0.5 amino acids will be a methionine or Aha. As a result, only ~50% of all growing peptides will contain one Aha moiety. Given that the speed of translation in human cells is ~5 amino acids per second [248], and the size of a typical human protein is ~400 amino acids [249,250], only 1/16 of all the peptides formed during 5 s can leave their synthesis sites (because only such peptides whose last 25 amino acids are being added during this 5 s can leave their

synthesis sites, and 5x5/400 = 1/16); then, 95% of peptides made during this 5-s pulse remain at their synthesis sites. Therefore, it is unlikely that nuclear signal could arise from peptides made in the cytoplasm.

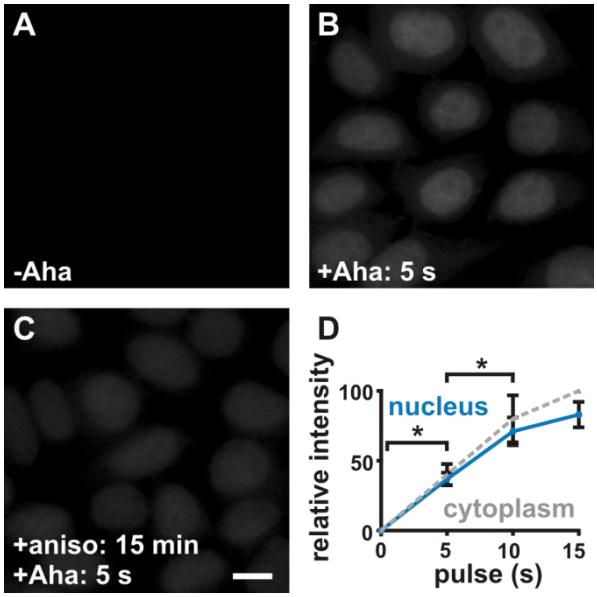


Fig. 3.4. During pulses of only 5-15 s, Aha is incorporated into nuclei. HeLa cells were starved of methionine for 30 min, and then pulsed with 2 mM Aha for 0-15 s; after fixation, incorporated Aha was labelled with Alexa555 using 'click' chemistry, DNA counterstained with DAPI, images collected using a wide-field microscope (typical ones are shown in A-C), and the intensities of Alexa555 in the cytoplasm and nucleus measured.

A. Omission of Aha yields background fluorescence.

B. A 5-sec pulse of Aha yields some cytoplasmic and nuclear signal.

C. Adding 100 µg/ml anisomycin for 15 min before a 5-s Aha pulse reduces the weak signals in both cytoplasm and nucleus (bar: 10 µm).

D. After subtracting background (here a 0-s pulse), intensities (± SD) seen with Aha are expressed relative to the value found in the cytoplasm after a 15-s pulse (*: $P < 0.0001$, Student's two-tailed t test, $n = 20$ cells); signal increases in both nucleus and cytoplasm. Adding anisomycin for 15 min before a 5-s pulse (as in C) reduces significantly the relative intensity in the nucleus by 54% , and in the cytoplasm by 85% ($P < 0.0001$, Student's two-tailed t test, $n = 20$ cells).

3.3 Approach 2: Puromycin incorporation

Protein synthesis is known to be most efficient when more than one ribosome (i.e., a polyribosome or polysome) tandemly translate a transcript [253,254]. Polysomes have only been observed in the cytoplasm. Puromycin is a translation inhibitor that mimics aminoacyl-tRNA [255]; it is incorporated at the ribosomal A-site into the C-terminus of the nascent peptide chain *via* a peptide bond (amide bond) [256,257]. As it lacks a carboxyl group, synthesis of the peptide chain terminates prematurely. Also, since puromycin has only a weak affinity for the A-site of the 80S ribosome (K_d = 2.5 x 10^{-5} M) [258] compared to the amino acyl-tRNA (K_d = 7.9 x 10^{-8} [259]), the resulting puromycylated (nascent) peptide soon dissociates from the ribosome [260] (presumably in much less than 1 min, by which time exchange of ribosomal subunits is already effected [261]). However, except in the presence of high concentrations of puromycin (>25 µM) and a sufficient energy supply, engaged ribosomes continue synthesizing new puromycylated peptides without reinitiating [260]. Eventually, however, polysomes dissociate into monosomes, which then dissociate from the mRNA into their constituent sub-units.

This dissociation of puromycylated (nascent) peptide chains from the ribosome, and the subsequent dissociation of the engaged ribosome from the transcript, can be decreased by slowing down and retaining the translating poly-ribosomes using another inhibitor of translation – cycloheximide. Cycloheximide acts by associating with the E-site of the 80S ribosome and blocks eEF2-GTP-dependent translocation of the amino-acylated-tRNA from A-site to P-site, and in the process stabilizes polysomes and apparently stimulates their formation [262,263]. The combined use of the two inhibitors provides a small window of opportunity when puromycylated peptides remain associated

with actively-translating ribosomes; then, the puromycylated peptides can be immuno-localized using an antibody targeting puromycin [264].

3.3.1 Puromycylated (nascent) peptides are found in nuclear foci

Cells are treated without or with cycloheximide (15 min, 37˚C), grown in puromycin for 5 s at 37˚C, fixed, and puromycylated peptides indirectly immuno-labelled. In the absence of cycloheximide, puromycylated peptides can be seen in discrete foci in both the cytoplasm and nuclei (**Fig. 3.5A**). In the presence of cycloheximide, signal is reduced in both nucleus and cytoplasm (**Fig. 3.5B**), but more prominently in the latter (**Fig. 3.5C**).

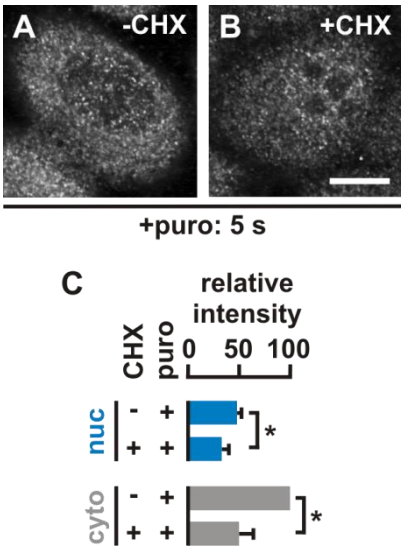


Fig. 3.5. Puromycylated (nascent) peptides are found in nuclear foci. HeLa cells were treated without or with 100 µg/ml cycloheximide for 15 min (to slow ribosomes without stopping them completely), pulsed with 91 µM puromycin for 5 s, fixed, puromycylated peptides indirectly immuno-labelled, DNA counterstained with DAPI, and images collected using a confocal microscope (typical optical sections through the centre of nuclei are shown in A and B).

A. Without cycloheximide pre-incubation, puromycylated peptides are seen in foci in both cytoplasm and nucleus.

B. With cycloheximide pre-incubation, puromycylated peptides are seen in foci in both cytoplasm and nucleus, but the signal has been reduced in both compartments (bar: 10 μm).

C. After subtracting background, intensities in individual sections are expressed relative to the value found in the cytoplasm after a 5-s puromycin pulse without cycloheximide pre-incubation. Cycloheximide reduces both nuclear and cytoplasmic signal significantly (*: $P < 0.0001$, Student's two-tailed *t* test, $n = 20$ cells), but the drug has a stronger effect on the cytoplasmic signal.

As puromycylated peptides are unstably associated with ribosomes, I next used longer puromycin pulses; cells were pre-treated with cycloheximide to slow ribosomes, and cells grown in puromycin for 5–60 s (**Fig. 3.6**). After 5 s, foci were seen in both nucleus and cytoplasm (**Fig. 3.6A**,**B**). After 30 s, the nuclear signal decreases and becomes more diffuse (**Fig. 3.6C**). After 60 s, both nuclear and cytoplasmic signals rise, with most being concentrated in the peri-nuclear region of the cytoplasm (**Fig. 3.6D**,**E**). This is consistent with previous observations showing that puromycylated peptides accumulate at protein exit-sites on the smooth endoplasmic reticulum (SER) prior to secretion through the Golgi apparatus [265,266].

These fluctuations in signal after pretreatment with cycloheximide (**Fig. 3.6E**) can be interpreted as follows (**Fig. 3.6F**), assuming that monosomes in the nucleus proof-read the nascent transcript, whilst polysomes in the cytoplasm engage in bulk protein production. Thus, after 5 s, puromycylated peptides are found mainly at sites active in protein synthesis, in both nucleus and cytoplasm. But after 30 s, many puromycylated peptides dissociate from nucleoplasmic monosomes and cytoplasmic polysomes [260]; this creates a more uniform distribution in both compartments. [Puromycin induces polysomes attached to the RER [267] to detach (since translating polysomes are anchored *via* extending peptides [268-270]) and dissociate into monosomes [271], ribosomes to exchange their subunits [261,272], and the small ribosomal subunit to disengage from its transcript [273].] After 60 s, some puromycylation continues (mainly due to polysomes in

the cytoplasm), and – as more and more puromycylated peptides detach – they accumulate at exit sites in the SER [274,275]. As a result, signal builds up around the nucleus. In the case involving no pre-treatment with cycloheximide and a 5-s pulse (as in **Fig. 3.5A**,**B**), the cytoplasmic signal arises simply because uninhibited polysomes are active on the RER [267]. Note that cycloheximide reduces the nuclear puromycylated peptides significantly (although the effect is small), which might be produced by monosomes [276-279] in transcription factories [280]; however, the large reduction in the cytoplasmic signal with cycloheximide inevitably makes the nuclear signal appear more prominent (compare **Fig. 3.5A** with **B**).

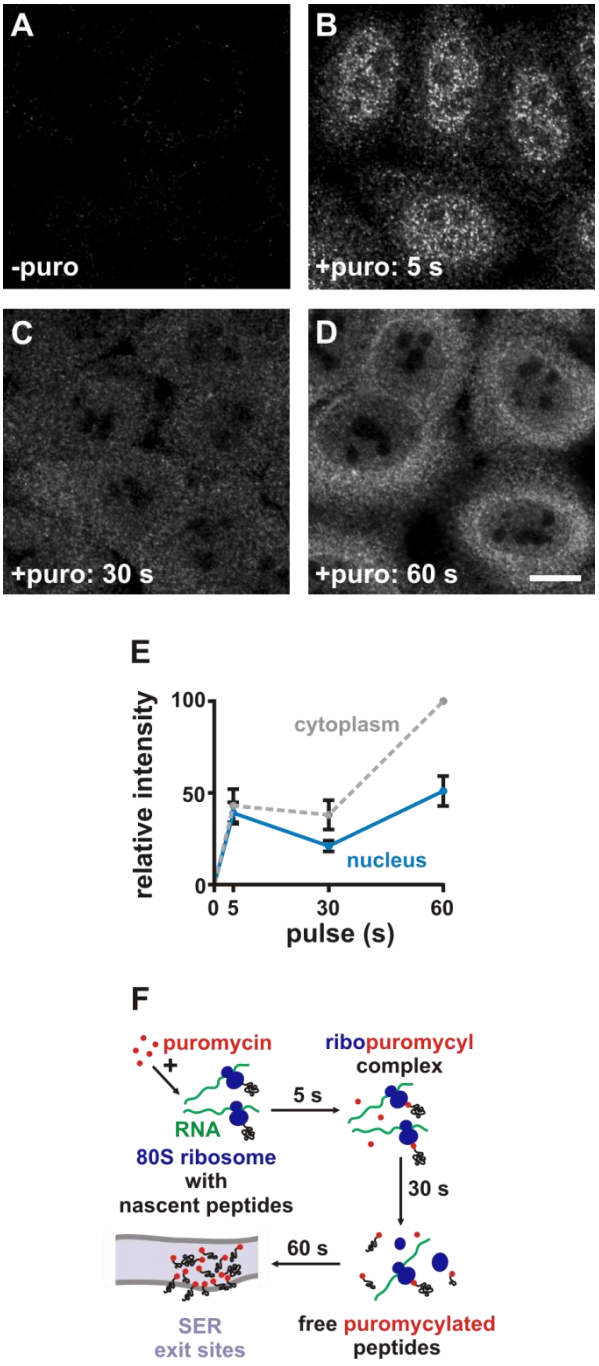


Fig. 3.6. Spatial dynamics of puromycylated (nascent) peptides. HeLa cells were treated with 100 µg/ml cycloheximide for 15 min (to slow ribosomes without stopping them completely), pulsed with 91 µM puromycin for 0-60 s, fixed, puromycylated peptides indirectly immuno-labelled, DNA counterstained with DAPI, and images collected using a confocal microscope (typical optical sections through the centre of nuclei are shown in A-D).

A. With no puromycin pulse, background levels of signal are seen.

B. With a 5-s pulse, puromycylated peptides are seen in foci in both cytoplasm and nucleus.

C. With a 30-s pulse, cytoplasmic and nuclear signals have become more similar, and both are more diffuse.

D. With a 60-s pulse, the cytoplasmic signal increases, especially in the peri-nuclear region (bar: 10 μm).

E. After subtracting background, intensities (± SD; *n* = 20 cells) in individual sections are expressed relative to the value found in the cytoplasm after a 60-s pulse; nuclear signal declines sharply after 5 s, while cytoplasmic signal increases progressively.

F. Interpretation. During a 5-s puromycin pulse, the nascent peptide made by a translating ribosome (slowed by cycloheximide) becomes puromycylated (red spot), and signal is seen in discrete foci (where many active ribosomes are found). After a 30-s pulse, some puromycylated peptides disengage from ribosomes, so the focal signal becomes more diffuse. After 60 s, released and tagged peptides become concentrated in the cytoplasm around the nucleus – probably at exit sites in the smooth endoplasmic reticulum (SER) – prior to export to the Golgi.

David *et al*. [266] observed significant perinuclear and nucleolar labelling after performing an analogous experiment (in the presence of cycloheximide) with two significant differences: cells were permeabilized before adding puromycin, and the puromycin pulse was for 5 min. Therefore I examined the effects of a longer puromycin pulse in the absence of permeabilization before fixation; growth in puromycin for 5 min led to strong labelling within and around the nucleus (**Fig. 3.7**) – and this labelling was sensitive to anisomycin (**Fig. 3.7C**,**D**). These results are consistent with synthesis in both nucleus and cytoplasm, and accumulation of released puromycylated peptides at exit sites on the SER. They are like those of David *et al.* [266] except that I saw no nucleolar labelling (**Fig. 3.7B**); this difference is discussed in **Section 4.6**.

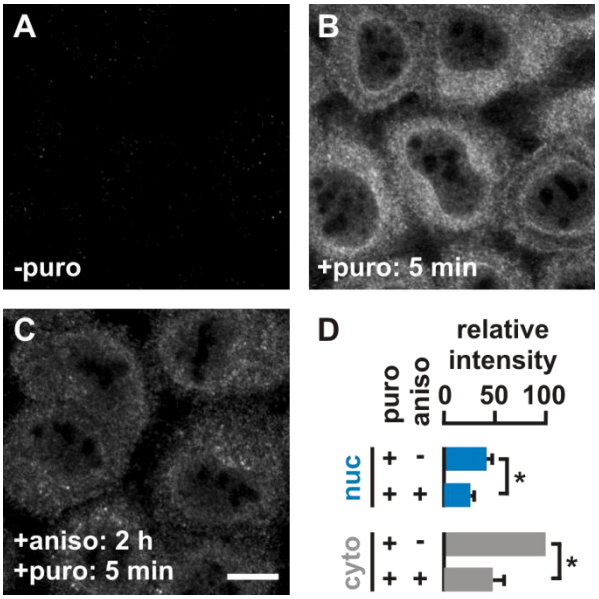


Fig. 3.7. Puromycin is incorporated into nascent peptides. HeLa cells were treated with 100 µg/ml cycloheximide for 15 min (to slow ribosomes without stopping them completely), pulsed with 91 µM puromycin for 5 min, and fixed; in some cases, 100 µg/ml anisomycin was added 2 h before fixation. Next, puromycylated peptides were indirectly immuno-labelled with Alexa555, DNA counterstained with DAPI, and images collected using a confocal (A-C; typical optical sections through the centre of nuclei are shown) or wide-field microscope (D).

A. With no puromycin pulse, background levels of signal are seen.

B. A 5-min pulse of puromycin yields significant peri-nuclear (but little nucleolar) labelling.

C. Pretreatment with anisomycin reduces both cytoplasmic and nuclear signal (bar: 10 µm).

D. After subtracting background, Alexa555 intensities (± SD) are expressed relative to the value found in the untreated (- anisomycin) cytoplasm (*: $P < 0.0001$, Student's two-tailed *t* test, $n = 20$ cells). Anisomycin reduces significantly ($P < 0.0001$) the relative intensity in both nucleus and cytoplasm.

3.3.2 Puromycylated (nascent) peptides are found close to transcription sites

The local high concentration of puromycylated peptides in nuclear foci suggested that those peptides might be made co-transcriptionally in active transcription sites or factories [280]. Therefore, I examined whether the puromycylation of the nascent nuclear peptides was sensitive to an inhibitor of RNA polymerase II, DRB (5,6-dichloro-1-β-D-

ribofuranosylbenzimidazole) [281,282]. Cells were treated without and with DRB (or anisomycin used as a control), pulsed for 5 s with puromycin, and puromycylated peptides immuno-localized; DRB decreased the intensity of signal seen in the nucleus, without much effect on the cytoplasmic signal (**Fig. 3.8Ai-iv**). This result is consistent with some translation occurring in nuclei closely coupled to transcription. If translation of RNA occurred only in the cytoplasm and DRB worked in the expected way (i.e., inhibited transcription in the nucleus), then the reduction in nuclear puromycylated peptide signal due to DRB can only be explained if splicing, mRNA export, cytoplasmic translation, and subsequent protein import into nucleus all happen in 5 s; this is highly unlikely.

The effect of DRB on the nuclear signal suggests that nuclear translation might be closely coupled to transcription. Therefore, I tried to see if I could localize the puromycylated peptides relative to the initiating form of RNA polymerase II (RNAPII) [283]. The largest catalytic subunit of this enzyme contains heptad repeats in its C-terminal domain (CTD) [284,285], and serine-5 in many of these heptads becomes phosphorylated on transcriptional initiation [286-288]; therefore, antibodies targeting this phospho-serine-5 (RNAPII-S5^{P}) are often used to mark the active polymerase [287,289,290]. Cells were incubated for 5 s in puromycin, fixed, and both puromycylated peptides and RNAPII-S5^{P} immuno-labelled with red and green fluors, respectively. Each target was found in discrete foci (**Fig. 3.8Bii-iii**), some of which overlapped (**Fig. 3.8Biv**; the inset shows one overlapping focus). However, the patterns are complicated, and it is difficult to conclude whether the yellow foci seen result from a true colocalization of red and green foci or just a chance overlap.

As patterns were so complex, I applied a general ‘super-resolution’ method (**Section 2.3.6**) to determine whether the red and green foci were closer to each other than expected by chance. The method [291] is based upon the idea of determining the distance

between the peaks in intensities given by the diffraction-limited spots in red and green channels [292-295]. The method begins by automatically identifying diffraction-limited spots within an image, goes on to localize the peak intensity within each with ~15-nm precision, and ends with a measurement of the distance from one peak in the green channel to its nearest neighbor (NN) [296] in the red channel – and *vice versa*. As each distance measurement is repeated for every focus in the image, the resulting distribution of distances can be compared with a random set (which is created for each nucleus using the same density of spots as observed in the nucleoplasmic area), and any difference analyzed statistically using the Kolmogorov-Smirnov (K-S) test [297,298]. The output of the analysis tells us that foci in the green and red channels are either distributed in a random pattern, or that they lie significantly closer to each other than expected by chance. In our case, we can then decide whether the puromycylated peptides lie closer to the initiating polymerase than expected by chance.

The distribution of such NN distances obtained from the red and green foci as seen in **Figure 3.8Biv** was plotted as a 'probability density function' (PDF) and 'cumulative distribution function' (CDF), and results compared to the PDF and CDF generated from red and green spots randomly distributed in a 250x250 pixel area with the same density as in the experiment (**Fig. 3.8Bv**,**vi**). The PDF for the experimental foci shows that ~85% of the red and green foci are within 0.75 μm of each other (this limit is selected because the PDF for the observed NN distances is higher than the PDF for the NN distances of the random set up to 0.75 μm). The CDF for the experimental foci was significantly shifted (two-sample K-S test) towards shorter NN distances, compared to the CDF for the randomly-generated foci. This suggests that nuclear puromycylated peptides lie significantly closer to sites of transcription (i.e., those containing the phosphorylated polymerase) than expected by random chance. However, the presence of many red and

green foci (and few yellow foci) in the merge in **Figure 3.8Biv** indicates that there is only incomplete overlap between puromycylated peptides and the polymerase; this apparent discrepancy could be because the 'random set' simulated for the NN-analysis might not be an ideal representation of a random distribution of molecules in the nucleus – as the nucleus has such a variable topology which may restrict the distribution of molecules. Such topological restrictions cannot be accounted for, while simulating the 'random set', and hence this may explain the significant shift of the CDF for experimental foci towards shorter NN distances. This will be discussed in **Section 4.3**.

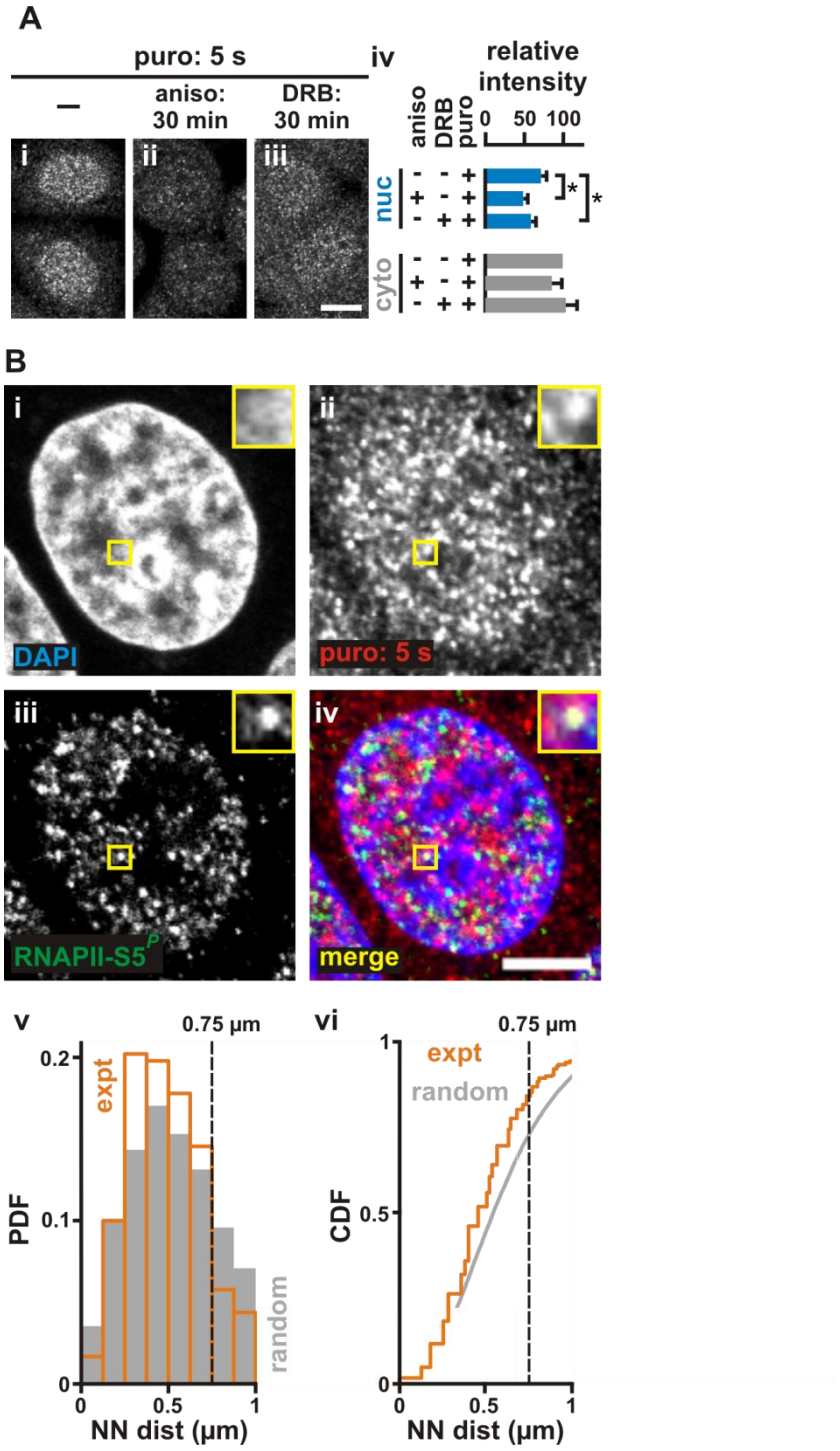


Fig. 3.8. Puromycylated (nascent) peptides are found close to transcription sites. HeLa cells were treated with 100 µg/ml cycloheximide for 15 min, pulsed with 91 µM puromycin for 5 s, fixed, puromycylated peptides indirectly immuno-labelled, DNA counterstained with DAPI, and images collected using a confocal microscope (typical optical sections through the centre of nuclei are shown).

A. Pulse-labelling for 5 s with puromycin.

i. Puromycylated peptides are seen in foci in both cytoplasm and nucleus, with the latter yielding the brightest foci.

ii. 100 µg/ml anisomycin added 30 min before puromycin reduces nuclear signal.

iii. 100 µM DRB added 30 min before puromycin reduces only the nuclear signal (bar: 10 µm).

iv. After subtracting background, intensities (± SD) seen with puromycin are expressed relative to the value found in the untreated (- anisomycin/-DRB) cytoplasm (*: $P < 0.0002$, Student's two-tailed t test, $n = 20$ cells). While adding anisomycin reduces both cytoplasmic and nuclear signal, but adding DRB reduces only nuclear signal.

B. Immuno-labelling both puromycylated peptides and (initiating) RNA polymerase II (tagged with an antibody targeting phospho-S5 in the CTD). Bar: 5 µm.

i-iv. Four views of one typical section; some puromycylated peptides (red) colocalize with polymerase (green) in discrete nuclear foci (inset shows one focus).

v,vi. Super-resolution localization shows red and green foci lie close together. Foci with a Gaussian intensity profile were selected, peaks localized with 15-nm precision, and nearest-neighbour (NN) distances calculated (from each red peak to the nearest green peak, and *vice-versa*) with 10-50 nm precision. The probability density function (PDF) and cumulative distribution function (CDF) of NN distances is compared to the simulated PDF and CDF (line includes 99%-confidence limit) of the same nuclear density of green and red foci distributed randomly (random); ~85% of green and red foci lie closer than 0.75 µm; the experimental CDF curve (expt) lies to the left of the random one, indicative of significant proximity ($P < 0.0001$, 1-sample Kolmogorov-Smirnov test, $n_{expt} = 4558$, $n_{random} = 23866$, cells = 16).

I went on to localize puromycylated peptides relative to a second marker for the active transcription complex – nascent RNA. Nascent RNA can be immuno-localized after gently permeabilizing cells in a 'physiological' buffer (PB^+; **Section 2.3.3**), and allowing still-engaged polymerases to extend their transcripts in the UTP analogue, BrUTP [299]. Cycloheximide-treated cells were grown in puromycin for 5 s, cells permeabilized gently, BrUTP added, and engaged polymerases allowed to extend their nascent transcripts by ~500 nucleotides [300]; the resulting BrRNA cannot be spliced and

does not leave transcription sites [301,302]. Cells were then fixed, and puromycylated peptides and (nascent) BrRNA indirectly immuno-labelled with red and green fluors, respectively. The puromycylated peptides and BrRNA are found in discrete nuclear foci, some of which colocalize (**Fig. 3.9B-D**). Again, given the two complicated patterns, it is difficult to infer whether the colocalization seen merely results from chance overlap of many randomly-distributed foci. Therefore, I used the same method of analysis as before in **Figure 3.8B**. The PDF for the experimental foci again shows that ~85% of the red and green foci are within 0.75 μm of each other (as before, this limit is selected because the PDF for the observed NN distances is higher than the PDF for the NN distances of the random set up to 0.75 μm). The CDF for the experimental foci was significantly shifted (two-sample K-S test) towards shorter NN distances, compared to the CDF for the randomly-generated foci (**Fig. 3.9E**,**F**). This suggests that the nuclear puromycylated peptides lie significantly closer to the sites of transcription (marked by nascent BrRNA) than expected by random chance. Once again, the presence of many red and green foci (and few yellow foci) in the merge in **Figure 3.9B** indicates that there is only incomplete overlap between puromycylated peptides and the BrRNA; this apparent discrepancy has been discussed previously in this section, and will be discussed further in **Section 4.3**.

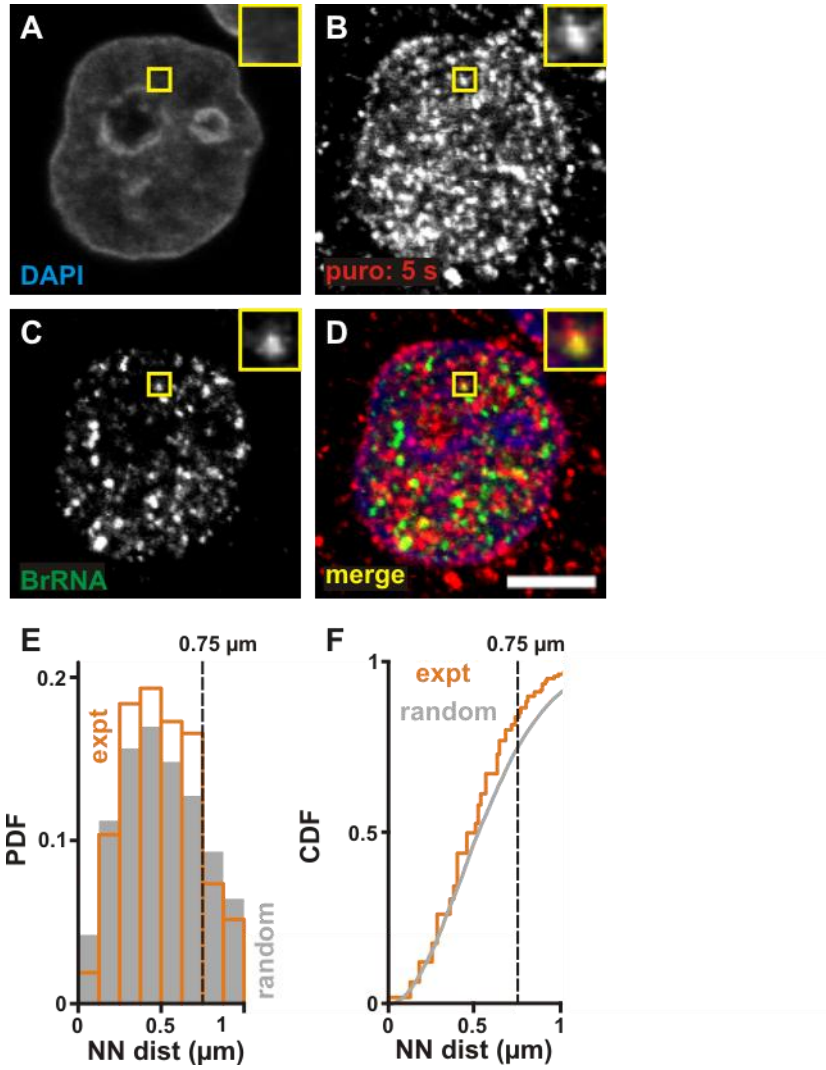


Fig. 3.9. Puromycylated peptides are found close to nascent BrRNA. HeLa cells were treated with 100 µg/ml cycloheximide for 15 min, and pulsed with 91 µM puromycin for 5 s; cells were now washed at 4°C, permeabilized gently, BrUTP added, and engaged polymerases allowed to extend their nascent transcripts by ~500 nucleotides at 33 °C for 15 min. After fixation, puromycylated peptides and (nascent) BrRNA were indirectly immuno-labelled, DNA counterstained with DAPI, and images collected using a confocal microscope (typical optical sections through the centre of nuclei are shown).

A-D. Four views of one typical section; puromycylated peptides (red; cytoplasmic signal is reduced by permeabilization) and BrRNA (green) are found in discrete nuclear foci, some of which overlap to give yellow in the merge (inset). Bar: 5 µm.

E,F. Super-resolution localization shows red and green foci lie close together. Foci with a Gaussian intensity profile were selected, peaks localized with 15-nm precision, and nearest-neighbour (NN) distances (from each red peak to the nearest green peak, and *vice-versa*) with 10-50 nm precision. The probability

density function (PDF) and cumulative distribution function (CDF) of NN distances is compared to the simulated PDF and CDF (line includes 99%-confidence limit) of the same nuclear density of green and red foci distributed randomly (random); ~85% of green and red foci lie closer than 0.75 μm; the experimental CDF curve (expt) lies to the left of the random one, indicative of significant proximity ($P < 0.0001$, 1-sample Kolmogorov-Smirnov test, $n_{expt} = 4121$, $n_{random} = 25478$, cells = 16).

3.4 Approach 3: Immuno-localizing newly-made CD2

If some translation occurs within nuclei closely coupled to transcription, some newly-made protein should be found close to the nascent transcript that encoded it. Performing such an experiment is challenging, as the optimal approach requires efficient detection of both a single nascent transcript and a single nascent peptide. Whilst RNA FISH applied with intronic probes permits the former, there remains no method that allows detection of a single nascent protein with high efficiency. Therefore, I used a multi-copy system, and I now outline the details of this system.

I began by choosing a well-characterized non-nuclear protein – one that would never be expected to be found in the nucleus if translation occurred only in the cytoplasm; I chose a cell-surface antigen. As efficient detection is critical, this antigen would be immuno-detected using a well-characterized monoclonal antibody that bound with high affinity. Therefore, I selected the rat (*Rattus norvegicus*) CD2 protein. CD2 (cluster of differentiation 2 antigen) – also known as LFA-2 (lymphocyte function-associated antigen-2) – is a cell surface antigen found in lymphocytes [303-305]; it contains a signal sequence plus a trans-membrane domain, and so is normally inserted into the endoplasmic reticulum during synthesis by a cytoplasmic ribosome. As previous work had shown that some CD2 is nevertheless found in a diffuse cytoplasmic pool [235], I wanted to further restrict the protein to a defined non-nuclear organelle. Therefore, I

replaced the N-terminal leader peptide sequence of rat CD2 with a mitochondrial import signal sequence (from human pyruvate dehydrogenase 1α) [306], which should then lead to the concentration of CD2 in mitochondria. Rat CD2 can be detected using a high-affinity and well-characterized antibody, OX34, that recognizes an epitope near the N-terminus (encoded by exons 2 and 3) [307,308]; as this epitope of CD2 protein is known to fold rapidly [309], this increases the chances that I might detect the N-terminal antigen whilst the protein is still being made. [In **Ref** [309], the N-terminal domain of CD2 containing this epitope – CD2.D1 [304] – is incorrectly named as the C-terminal domain of CD2 protein.]

As there remains no robust way of detecting a single protein, let alone a nascent single protein (especially when the target is embedded in dense chromatin that would be expected to reduce probe access), I used a multi-copy system [310] based on the SV40 *ori* (origin of replication) [311] and Cos7 cells (kidney fibroblast cells from the African green monkey, *Cercopithecus aethiops*) which express the SV40 large T antigen [312]. When a vector encoding the *ori* is transfected into Cos7 cells, the T antigen drives replication from the *ori* to generate thousands of copies of the vector; these copies are incorporated into 'minichromosomes' [313,314] which are covered with nucleosomes and transcribed by the cellular machinery. As they all encode the same promoters, they tend to be co-transcribed in the same transcription factories [315]. If the minichromosomes encode CD2 protein, and if some translation occurs co-transcriptionally, one might expect a number of copies of nascent CD2 RNA to lie next to a number of copies of nascent CD2 protein.

I also incorporated a gene switch into the expression vector, so that I could turn on CD2 transcription rapidly. To this end, I used the commercially-available 'TetOn

Advanced' system; the CD2 gene was driven by a doxycycline-inducible promoter – a modified 'Tet' promoter (P^{tight}) – that is normally repressed in cells expressing a modified Tet-repressor – rtTA-Advanced (under the control of a constitutive cytomegalo-virus, CMV, promoter). When doxycycline is added, the 'TetOn Advanced' repressor switches to become an activator and fires expression from the 'Tet' promoter [316,317].

In order to facilitate detection of cells expressing CD2 protein, I fused a sequence encoding EGFP [318-321] with the C-terminus of the CD2 sequence. At its 3′-end, the EGFP coding sequence was followed by an SV40 polyA (polyadenylation signal) [322,323], which reduces transcriptional read-through into other coding sequences in the vector, and helps post-transcriptional processing of the synthesized RNA. I also inserted a ribosome pause sequence (rps) [324,325] just after the CD2 sequence and before EGFP, which is expected to slow down a translating ribosome after it has translated CD2. Then, as the ribosome translates the EGFP-containing portion of the nascent peptide, the CD2-containing portion should get more time to fold and display the OX34-recognized epitope. I also included parts of intron 2 of the *Cd2* gene in my vector, and the use I make of this is discussed below. All the above-mentioned components constitute my basic expression vector – Vector 1 (**Fig. 3.10**).

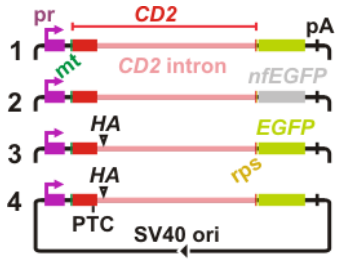


Fig 3.10. CD2-EGFP expression constructs. Each construct encodes an SV40 *ori* (to permit replication) and a modified 'Tet' promoter inducible with doxycycline. In vector 1, the 'Tet' promoter drives expression of a protein containing an N-terminal mitochondrial signal sequence (mt), the rat CD2 epitope 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 increase the time available for CD2 to fold as the protein is being made), and a C-terminal EGFP. Vector 2 encodes an EGFP sequence with the segment encoding the fluorochrome deleted (nfEGFP, to free the green channel for other use). Vector 3 encodes a haemagglutinin (HA) tag inserted within the intron upstream of the first stop codon. Vector 4 also bears a PTC, 67 nucleotides upstream of the intron-exon junction (a position that should induce NMD).

I wished to determine whether some CD2 protein might be found close to the nascent transcript that encoded it; as nascent RNA is found only in the nucleus, finding CD2 protein at the same site would provide strong circumstantial evidence that the protein was made at the transcription site. To label nascent CD2 transcripts, I intended to perform RNA FISH (fluorescent *in situ* hybridization) [326,327] with probes targeting a part of intron 2 of the *Cd2* gene included in my vector. I reduced the length of CD2 intron-2 to ~2.5 kb (to facilitate cloning) by removing a central region whilst retaining the splice donor and acceptor sites [328]. Intronic RNA is degraded rapidly after synthesis and not found in the cytoplasm [329-335]; therefore, it is often used as a marker for nascent RNA and so the site of transcription [330,331,335,336].

Vector 1 encodes EGFP, which facilitates detection of transfected cells (which auto-fluoresce green). As I intended to use the blue channel to detect nuclei (after counterstaining with DAPI), the red channel to immuno-detect CD2 protein, and the far-red channel to detect mitochondria (after staining with MitoTracker Deep Red), I deleted the EGFP-chromophore coding sequence [337-339] to free the green channel for detection of intronic CD2 RNA. The resulting Vector 2 encodes a non-fluorescent EGFP (nfEGFP; **Fig. 3.10**).

In Vectors 3 and 4 (**Fig. 3.10**), I inserted a short coding sequence for the haemagglutinin tag (HA tag) [340] in the CD2 intron (before the first termination codon

within this intron); this enables testing whether ribosomes could translate into this intron to express an HA-tagged protein, which can then be detected by indirect immuno-labelling. Vector 4 additionally has a premature termination codon (PTC) 67 nucleotides upstream of the exon-intron junction, which is expected to trigger degradation of CD2-EGFP RNA transcribed from this vector by the NMD (nonsense-mediated mRNA decay) machinery [244,245]. Since NMD depends on active translation by ribosomes to detect PTC-bearing RNA [232,341], any effect of the PTC on the stability of the intronic RNA would point to nuclear translation playing a role in NMD.

3.4.1 The mitochondria-targeted non-nuclear CD2 protein is also found in nuclei

I co-transfected Cos7 cells with Vector 1 and the plasmid encoding the rtTA-Advanced repressor, grew the cells for 24 h, and induced CD2-EGFP expression from Vector 1 with doxycycline for 45 min; after staining cells with MitoTracker, cells were fixed, and CD2 indirectly immuno-labelled. In transfected cells (**Fig. 3.11**, cell on right), most CD2 protein (red) is found in mitochondria ('MitoTracker Deep Red' fluoresces in the far red and is pseudo-coloured green) to give yellow in the 'merge' (**Fig. 3.11B-D**); however, some punctate CD2 signal is found in nuclei (**Fig. 3.11A**,**C**,**D**). The lymphocyte antigen, CD2, is not expressed in Cos7 cells, and – as the OX34 antibody is highly specific – there is little background in untransfected cells (**Fig. 3.11**, cell on left). The detection of some CD2 – a non-nuclear protein targeted to mitochondria – in nuclei raises the question as to how it got there.

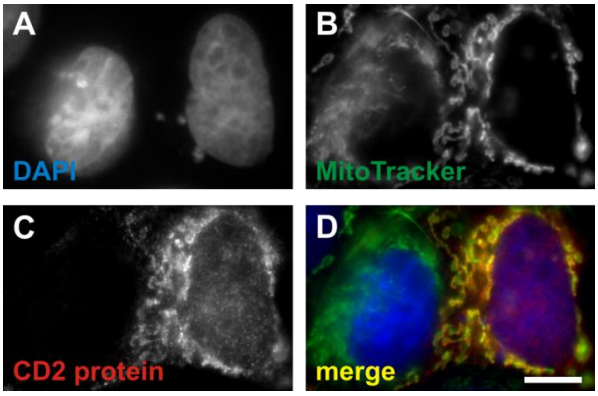


Fig. 3.11. A non-nuclear protein bearing a mitochondrial-import sequence can be found in nuclei. Cos7 cells (which express the SV40 T antigen) were co-transfected with one construct encoding the 'TetOn Advanced' repressor, plus a second encoding the SV40 *ori* and the 'Tet' promoter driving an artificial CD2-EGFP gene (i.e., vector 1). By 24 h, the *ori*-containing vector replicates to give ~8,000 minichromosomes per cell; the 'Tet' promoter is expressed only at basal levels and essentially no CD2 epitope or EGFP fluorescence can be detected. 10 µM doxycycline was now added for 45 min to induce CD2-EGFP expression, and 'MitoTracker' added during the last 15 min to stain mitochondria. After fixation, CD2 was indirectly immuno-labelled, DNA counterstained with DAPI, and images of cells expressing the CD2 epitope collected using a wide-field microscope.

A-D. Four views of two cells are shown. Only the cell on the right has been transfected and expresses CD2 protein (which is mainly found in mitochondria); however, some CD2 is seen in the nucleus. Bar: 10 µm.

3.4.2 CD2 protein foci in nuclei are close to their encoding nascent RNA

If some translation occurs coupled to transcription in nuclei, then I should observe nascent CD2 protein close to the nascent RNA that encoded it (**Fig. 3.12A**). Therefore, I co-transfected Cos7 cells with Vector 2 and the plasmid encoding the 'TetOn Advanced' repressor, induced expression of CD2-nfEGFP from Vector 2 with doxycycline, and performed an immuno-FISH experiment to detect CD2 protein and intronic RNA. CD2 was present in many nuclear foci (**Fig. 3.12C**), as was nascent (intronic) CD2 RNA (**Fig. 3.12D**); the 'merge' (**Fig. 3.12E**) indicates that some CD2 protein (red) lies close to CD2

RNA (green) to give yellow. As found previously, it is difficult to infer whether the degree of colocalization is significantly more than expected by chance; therefore, I applied the 'super-resolution' method of analysis used previously (**Fig. 3.8B**); nuclear CD2 protein lay significantly closer to nascent CD2 RNA than expected by chance (**Fig. 3.12F**,**G**). The PDF for the experimental foci shows that ~60% of the red and green foci are within 0.54 μm (this limit is selected because the PDF for the observed NN distances is higher than the PDF for the NN distances of the random set up to 0.54 μm). Once again, the presence of many red and green foci (and few yellow foci) in the merge in **Figure 3.12E** indicates that there is only incomplete overlap between the red and green foci; this apparent discrepancy has been discussed previously in **Section 3.3.2**, and will be discussed further in **Section 4.3**.

As expected, doxycycline treatment (which allows the promoter to fire) increased nuclear CD2 protein and intronic RNA levels (**Fig. 3.12H**). Pre-treatment with anisomycin (an inhibitor of translational elongation) plus pactamycin (an inhibitor of both initiation and elongation) [342-344] decreased nuclear levels of CD2 protein more than intronic RNA (**Fig. 3.12H**); this is consistent with the inhibitors having a direct effect on translation (whether in the cytoplasm or nucleus), and only an indirect effect on transcription.

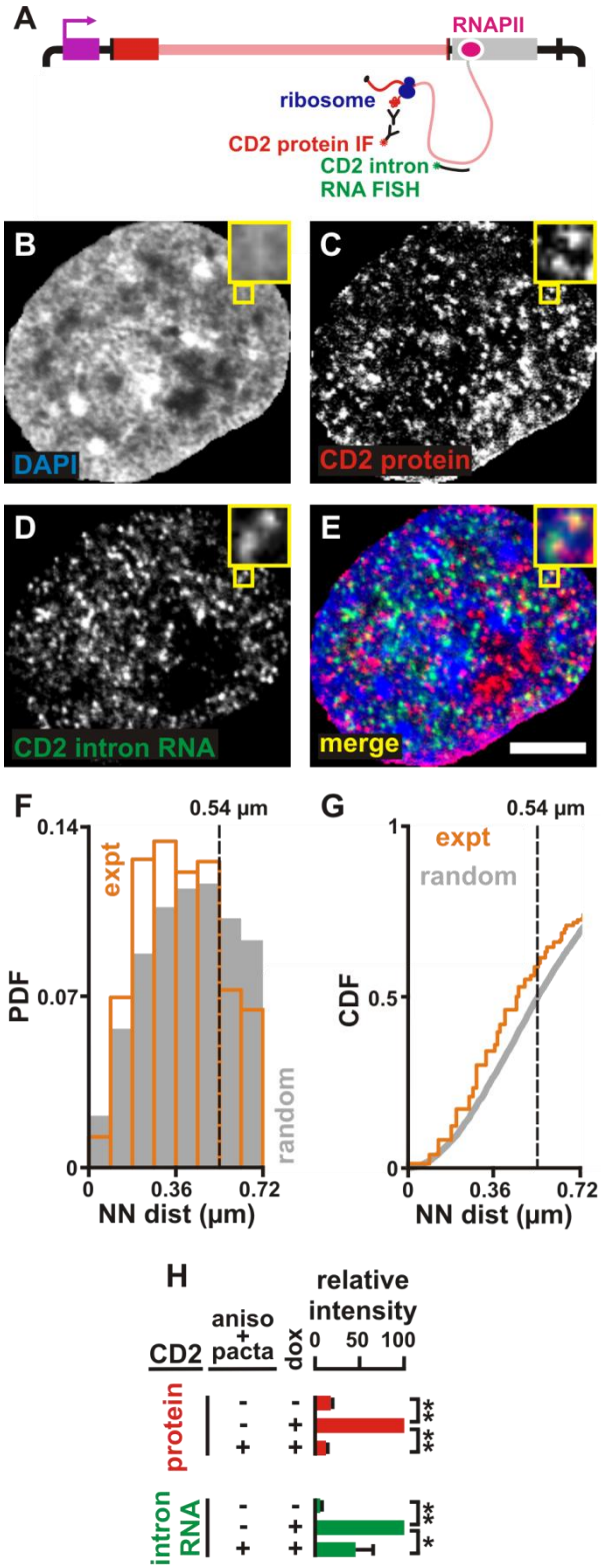


Fig. 3.12. A non-nuclear protein bearing a mitochondrial import sequence can be found close to the intronic (nascent) RNA that encoded it. Cos7 cells were co-transfected with one construct encoding the 'TetOn Advanced' repressor, plus a second (i.e., vector 2) that encodes the SV40 *ori* and the 'Tet' promoter driving an artificial *CD2* gene containing an intron. By 24 h, the *ori*-containing vector replicates to give

~8,000 minichromosomes per cell; the 'Tet' promoter is silent and essentially no CD2 can be detected. 10 µM doxycycline was now added for 45 min to induce CD2 expression; after fixation, CD2 is indirectly immuno-labelled, RNA FISH performed using probes targeting the intron, DNA counterstained with DAPI, and equatorial sections of CD2-expressing nuclei collected using a confocal microscope (B-E) or a wide-field microscope (F-H).

A. Strategy. The cartoon illustrates the transcription unit in vector 2 being transcribed as a ribosome translates the nascent transcript. If such coupled transcription and translation occurs, the nascent CD2 epitope (detected by immuno-fluorescence, IF; first and second antibodies shown) should lie close to the nascent RNA that encodes it (detected by RNA FISH using an intronic probe).

B-E. Four views of one typical section (only nucleus shown); the CD2 protein (red) and intronic RNA (green) are both found in discrete foci in the nucleus (blue), some of which overlap to give yellow in the merge (inset: a yellow focus). Bar: 5 µm.

F,G. Super-resolution localization shows red and green foci lie close together. Foci with a Gaussian intensity profile were selected, peaks localized with 15-nm precision, and nearest-neighbour (NN) distances calculated (from each red peak to the nearest green peak, and *vice-versa*) with 10-50 nm precision. The probability density function (PDF) and cumulative distribution function (CDF) of NN distances is compared to the simulated PDF and CDF (line includes 99%-confidence limit) of the same nuclear density of green and red foci distributed randomly (random); ~60% of green and red foci lie closer than 0.54 µm; the experimental CDF curve (expt) lies to the left of the random one, indicative of significant proximity ($P < 0.0001$, 1-sample Kolmogorov-Smirnov test, $n_{expt} = 6224$, $n_{random} = 11154$, cells = 14).

H. After subtracting background, intensities (± SD) of CD2 protein and CD2 intron RNA are expressed relative to the value found in the doxycycline-induced nucleus (**: $P < 0.0001$, *: $P < 0.005$, Student's two-tailed t test, $n = 15$ cells). Doxycycline induces expression of both CD2 RNA and protein in the nucleus; pretreatment with 100 µg/ml anisomycin + pactamycin 45 min before adding doxycycline reduces CD2 protein signal in the nucleus, without significantly affecting CD2 intron RNA.

3.4.3 Nascent CD2 RNA harbouring PTCs is a substrate for nuclear NMD

Any RNA encoding a PTC is generally degraded by NMD. As a translating ribosome is the only mechanism known that can detect a PTC, and if such a ribosome proof-reads the nascent transcript for PTCs [232,341] that might arise through mis-

splicing [345,346], then introducing a PTC into the vector might have an effect on the nuclear signal. Therefore, I co-transfected Cos7 cells with Vector 3 or 4 (without and with a PTC, respectively) and with the plasmid encoding the 'TetOn Advanced' repressor, induced expression of CD2-EGFP, and monitored levels of intronic CD2 RNA using FISH. Cells transfected with the (control) PTC^- vector contained considerable amounts of intronic CD2 RNA in discrete nucleoplasmic foci (**Fig. 3.13A**,**B**; the cell at the top is the only transfected cell in the field, and only it contains CD2 RNA). Cells transfected with the PTC^+ vector yielded significantly less nuclear signal (**Fig. 3.13Ci**); this is consistent with nascent (intronic) PTC^+ RNA being destroyed by NMD. Significantly, pretreatment with cycloheximide increased signal even above that seen with the control (**Fig. 3.13Ci**); this is consistent with the inhibitor preventing detection of PTC^+ transcripts, so that they are no longer destroyed by NMD.

It is widely assumed that PTCs might arise due to faulty splicing [232,341]. Therefore, I examined how spliceostatin A (SSA) – a splicing inhibitor [347] – affected levels of intronic PTC^- RNA; SSA reduced levels (**Fig. 3.13Cii**). Presumably, the now-unspliced transcripts contain many (intronic) PTCs that are degraded by NMD. Significantly, pretreatment with anisomycin reversed this reduction (**Fig. 3.13Cii**). This is consistent with inhibition of a translating ribosome preventing detection of the PTCs [232,341], so that the intron-containing transcripts are no longer degraded by NMD. These experiments point to a translating ribosome acting on intronic RNA during NMD. As intronic RNA is found only in the nucleus close to the active site of transcription, this means that some translation must occur close to the transcription site.

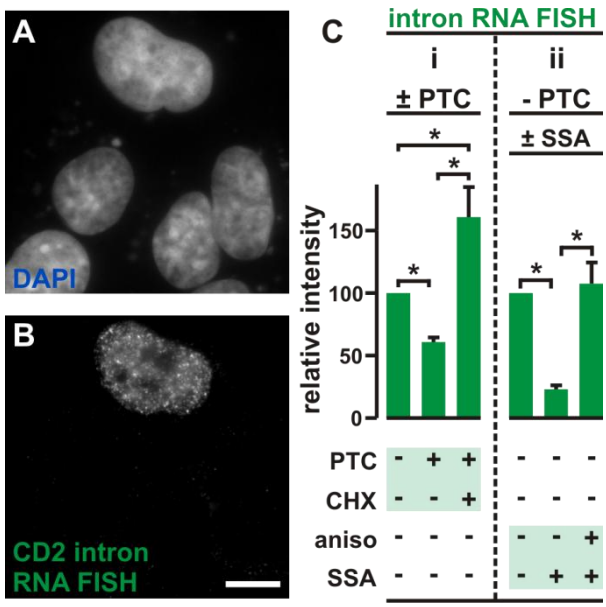


Fig. 3.13. The effects of a premature termination codon (PTC) and a splicing inhibitor – spliceostatin A (SSA) – on the stability of nascent CD2 RNA. Cos-7 cells were co-transfected with one construct encoding the 'TetOn Advanced' repressor plus the vector indicated, 10 μM doxycycline added for the last 45 min, and cells fixed 26 h after transfection. In (C), inhibitors were added at the times indicated. Next, intronic CD2 RNA was detected by RNA FISH, DNA counterstained with DAPI, and images of cells expressing CD2 RNA collected using a wide-field microscope.

A-B. Two views of a group of cells after a co-transfection using vector 3; only the cell at the top has been transfected and so expresses intronic CD2 RNA (bar: 10 μm).

C. Effects of a PTC and spliceostatin A.

(i). Cells were transfected with vector 3 (PTC^-) or vector 4 (PTC^+), and cells treated ± 100 μg/ml cycloheximide 2 h prior to fixation; after subtracting background, intensities (± SD) seen in nuclei of cells transfected with vector 4 are expressed relative to the value found in cells transfected with vector 3. The presence of a PTC significantly reduces levels of intronic RNA, and pre-treatment with cycloheximide reverses the effect (*: $P < 0.0004$, Student's two-tailed t test, $n = 20$ cells). This is consistent with a cycloheximide-sensitive ribosome acting on intronic RNA as the detector that triggers NMD in the nucleus (as intronic RNA is not found in the cytoplasm).

(ii). Cells were transfected with vector 2 (PTC^-), and cells treated ± 100 μg/ml anisomycin (2 h before fixation) and ± 100 ng/ml SSA (45 min before fixation); after subtracting background, intensities (± SD) seen in nuclei treated with SSA ± anisomycin are expressed relative to the value found in untreated cells. SSA pre-treatment significantly reduces levels of intronic RNA (*: $P < 0.0001$, Student's two-tailed t test, n

= 20 cells), and pre-treatment with anisomycin reverses this effect. This is consistent with SSA increasing the numbers of intron-containing transcripts that contain many stop codons (which are then immediately degraded by NMD), and with an anisomycin-sensitive ribosome acting on intronic RNA as the detector that triggers NMD in the nucleus (as intronic RNA is not found in the cytoplasm).

3.4.4 Some intronic RNA is translated by ribosomes

The presence of introns in eukaryotic genes [348] underpins the idea that translation cannot occur in nuclei [349]: if nuclear ribosomes translated through an intron before it was removed by splicing, the presence of stop codons in the intron would inevitably lead to the production of truncated peptides – and many of these would surely be toxic to the cell. Out of curiosity, I investigated whether a ribosome – in either the cytoplasm or nucleus – could indeed translate an intron in the cell. Therefore, I co-transfected Cos7 cells with Vector 3 (which contains a short sequence encoding the HA tag [340] in the CD2 intron, which is in-frame with CD2 exon-2, and which lies before the first intronic stop codon) plus the plasmid encoding the repressor. In order to drive over-production of intronic RNA, transfected cells were grown for 50 h (instead of the usual 24 h) to allow even further minichromosomal replication; moreover, cells were grown in the continual presence of doxycycline (instead of just the last 45 min before fixation). Finally, the HA tag was detected by indirect immuno-labelling. **Figure 14A** illustrates two images of one transfected cell; the image on the right has been 'contrast-stretched' six-fold relative to the one on the left. Although signal is faint, some can nevertheless be detected in mitochondria (as the HA tag is in-frame with CD2 exon-2, which in turn is in-frame with the N-terminal mitochondrial signal sequence). The presence of the splicing inhibitor, SSA [347], increases this cytoplasmic signal (**Fig. 3.14B**,**C**). Clearly, intronic sequences can be translated, as has been shown previously in

prokaryotes [350-352] and eukaryotes [347]. Whilst this experiment involves prolonged incubations with drugs, it nevertheless indicates that vector and inhibitor are behaving in the expected manner.

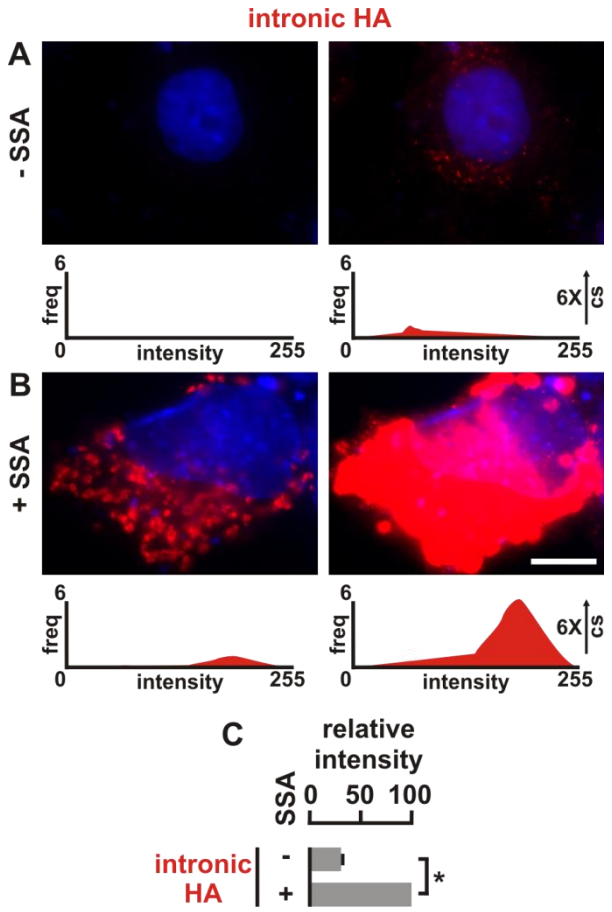


Fig. 3.14. Some intronic RNA is translated by ribosomes. Cos7 cells were co-transfected with one construct encoding the 'TetOn Advanced' repressor, plus a second (i.e., vector 3) that encodes the SV40 *ori* and the 'Tet' promoter driving an artificial CD2-EGFP gene and an intronic haemagglutinin (HA) tag; immediately after transfection, 10 µM doxycycline ± 100 ng/ml SSA were added. By 50 h, the *ori*-containing vector replicates to give ~20,000 minichromosomes per cell, which are actively transcribed. After fixation, HA was indirectly immuno-labelled (red), DNA counterstained with DAPI (blue), and images of cells expressing the HA epitope collected using a wide-field microscope.

A. Even in the absence of SSA, some translation of intronic HA sequence occurs. Cells express HA epitope which is concentrated in the mitochondria (a 6-fold 'contrast stretch' from **Ai** to **Aii** helps to see this faint signal).

B. Adding SSA induces mis-splicing of CD2-EGFP gene, thus increasing the chances of intronic HA sequence being subjected to translation. Cells express HA epitope (HA signal can be easily seen in **Bi**, without a 6-fold 'contrast stretch' – in contrast to **A**) which is concentrated in mitochondria. Bar: 10 µm.

C. After subtracting background, intensities (± SD) seen with HA epitope are expressed relative to the value found in the cytoplasm of SSA treated cells (*: $P < 0.0001$, Student's two-tailed t test, n = 17-19 cells). Some intronic HA is expressed, and SSA increases levels; the HA signal is significantly more than the background and mostly accumulates in the mitochondria.

3.5 References

214. **Hein JE, Fokin VV.** Copper-catalyzed azide-alkyne cycloaddition (CuAAC) and beyond: new reactivity of copper(I) acetylides. *Chem Soc Rev*. 2010;**39**(4):1302-1315.

215. **Sletten EM, Bertozzi CR.** Bioorthogonal chemistry: fishing for selectivity in a sea of functionality. *Angew Chem Int Ed Engl*. 2009;**48**(38):6974-6998.

216. **Johnson JA, Lu YY, Van Deventer JA, Tirrell DA.** Residue-specific incorporation of non-canonical amino acids into proteins: recent developments and applications. *Curr Opin Chem Biol.* 2010;**14**(6):774-780.

217. **Dieterich DC, Link AJ, Graumann J, Tirrell DA, Schuman EM.** Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc Natl Acad Sci U S A.* 2006;**103**(25):9482-9487.

218. **Beatty KE, Liu JC, Xie F, Dieterich DC, Schuman EM, Wang Q, Tirrell DA.** Fluorescence visualization of newly synthesized proteins in mammalian cells. *Angew Chem Int Ed Engl*. 2006;**45**(44):7364-7367.

219. **Beatty KE, Tirrell DA.** Two-color labeling of temporally defined protein populations in mammalian cells. *Bioorg Med Chem Lett*. 2008;**18**(22):5995-5999.

220. **Dieterich DC, Hodas JJ, Gouzer G, Shadrin IY, Ngo JT, Triller A, Tirrell DA, Schuman EM.** *In situ* visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nat Neurosci*. 2010;**13**(7):897-905.

221. **Chen PJ, Huang YS.** CPEB2-eEF2 interaction impedes HIF-1α RNA translation. *EMBO J*. 2011;**31**(4):959-971.

222. **Link AJ, Vink MK, Tirrell DA.** Presentation and detection of azide functionality in bacterial cell surface proteins. *J Am Chem Soc*. 2004;**126**(34):10598-10602.

223. **Kramer G, Sprenger RR, Back J, Dekker HL, Nessen MA, van Maarseveen JH, de Koning LJ, Hellingwerf KJ, de Jong L, de Koster CG.** Identification and quantitation of newly synthesized proteins in Escherichia coli by enrichment of azidohomoalanine-labeled peptides with diagonal chromatography. *Mol Cell Proteomics*. 2009;**8**(7):1599-1611.

224. **Banerjee PS, Ostapchuk P, Hearing P, Carrico IS.** Unnatural amino acid incorporation onto adenoviral (Ad) coat proteins facilitates chemoselective modification and retargeting of Ad type 5 vectors. *J Virol*. 2011;**85**(15):7546-7554.

225. **Qin Q, Carroll K, Hastings C, Miller CL.** Mammalian orthoreovirus escape from host translational shutoff correlates with stress granule disruption and is independent of eIF2alpha phosphorylation and PKR. *J Virol*. 2011;**85**(17):8798-8810.

226. **Teves SS, Henikoff S.** Heat shock reduces stalled RNA polymerase II and nucleosome turnover genome-wide. *Genes Dev*. 2011;**25**(22):2387-2397.

227. **Hinz FI, Dieterich DC, Tirrell DA, Schuman EM.** Non-canonical amino acid labeling *in vivo* to visualize and affinity purify newly synthesized proteins in larval zebrafish. *ACS Chem Neurosci*. 2012;**3**(1):40-49.

228. **Nagarajan S, Taskent-Sezgin H, Parul D, Carrico I, Raleigh DP, Dyer RB.** Differential ordering of the protein backbone and side chains during protein folding revealed by site-specific recombinant infrared probes. *J Am Chem Soc*. 2011;**133**(50):20335-20340.

229. **Hodas JJ, Nehring A, Höche N, Sweredoski MJ, Pielot R, Hess S, Tirrell DA, Dieterich DC, Schuman EM.** Dopaminergic modulation of the hippocampal neuropil proteome identified by bio-orthogonal non-canonical amino-acid tagging (BONCAT). *Proteomics*. 2012; **12**(15-16): 2464-2476.

230. **Grollman AP.** Inhibitors of protein biosynthesis. II. Mode of action of anisomycin. *J Biol Chem*. 1967;**242**(13):3226-3233.

231. **Hansen JL, Moore PB, Steitz TA.** Structures of five antibiotics bound at the peptidyl transferase center of the large ribosomal subunit. *J Mol Biol*. 2003;**330**(5):1061-1075.

232. **Carter MS, Doscow J, Morris P, Li S, Nhim RP, Sandstedt S, Wilkinson MF.** 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.* 1995;**270**(48):28995-29003.

233. **Qian SB, Princiotta MF, Bennink JR, Yewdell JW.** Characterization of rapidly degraded polypeptides in mammalian cells reveals a novel layer of nascent protein quality control. *J Biol Chem*. 2006;**281**(1):392-400.

234. **Reits E, Griekspoor A, Neijssen J, Groothuis T, Jalink K, van Veelen P, Janssen H, Calafat J, Drijfhout JW, Neefjes J.** Peptide diffusion, protection, and degradation in nuclear and cytoplasmic compartments before antigen presentation by MHC class I. *Immunity*. 2003;**18**(1):97-108.

235. **Iborra FJ, Escargueil AE, Kwek KY, Akoulitchev A, Cook PR.** Molecular cross-talk between the transcription, translation, and nonsense-mediated decay machineries. *J Cell Sci*. 2004;**117**(6):899-906.

236. **Dolan BP, Bennink JR, Yewdell JW.** Translating DRiPs: progress in understanding viral and cellular sources of MHC class I peptide ligands. *Cell Mol Life Sci*. 2011;**68**(9):1481-1489.

237. **Reits EA, Benham AM, Plougastel B, Neefjes J, Trowsdale J.** Dynamics of proteasome distribution in living cells. *EMBO J*. 1997;**16**(20):6087-6094.

238. **Baumeister W, Walz J, Zühl F, Seemüller E.** The proteasome: paradigm of a self-compartmentalizing protease. *Cell*. 1998;**92**(3):367-380.

239. **Geng F, Wenzel S, Tansey WP.** Ubiquitin and proteasomes in transcription. *Annu Rev Biochem*. 2012;**81**:177-201.

240. **Yao T, Ndoja A.** Regulation of gene expression by the ubiquitin-proteasome system. *Semin Cell Dev Biol*. 2012;**23**(5):523-529.

241. **Gillette TG, Gonzalez F, Delahodde A, Johnston SA, Kodadek T.** Physical and functional association of RNA polymerase II and the proteasome. *Proc Natl Acad Sci U S A*. 2004;**101**(16):5904-5909.

242. **Rockel TD, Stuhlmann D, von Mikecz A.** Proteasomes degrade proteins in focal subdomains of the human cell nucleus. *J Cell Sci*. 2005;**118**(22):5231-5242.

243. **Palombella VJ, Rando OJ, Goldberg AL, Maniatis T.** The ubiquitin-proteasome pathway is required for processing the NF-kappa B1 precursor protein and the activation of NF-kappa B. *Cell*. 1994;**78**(5):773-785.

244. **Chang Y-F, Imam JS, Wilkinson MF.** The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem.* 2007;**76**:51-74.

245. **Brogna S, Wen J.** Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat Struct Mol Biol.* 2009;**16**(2):107-113.

246. **Kuroha K, Tatematsu T, Inada T.** Upf1 stimulates degradation of the product derived from aberrant messenger RNA containing a specific nonsense mutation by the proteasome. *EMBO Rep*. 2009;**10**(11):1265-1271.

247. **Apcher S, Daskalogianni C, Lejeune F, Manoury B, Imhoos G, Heslop L, Fåhraeus R.** Major source of antigenic peptides for the MHC class I pathway is produced during the pioneer round of mRNA translation. *Proc Natl Acad Sci U S A*. 2011;**108**(28):11572-11577.

248. **Nielsen PJ, McConkey EH.** Evidence for control of protein synthesis in HeLa cells *via* the elongation rate. *J Cell Physiol*. 1980;**104**(3):269-281.

249. **Scherer S.** *A Short Guide to the Human Genome*. 2008; Cold Spring Harbor Laboratory Press, New York.

250. **Brocchieri L, Karlin S.** Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Res*. 2005;**33**(10):3390-3400.

251. **Tan T, Frenkel D, Gupta V, Deem MW.** Length, protein-protein interactions, and complexity. *Physica A*. 2005;**350**(1):52–62.

252. **Ainslie SJ, Fox DG, Perry TC, Ketchen DJ, Barry MC.** Predicting amino acid adequacy of diets fed to Holstein steers. *J Anim Sci*. 1993;**71**(5):1312-1319.

253. **Wettstein FO, Staehelin T, Noll H.** Ribosomal aggregate engaged in protein synthesis: characterization of the ergosome. *Nature*. 1963;**197**(4866):430-435.

254. **Warner JR, Rich A, Hall CE.** Electron microscope studies of ribosomal clusters synthesizing hemoglobin. *Science*. 1962;**138**(3548):1399-1403.

255. **Yarmolinsky MB, Haba GL.** Inhibition by puromycin of amino acid incorporation into protein. *Proc Natl Acad Sci U S A*. 1959;**45**(12):1721-1729.

256. **Nathans D.** Puromycin inhibition of protein synthesis: incorporation of puromycin into peptide chains. *Proc Natl Acad Sci U S A*. 1964;**51**(4):585-592.

257. **Nissen P, Hansen J, Ban N, Moore PB, Steitz TA.** The structural basis of ribosome activity in peptide bond synthesis. *Science*. 2000;**289**(5481):920-930.

258. **Graifer DM, Fedorova OS, Karpova GG.** Interaction of puromycin with acceptor site of human placenta 80 S ribosomes. *FEBS Lett*. 1990;**277**(1):4-6.

259. **Meskauskas A, Dinman JD.** Ribosomal protein L3 functions as a 'rocker switch' to aid in coordinating of large subunit-associated functions in eukaryotes and Archaea. *Nucleic Acids Res*. 2008;**36**(19):6175-6186.

260. **Williamson AR, Schweet R.** Role of the genetic message in polyribosome function. *J Mol Biol*. 1965;**11**(2):358-372.

261. **Kaempfer R, Meselson M.** Studies of ribosomal subunit exchange. *Cold Spring Harb Symp Quant Biol*. 1969;**34**:209-220.

262. **Obrig TG, Culp WJ, McKeehan WL, Hardesty B.** The mechanism by which cycloheximide and related glutarimide antibiotics inhibit peptide synthesis on reticulocyte ribosomes. *J Biol Chem*. 1971;**246**(1):174-181.

263. **Schneider-Poetsch T, Ju J, Eyler DE, Dang Y, Bhat S, Merrick WC, Green R, Shen B, Liu JO.** Inhibition of eukaryotic translation elongation by cycloheximide and lactimidomycin. *Nat Chem Biol*. 2010;**6**(3):209-217.

264. **Schmidt EK, Clavarino G, Ceppi M, Pierre P.** SUnSET, a nonradioactive method to monitor protein synthesis. *Nat Methods*. 2009;**6**(4):275-277.

265. **Siuta-Mangano P, Lane MD.** 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*. 1981;**256**(5):2094-2097.

266. **David A, Dolan BP, Hickman HD, Knowlton JJ, Clavarino G, Pierre P, Bennink JR, Yewdell JW.** Nuclear translation visualized by ribosome-bound nascent chain puromycylation. *J Cell Biol*. 2012;**197**(1):45-57.

267. **Palade GE.** A small particulate component of the cytoplasm. *J Biophys Biochem Cytol*. 1955;**1**(1):59-68.

268. **Adelman MR, Sabatini DD, Blobel G.** Ribosome-membrane interaction. Nondestructive disassembly of rat liver rough microsomes into ribosomal and membranous components. *J Cell Biol*. 1973;**56**(1):206-229.

269. **Redman CM, Sabatini DD.** Vectorial discharge of peptides released by puromycin from attached ribosomes. *Proc Natl Acad Sci U S A*. 1966;**56**(2):608-615.

270. **Redman CM.** Studies on the transfer of incomplete polypeptide chains across rat liver microsomal membranes in vitro. *J Biol Chem*. 1967;242(4):761-768.

271. **Burka ER, Marks PA.** Protein synthesis in erythroid cells. II. Polyribosome function in intact reticulocytes. *J Mol Biol*. 1964;**9**(2):439-451.

272. **Blobel G, Sabatini D.** Dissociation of mammalian polyribosomes into subunits by puromycin. *Proc Natl Acad Sci U S A*. 1971;**68**(2):390-394.

273. **Blobel G.** Release, identification, and isolation of messenger RNA from mammalian ribosomes. *Proc Natl Acad Sci U S A*. 1971;**68**(4):832-835.

274. **Caro LG, Palade GE.** Protein synthesis, storage, and discharge in the pancreatic exocrine cell. An autoradiographic study. *J Cell Biol*. 1964;**20**(3):473-495.

275. **Palade G.** Intracellular aspects of the process of protein synthesis. *Science*. 1975;**189**(4200):347-358.

276. **Munro AJ, Jackson RJ, Korner A.** Studies on the nature of polysomes. *Biochem J*. 1964;**92**(2):289-299.

277. **Lamfrom H, Knopf PM.** Initiation of haemoglobin synthesis in cell-free systems. *J Mol Biol*. 1964;**9**:558-575.

278. **Lamfrom H, Knopf PM.** Properties of reticulocyte 80 s ribosomes. *J Mol Biol*. 1965;**11**(3):589-599.

279. **Luzzatto L, Banks J, Marks PA.** Protein synthesis in erythroid cells. III. Monoribosome and polyribosome function in the cell-free system. *Biochim Biophys Acta*. 1965;**108**(3):434-446.

280. **Iborra FJ, Jackson DA, Cook PR.** Coupled transcription and translation within nuclei of mammalian cells. *Science*. 2001;**293**(5532):1139-1142.

281. **Allfrey VG, Mirsky AE, Osawa S.** Protein synthesis in isolated cell nuclei. *J Gen Physiol*. 1957;**40**(3):451-490.

282. **Stevens A, Maupin MK.** 5,6-Dichloro-1-beta-D-ribofuranosylbenzimidazole inhibits a HeLa protein kinase that phosphorylates an RNA polymerase II-derived peptide. *Biochem Biophys Res Commun*. 1989;**159**(2):508-515.

283. **Roeder RG.** The eukaryotic transcriptional machinery: complexities and mechanisms unforeseen. *Nat Med*. 2003;**9**(10):1239-1244.

284. **Corden JL, Cadena DL, Ahearn JM Jr, Dahmus ME.** A unique structure at the carboxyl terminus of the largest subunit of eukaryotic RNA polymerase II. *Proc Natl Acad Sci U S A*. 1985;**82**(23):7934-7938.

285. **Cramer P, Bushnell DA, Kornberg RD.** Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science*. 2001;**292**(5523):1863-1876.

286. **Cadena DL, Dahmus ME.** Messenger RNA synthesis in mammalian cells is catalyzed by the phosphorylated form of RNA polymerase II. *J Biol Chem*. 1987;**262**(26):12468-12474.

287. **Komarnitsky P, Cho EJ, Buratowski S.** Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev*. 2000;**14**(19):2452-2460.

288. **Egloff S, Dienstbier M, Murphy S.** Updating the RNA polymerase CTD code: adding gene-specific layers. *Trends Genet*. 2012;**28**(7):333-341.

289. **Bregman DB, Du L, van der Zee S, Warren SL.** Transcription-dependent redistribution of the large subunit of RNA polymerase II to discrete nuclear domains. *J Cell Biol*. 1995;**129**(2):287-298.

290. **Rahl PB, Lin CY, Seila AC, Flynn RA, McCuine S, Burge CB, Sharp PA, Young RA.** c-Myc regulates transcriptional pause release. *Cell*. 2010;**141**(3):432-445.

291. **Larkin JD, Papantonis A, Cook PR.** Promoter type influences transcriptional topography by targeting genes to distinct nucleoplasmic sites. *J Cell Sci*. 2013;**126**(9):2052-2059.

292. **Larkin JD, Cook PR.** Maximum precision closed-form solution for localizing diffraction-limited spots in noisy images. *Opt Express*. 2012;**20**(16):18478-18493.

293. **Larkin JD, Cook PR, Papantonis A.** Dynamic reconfiguration of long human genes during one transcription cycle. *Mol Cell Biol*. 2012;**32**(14): 2738-2747.

294. **Thompson RE, Larson DR, Webb WW.** Precise nanometer localization analysis for individual fluorescent probes. *Biophys J*. 2002;**82**(5), 2775-2783.

295. **Yildiz A, Forkey JN, McKinney SA, Ha T, Goldman YE, Selvin PR.** Myosin V walks hand-over-hand: single fluorophore imaging with 1.5-nm localization. *Science*. 2003; **300**(5628), 2061-2065.

296. **Barbini P, Cevenini G, Massai MR.** Nearest-neighbor analysis of spatial point patterns: application to biomedical image interpretation. *Comput Biomed Res*. 1996;**29**(6):482-493.

297. **Hodges JL.** The significance probability of the smirnov two-sample test. *Ark Mat*. 1958;**3**(5):469-486.

298. **McManus KJ, Stephens DA, Adams NM, Islam SA, Freemont PS, Hendzel MJ.** The transcriptional regulator CBP has defined spatial associations within interphase nuclei. *PLoS Comput Biol*. 2006;**2**(10):e139.

299. **Pombo A, Jackson DA, Hollinshead M, Wang Z, Roeder RG, Cook PR.** Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. *EMBO J*. 1999;**18**(8):2241-2253.

300. **Iborra FJ, Pombo A, Jackson DA, Cook PR.** Active RNA polymerases are localized within discrete transcription 'factories' in human nuclei. *J Cell Sci*. 1996;**109**(6):1427-1436.

301. **Wansink DG, Schul W, van der Kraan I, van Steensel B, van Driel R, de Jong L.** Fluorescent labeling of nascent RNA reveals transcription by RNA polymerase II in domains scattered throughout the nucleus. *J Cell Biol*. 1993;**122**(2):283-293.

302. **Wansink DG, Nelissen RL, de Jong L.** In vitro splicing of pre-mRNA containing bromouridine. *Mol Biol Rep*. 1994;**19**(2):109-113.

303. **Bierer BE, Sleckman BP, Ratnofsky SE, Burakoff SJ.** The biologic roles of CD2, CD4, and CD8 in T-cell activation. *Annu Rev Immunol*. 1989;**7**:579-599.

304. **Driscoll PC, Cyster JG, Campbell ID, Williams AF.** Structure of domain 1 of rat T lymphocyte CD2 antigen. *Nature*. 1991;**353**(6346):762-765.

305. **Jones EY, Davis SJ, Williams AF, Harlos K, Stuart DI.** Crystal structure at 2.8 Å resolution of a soluble form of the cell adhesion molecule CD2. *Nature*. 1992;**360**(6401):232-239.

306. **Margineantu DH, Brown RM, Brown GK, Marcus AH, Capaldi RA.** Heterogeneous distribution of pyruvate dehydrogenase in the matrix of mitochondria. *Mitochondrion*. 2002;**1**(4):327-338.

307. **Williams AF, Barclay AN, Clark SJ, Paterson DJ, Willis AC.** Similarities in sequences and cellular expression between rat CD2 and CD4 antigens. *J Exp Med*. 1987;**165**(2):368-380.

308. **Davis SJ, Davies EA, van der Merwe PA.** Mutational analysis of the epitopes recognized by anti-(rat CD2) and anti-(rat CD48) monoclonal antibodies. *Biochem Soc Trans*. 1995;**23**(1):188-194.

309. **Parker MJ, Dempsey CE, Lorch M, Clarke AR.** Acquisition of native β-strand topology during the rapid collapse phase of protein folding. *Biochemistry*. 1997;**36**(43):13396-13405.

310. **Mellon P, Parker V, Gluzman Y, Maniatis T.** Identification of DNA sequences required for transcription of the human alpha 1-globin gene in a new SV40 host-vector system. *Cell*. 1981;**27**(2):279-288.

311. **Myers RM, Rio DC, Robbins AK, Tjian R.** SV40 gene expression is modulated by the cooperative binding of T antigen to DNA. *Cell*. 1981;**25**(2):373-384.

312. **Gluzman Y.** SV40-transformed simian cells support the replication of early SV40 mutants. *Cell*. 1981;**23**(1):175-182.

313. **Jackson DA, Cook PR.** Transcriptionally active minichromosomes are attached transiently in nuclei through transcription units. *J Cell Sci*. 1993;**105**(4):1143-1150.

314. **Pombo A, Ferreira J, Bridge E, Carmo-Fonseca M.** Adenovirus replication and transcription sites are spatially separated in the nucleus of infected cells. *EMBO J*. 1994;**13**(21):5075-5085.

315. **Xu M, Cook PR.** Similar active genes cluster in specialized transcription factories. *J Cell Biol*. 2008;**181**(4):615-623.

316. **Gossen M, Freundlieb S, Bender G, Müller G, Hillen W, Bujard H.** Transcriptional activation by tetracyclines in mammalian cells. *Science*. 1995;**268**(5218):1766-1769.

317. **Urlinger S, Baron U, Thellmann M, Hasan MT, Bujard H, Hillen W.** Exploring the sequence space for tetracycline-dependent transcriptional activators: novel mutations yield expanded range and sensitivity. *Proc Natl Acad Sci U S A*. 2000;**97**(14):7963-7968.

318. **Chalfie M, Tu Y, Euskirchen G, Ward WW, Prasher DC.** Green fluorescent protein as a marker for gene expression. *Science*. 1994;**263**(5148):802-805.

319. **Cormack BP, Valdivia RH, Falkow S.** FACS-optimized mutants of the green fluorescent protein (GFP). *Gene*. 1996;**173(**1):33-38.

320. **Haas J, Park EC, Seed B.** Codon usage limitation in the expression of HIV-1 envelope glycoprotein. *Curr Biol*. 1996;**6**(3):315-324.

321. **Tsien RY.** The green fluorescent protein. *Annu Rev Biochem*. 1998;**67**:509-544.

322. **Fitzgerald M, Shenk T.** The sequence 5'-AAUAAA-3'forms parts of the recognition site for polyadenylation of late SV40 mRNAs. *Cell*. 1981;**24**(1):251-260.

323. **Carswell S, Alwine JC.** Efficiency of utilization of the simian virus 40 late polyadenylation site: effects of upstream sequences. *Mol Cell Biol*. 1989;**9**(10):4248-4258.

324. **Lemm I, Ross J.** Regulation of c-myc mRNA decay by translational pausing in a coding region instability determinant. *Mol Cell Biol*. 2002;**22**(12):3959-3969.

325. **Fernandez J, Yaman I, Huang C, Liu H, Lopez AB, Komar AA, Caprara MG, Merrick WC, Snider MD, Kaufman RJ, Lamers WH, Hatzoglou M.** Ribosome stalling regulates IRES-mediated translation in eukaryotes, a parallel to prokaryotic attenuation. *Mol Cell*. 2005;**17**(3):405-416.

326. **Lawrence JB, Singer RH.** Quantitative analysis of *in situ* hybridization methods for the detection of actin gene expression. *Nucleic Acids Res*. 1985;**13**(5):1777-1799.

327. **Zenklusen D, Singer RH.** Analyzing mRNA expression using single mRNA resolution fluorescent *in situ* hybridization. *Methods Enzymol*. 2010;**470**:641-659.

328. **Pret AM, Fiszman MY.** Sequence divergence associated with species-specific splicing of the nonmuscle β-tropomyosin alternative exon. *J Biol Chem*. 1996;**271**(19):11511-11517.

329. **Zhang G, Taneja KL, Singer RH, Green MR.** Localization of pre-mRNA splicing in mammalian nuclei. *Nature*. 1994;**372**(6508):809-812.

330. **Brody Y, Neufeld N, Bieberstein N, Causse SZ, Böhnlein EM, Neugebauer KM, Darzacq X, Shav-Tal Y.** The *in vivo* kinetics of RNA polymerase II elongation during co-transcriptional splicing. *PLoS Biol*. 2011;**9**(1):e1000573.

331. **Schmidt U, Basyuk E, Robert MC, Yoshida M, Villemin JP, Auboeuf D, Aitken S, Bertrand E.** Real-time imaging of cotranscriptional splicing reveals a

kinetic model that reduces noise: implications for alternative splicing regulation. *J Cell Biol*. 2011;**193**(5):819-829.

332. **Moore MJ.** Nuclear RNA turnover. *Cell*. 2002;**108**(4):431-434.
333. **Danin-Kreiselman M, Lee CY, Chanfreau G.** RNAse III-mediated degradation of unspliced pre-mRNAs and lariat introns. *Mol Cell*. 2003;**11**(5):1279-1289.
334. **Dye MJ, Gromak N, Haussecker D, West S, Proudfoot NJ.** Turnover and function of noncoding RNA polymerase II transcripts. *Cold Spring Harb Symp Quant Biol*. 2006;**71**:275-284.
335. **Wada Y, Ohta Y, Xu M, Tsutsumi S, Minami T, Inoue K, Komura D, Kitakami J, Oshida N, Papantonis A, Izumi A, Kobayashi M, Meguro H, Kanki Y, Mimura I, Yamamoto K, Mataki C, Hamakubo T, Shirahige K, Aburatani H, Kimura H, Kodama T, Cook PR, Ihara S.** A wave of nascent transcription on activated human genes. *Proc Natl Acad Sci U S A*. 2009;**106**(43):18357-61.
336. **Papantonis A, Larkin JD, Wada Y, Ohta Y, Ihara S, Kodama T, Cook PR.** Active RNA polymerases: mobile or immobile molecular machines? *PLoS Biol*. 2010;**8**(7):e1000419.
337. **Shimomura O.** Structure of the chromophore of *Aequorea* green fluorescent protein. *FEBS Lett*. 1979;**104**(2):220-222.
338. **Cody CW, Prasher DC, Westler WM, Prendergast FG, Ward WW.** Chemical structure of the hexapeptide chromophore of the *Aequorea* green-fluorescent protein. *Biochemistry*. 1993;**32**(5):1212-1218.

339. **Niwa H, Inouye S, Hirano T, Matsuno T, Kojima S, Kubota M, Ohashi M, Tsuji FI.** Chemical nature of the light emitter of the *Aequorea* green fluorescent protein. *Proc Natl Acad Sci U S A.* 1996;**93**(24):13617-13622.

340. **Wilson IA, Niman HL, Houghten RA, Cherenson AR, Connolly ML, Lerner RA.** The structure of an antigenic determinant in a protein. *Cell*. 1984;**37**(3):767-778.

341. **Stalder L, Mühlemann O.** Transcriptional silencing of nonsense codon-containing immunoglobulin micro genes requires translation of its mRNA. *J Biol Chem.* 2007;**282**(22):16079-16085.

342. **Colombo B, Felicetti L, Baglioni C.** Inhibition of protein synthesis in reticulocytes by antibiotics. I. Effects on polysomes. *Biochim Biophys Acta*. 1966;**119**(1):109-119.

343. **Kappen LS, Goldberg IH.** Analysis of the two steps in polypeptide chain initiation inhibited by pactamycin. *Biochemistry*. 1976;**15**(4):811-818.

344. **Brodersen DE, Clemons WM Jr, Carter AP, Morgan-Warren RJ, Wimberly BT, Ramakrishnan V.** The structural basis for the action of the antibiotics tetracycline, pactamycin, and hygromycin B on the 30S ribosomal subunit. *Cell*. 2000;**103**(7):1143-1154.

345. **Jaillon O, Bouhouche K, Gout JF, Aury JM, Noel B, Saudemont B, Nowacki M, Serrano V, Porcel BM, Ségurens B, Le Mouël A, Lepère G, Schächter V, Bétermier M, Cohen J, Wincker P, Sperling L, Duret L, Meyer E.** Translational control of intron splicing in eukaryotes. *Nature*. 2008;**451**(7176):359-362.

346. **Miriami E, Sperling R, Sperling J, Motro U.** Regulation of splicing: the importance of being translatable. *RNA.* 2004;**10**(1):1-4.

347. **Kaida D, Motoyoshi H, Tashiro E, Nojima T, Hagiwara M, Ishigami K, Watanabe H, Kitahara T, Yoshida T, Nakajima H, Tani T, Horinouchi S, Yoshida M.** Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat Chem Biol*. 2007;**3**(9):576-583.

348. **Gilbert W.** Why genes in pieces? *Nature.* 1978;**271**(5645):501.

349. **Martin W, Koonin EV.** Introns and the origin of nucleus–cytosol compartmentalization. *Nature*. 2006;**440**(7080):41-45.

350. **Woodson SA.** Ironing out the kinks: splicing and translation in bacteria. *Genes Dev.* 1998;**12**(9):1243-1247.

351. **Semrad K, Schroeder R.** A ribosomal function is necessary for efficient splicing of the T4 phage thymidylate synthase intron *in vivo*. *Genes Dev*. 1998;**12**(9):1327-1337.

352. **Sandegren L, Sjöberg BM.** Self-splicing of the bacteriophage T4 group I introns requires efficient translation of the pre-mRNA in vivo and correlates with the growth state of the infected bacterium. *J Bacteriol.* 2007;**189**(3):980-990.

Chapter 4

Discussion

4.1 Introduction

It is accepted that translation can be coupled to transcription in prokaryotes [353-357]. The idea that such coupling also occurs in eukaryotes has resurfaced numerous times over the last 60 years, but has never been conclusively accepted [358-365]. It is interesting to note that those who first demonstrated that ribosomes associated with the endoplasmic reticulum (ER) were the most prominent sites of protein synthesis in eukaryotic cells [366], also claimed that some protein synthesis occurs in nuclei [358] functionally coupled to transcription [367]. However, the discovery that radio-labelled amino acids were incorporated mainly into the rough endoplasmic reticulum (RER) [366] around the nucleus (then known as the microsomal fraction that could be isolated by differential centrifugation of the cytoplasm) led to the general consensus that no nuclear translation occurs. In the late 1970s, non-coding introns containing many stop codons were discovered to interrupt eukaryotic genes [368], and – as a ribosome translating an intron-containing transcript would inevitably encounter a stop codon within an intron to generate a truncated peptide that might be toxic to the cell – it seemed plausible that Nature might restrict translation to the cytoplasm to ensure that a ribosome would only ever encounter a spliced message (as intron-containing transcripts are never exported to the cytoplasm) [369]. Moreover, there seemed no possible need for nuclear translation if proteins were made efficiently in the cytoplasm. The argument ran that it is very unlikely for cellular evolution to retain such an energy-consuming [370] ‘vestigial’ organelle (a nuclear ribosome) if ribosomes did not actively translate in the nucleus, or perform some

other secondary function (the presence of inactive ribosomal subunits within the nucleus is explained by ribosomal biogenesis occurring within the nucleus [371]). However, the discovery of a mechanism for RNA quality control, NMD [372,373], – which utilizes a translating ribosome to detect a faulty message [374,375] and which occurs in the nucleus [376-378] – provides such a need [379-381].

In this work, I used three different approaches to label sites of translation. I find that substantial amounts of newly-made peptides in the cell are found within nuclei close to transcription sites. These newly-made peptides are short lived, and seem to result from the activity of the proof-reading machinery that checks the quality of newly-made messages.

4.2 Nuclear translation observed by metabolic labelling with Aha and puromycin

I incubated cells with the amino-acid analogue, Aha, or the structural mimic of aminoacyl-tRNA, puromycin, for periods as short as 5 s; both are incorporated into nascent peptides and would be expected to lie close to the ribosome that made them. After fixation, fluors were either 'clicked' on to the Aha-tagged peptides or attached to the puromycylated peptides by indirect immuno-labelling, and the now-fluorescent peptides localized by wide-field or confocal microscopy. Surprisingly, substantial (almost half) fluorescent signal was found in the nucleus, and inhibitors acting on the translating ribosome reduced this signal (**Figs 3.4-3.6**). Given the short labelling period and the known speed of translation (i.e., ~5 amino acids per second [382]), it is highly unlikely that the nuclear signal could result from peptides made in the cytoplasm and transported into nuclei. The simplest interpretation of these observations is that peptide synthesis also occurs in the nucleus.

The two precursors gave different labelling kinetics. With increasing time of incubation in Aha, fluorescent signal increased progressively in both nucleus and cytoplasm (**Fig. 3.4D**), consistent with increasing incorporation in both compartments. With puromycin, the nuclear signal was maximal at 5 s, before it fell significantly by 30 s – unlike the cytoplasmic signal (mainly in the SER) which increases (**Fig. 3.6**). This is simply explained as follows: puromycin incorporation induces premature termination of nascent peptides [383], which then quickly diffuse away from translation sites (whether in the nucleus or cytoplasm) to become concentrated in the SER prior to export [364,384].

The two precursors also gave different nuclear labelling patterns – Aha a diffuse signal (**Figs 3.2-3.4**; though the pattern is punctate in a confocal section – not shown), and puromycin (at least initially) a punctate pattern (**Figs 3.5** and **3.6**). I attribute this to each Aha moiety in a peptide being clicked on to one fluor (at best), whilst each puromycylated peptide is tagged by one primary monoclonal antibody but many fluorescently-labelled secondaries (at best) – to give a more punctate pattern.

4.3 Newly-made CD2 protein lies close to the nascent RNA that encodes it

If some translation occurs coupled to transcription in nuclei, a nascent protein should lie close to the nascent RNA that encoded it. Using a multi-copy system, I switched on expression of a non-nuclear protein, CD2, fixed cells 45 min later, and performed an immuno-FISH experiment to detect CD2 protein and intronic CD2 RNA. CD2 protein was present in many nuclear foci (**Fig. 3.12C**), as was nascent (intronic) CD2 RNA (**Fig. 3.12D**); the ‘merge’ (**Fig. 3.12E**) indicates that some CD2 protein (red) lies close to CD2 RNA (green) to give yellow. I then applied a ‘super-resolution’ colocalization method (**Section 2.3.6**) to demonstrate that nuclear CD2 protein lay

significantly closer to nascent CD2 RNA than expected by chance (**Fig. 3.12F**,**G**). However, the control used for estimating colocalization by chance might not be ideal, and this has been discussed in **Section 4.4**. There are two simple explanations of these results – either CD2 protein was made by cytoplasmic ribosomes and entered the nucleus only after it was made (but then why should it become concentrated close to sites containing nascent CD2 RNA?), or the protein was made co-transcriptionally within the nucleus.

4.4 Nuclear translation is coupled to transcription

Previous work indicates that some nuclear translation occurs closely coupled to transcription in factories [362]. This is supported by the disappearance of the bright puromycylated foci when translation is inhibited by anisomycin, or when transcription is inhibited by DRB (**Fig. 3.8A**). Therefore, I used immunofluorescence coupled to 'super-resolution' localization to see whether puromycylated peptides were found with two markers of factories – the initiating form of RNAPII (which carries a phosphorylated serine 5 on the CTD of the catalytic subunit [385]), and nascent BrRNA [386] (**Figs 3.8B** and **3.9**). Although there was not much overlap of the resulting red and green foci (which would be indicated by yellow in the merge), puromycylated peptides nevertheless lay significantly closer to both markers than expected by chance (i.e., ~85% puromycylated foci lay within 0.75 μm of one or other marker; **Figs 3.8Bv-vi** and **3.9E-F**). A similar kind of analysis revealed that newly-made CD2 protein lay close to newly-made (intronic) CD2 RNA (**Fig. 3.12F**,**G**). I now discuss what might underlie such an 'imperfect' colocalization.

I first consider what might be expected if the two markers were found in the same factory. In HeLa cells, nascent transcripts (and presumably the polymerases that make them) lie on the surface of a ~90-nm protein-rich core – the factory, and the bulk of the

nascent RNA (in human endothelial cells) is found within ~35 nm of the surface [387,388]. If some translation occurs closely coupled to transcription, we might then expect to find puromycylated peptides within a 35-nm shell around a ~90-nm factory, and so would be found in the outer volume within a ~160-nm sphere centred on the factory. Given that the resolution of the light microscope is ~200 nm in the *x* and *y* dimensions (at best), double-immunolabelling the puromycylated peptide and polymerase or BrRNA in the same factory should yield yellow like those in **Figures 8Biv** and **9D**. However, little yellow is seen (despite the 'super-resolution' approach showing significant association).

I next discuss the general problem known as 'antibody blocking' that arises when using indirect labelling to demonstrate colocalization of two 'perfectly' colocalizing antigens [386,389]. When a primary antibody (dimensions ~9 nm) binds to one antigen, it will inevitably block access of the primary antibody being used to detect the second antigen. If red and green fluors attached to secondaries are then used to detect colocalization, yellow will never be seen in the resulting 'merge' (with steric hindrance between secondaries compounding the problem). [Such antibody blocking coupled to conventional fluorescence microscopy can be used to show that two antigens lie within a few nanometers of each other [386,389].] Moreover, in 'super-resolution' analyses like those illustrated in **Figures 8Bv**, **9E**, and **12F**, we would expect such antibody blocking to reduce the PDF seen at the shortest NN distances (i.e., in the first bin) – as is the case.

When determining if two markers lie closer together than expected by chance, a data set was generated by randomly distributing red and green foci at the same density as found experimentally; then NN distances in the randomly-generated set were compared with those found experimentally. However, it could be argued that the randomly-generated set only provides an imperfect control, perhaps because not all the volume might be accessible to the molecules under test [390]. Nevertheless, a protein of the size

of a GFP-tetramer (108 kDa) is free to diffuse throughout nuclear space irrespective of chromatin density [391].

Then, the 'imperfect' colocalization seen might result from (i) sites of nuclear translation being genuinely some distance apart from factories (i.e., up to 750 nm away), and/or (ii) puromycylated peptides (or 'spliced out' CD2 intron) made in factories might diffuse away to accumulate at some nearby site before exiting to the cytoplasm (or before nucleolytic degradation; but see **Refs** [392-394]), and/or the poor resolution of our approach.

4.5 Nuclear translation depends on ribosome activity

In all three approaches I used – Aha and puromycin incorporation plus CD2 expression – the appearance of newly-made nuclear protein was susceptible to translation inhibitors acting directly on the ribosome (**Figs 3.2**, **3.4**, **3.7**, **3.8**, and **3.12**). I now list all other known mechanisms of forming a peptide bond; none have the required sensitivities to the drugs.

1. Post-translational amidation [395-397]. [Any such α-amidation of an amino acid (Aha and puromycin in this case) to the C-terminus of a peptide chain would be independent of tRNA-mediated peptidyl-transferase activity, and thus should not be inhibited by anisomycin (which inhibits translation by acting on the peptidyl-transferase centre of the ribosome).]
2. Transglutamination [398]. [This is another post-translational method of peptide-bond formation (between the amine-group of lysine and carboxamide group of glutamine) by the enzyme, transglutaminase. Such a transglutaminase activity may be imagined to bond Aha or puromycin to free amine groups on peptides, and it should not be inhibited by anisomycin. It is also known not to be inhibited by

cycloheximide [399]. Moreover, transglutaminase activity is Ca^{2+} dependent [398], and the Ca^{2+} concentration is known to be higher in the cytoplasm of HeLa [400,401] (the cells I use for experiments with Aha and puromycin) as compared to the nucleus.]

3. L/F-transferase activity [402,403]. [L/F transferase is a prokaryotic enzyme that normally degrades proteins from the N-terminus; however, it can add amino acids at the N-terminus (but not the C terminus). It has not been found in eukaryotes, but – if present – it would not be expected to lead to puromycin incorporation (which is incorporated only at the C-terminus of peptides as it has an amine-group and so can form a peptide-bond [404-406], but not a carboxyl-group necessary for addition at the N-terminus). Although L/F transferase is stimulated by aminoacyl-tRNA [403], and puromycin is a competitive inhibitor (but is not incorporated into the growing peptide) [407], an equivalent activity cannot be responsible for the incorporation seen.]
4. For a review of all other known mechanisms of non-ribosomal peptide synthesis see **Ref** [408]. None are sensitive to translation inhibitors like anisomycin.

Recently, one of the strongest critics of ribosome-mediated nuclear translation reviewed alternative mechanisms for nuclear peptide synthesis, and could suggest nothing other than a peptidyl-transferase activity of the large subunit (60S) of the eukaryotic ribosome [365] – with arguably the strongest evidence being the presence of this subunit in the nucleus (like the small subunit, it is made in the nucleolus [371]). However, there remains no evidence that the 60S subunit can independently add amino acids to peptides post-translationally, or condense free/bound amino acids with each other.

If translation does not occur in the nucleus, what other explanations might there be for the signals seen on incubation with Aha and puromycin? One possibility is that these

precursors might stick preferentially to some unknown nuclear binding site. This seems unlikely. First, when an Aha pulse is followed by a methionine chase, both nuclear and cytoplasmic signals fall with similar kinetics at both 4˚C and 37˚C (**Section 3.2.2; Fig. 3.3B**,**C**,**D**). If we assume the cytoplasmic signal is due to ribosome-mediated incorporation into peptides, then any hypothetical nuclear binding site must fortuitously release its bound Aha in exactly the same way (at both temperatures). However, the reduction in nuclear and cytoplasmic Aha signal is simply explained by degradation of Aha-tagged (nascent) proteins at 37˚C (but not at 4˚C). Second, the Aha signal increases in presence of the proteasomal inhibitor – MG132 (**Fig. 3.3F**); this is easily explained by proteasomes degrading peptides. Third, the changing patterns seen after a puromycin pulse require complicated changes in the affinity of any hypothetical binding site; after 5 s bright punctate nuclear foci are seen, after 30 s these have become fainter and more diffuse, and by 30 s the signal is more uniformly distributed throughout the nucleus and cytoplasm (**Fig. 3.6**). Such temporal fluctuations are simply explained if puromycin is incorporated into nascent peptides which soon dissociate to become concentrated at export sites in the SER – all as expected (**Fig. 3.6F**).

In the absence of any known mechanisms for adding Aha and puromycin to nascent peptides in the nucleus, and having excluded the trivial possibilities that the two compounds are binding to hypothetical sites within nuclei that have extraordinary properties, I conclude that translating nuclear ribosomes incorporate the two labels in the expected manner.

4.6 Nuclear translation is not induced by stress

As cells are starved of methionine before incubation with Aha, or pre-treated with cycloheximide before adding puromycin, it was possible that stress induced by the pre-

treatments [409,410] might be responsible for the nuclear translation. This is unlikely. First, Aha is still mainly incorporated into nuclei in the absence of starvation (**Fig. 3.2E**). [Methionine starvation for 5-15 min increased Aha incorporation maximally; starvation for 30 min led to less Aha incorporation (data not shown).] Second, a pulse of puromycin gave higher nuclear signal without cycloheximide pre-treatment (**Fig. 3.5**). [However, the cytoplasmic signal increased, and thus the number of puromycylated peptides transported to the SER increased, irrespective of their synthesis site.] I conclude that nuclear translation is not a stress-induced-response.

4.7 Translation in the nucleolus?

David *et al*. [364] observed significant nucleolar labelling after performing an analogous experiment to the one I performed (**Section 3.3.1**) – but my experiment differed in two significant ways: cells were permeabilized after (not before) adding puromycin (to avoid the criticism that cytoplasmic ribosomes might enter nuclei on permeabilization), and the puromycin pulse was for 5 s and not 5 min (to minimize the chance that puromycylated peptides could leave the synthetic site). Like David *et al*. [364], I saw bright nuclear signal (as well as cytoplasmic signal), but – unlike David *et al*. [364] – my nuclear signal was punctate and not diffuse (**Figs 3.5A**,**B** and **3.6B**). This is probably because puromycylated peptides soon leave their synthetic sites to become uniformly distributed throughout the nucleus and concentrated at the peri-nuclear SER. Unlike David *et al*. [364], I found no bright nucleolar signal (even when I did a 5-min puromycin pulse; **Fig. 3.7**). Presumably the differences arise from the effects of permeabilizing cells before adding puromycin and fixation – which might improve access of the anti-puromycin antibodies to the nucleolus (which might then bind non-specifically to the high concentration of nucleolar RNA). I also note that the prior permeabilization

used by David *et al*. [364] extracted essentially all cytoplasmic ribosomes (but presumably not those in the nucleolus), and this would inevitably reduce non-nucleolar signal (and so make it easier to see a faint nucleolar signal against the resulting lower background).

If some translation does indeed occur in nucleoli, what role might it play? Perhaps it reflects a quality-control mechanism; it would seem sensible to check that a newly-assembled ribosome could translate an mRNA before exporting that ribosome to the cytoplasm [411]. Alternatively, the nucleolar signal could simply represent the inevitable background found with any imperfect mechanism – nucleoli contain such a high concentration of ribosomes that it is inevitable that some mRNAs find their way to the nucleolus where they might be translated [412].

4.8 Comparing methods for localizing sites of translation

Translation sites were traditionally detected by growing cells in radioactive amino acids and localizing incorporated label by autoradiography. More recently, higher-resolution and non-radioactive alternatives have been developed involving analogues such as Aha [413-415], Hpg (homopropargylglycine) [414], puromycin [364,416,417], O-propargyl-puromycin [418], as well as amino acids containing non-radioactive but stable isotopes [419]. Unfortunately, the newly-incorporated analogues/isotopes are not always in the same places [420,421]. For example, immunodetection of puromycin [417] and click-detection of O-propargyl-puromycin [418] give different results, and it has been argued the latter gives the most accurate localizations [418,422]. I now review the controversy (which involves labelling for >10 min, and so provides plenty of time for newly-made proteins to move away from their site of synthesis).

First, O-propargyl-puromycin incorporation into newly-synthesized proteins gives a localization pattern similar to that observed after incorporation of Aha and Hpg (i.e., robust signals are seen in the nucleolus, nucleus, and cytoplasm) [413-415]. This is consistent with the high amounts of nuclear and nucleolar proteins detected by mass-spectrometry of newly-synthesized Aha-bearing peptides (using BONCAT) [423], and with the observation of newly-synthesized proteins in the nucleus (detecting using tritiated puromycin) [416]. Such a localization pattern is claimed to be the natural distribution of newly-synthesized proteins in the cell (as other methods give similar results). However, this pattern is unlike that observed by immuno-detecting puromycylated peptides [417,421].

Second, the cycloheximide pre-treatment used with puromycin might distort the pattern. Thus, the pre-treatment was assumed to limit the puromycin-bearing new peptides to actively-translating ribosomes, and permeabilization before fixation was assumed to wash away any free puromycin-bearing peptides [424]. But both could distort the localization pattern, and yield results incomparable with those obtained without cycloheximide pre-treatment [417,418,422].

Third, the probable reason for the robust nuclear and nucleolar signal obtained when using click-chemistry to detect O-propargyl-puromycin-bearing peptides (compared to the relatively low nuclear signal seen when detecting puromycin-bearing peptides using an anti-puromycin antibody) was claimed to result from the better access of the smaller fluorescently-tagged azide compared to the bulkier antibodies [422].

The group that uses puromycin and an anti-puromycin antibody to detect puromycylated proteins support their approach as follows [417,421]: (i) Starving cells of methionine before incubation with Aha or Hpg could stress them to alter the localization

pattern. (ii) Puromycin-bearing peptides detected by the anti-puromycin antibody were mostly localized to cytoplasm rather than plasma membrane as claimed by Liu *et al*. [418].

In my work, I used both Aha and puromycin. The above debate [421,422] may be resolved as follows:

1. Unlike Aha (which is incorporated like the natural precursor, methionine), puromycin causes premature release of nascent peptides [383], which are then exported through the ER-Golgi pathway [364,384]. Thus, on extended incubation (in the order of minutes or hours) in Aha, import of newly-synthesized Aha-tagged peptides from cytoplasm into the nucleus adds to the nascent Aha-bearing peptides synthesized in the nucleus; this increases nuclear signal. On extended incubation in puromycin, the puromycylated peptides dissociate from translating ribosomes (nuclear as well as cytoplasmic) to become concentrated at exit sites on the SER (so causing a higher signal in the peri-nuclear cytoplasm). These differences explain why Aha-bearing peptides tend to be more nuclear. Note that I have shown in **Figure 3.2E** that methionine starvation before the Aha pulse does not distort the localization pattern.
2. After a long puromycin pulse, puromycylated peptides are seen in nuclei using an anti-puromycin antibody (Supplementary Fig. 3 in Schmidt *et al*. [417]); similarly, after a long pulse with O-propargyl-puromycin and click chemistry, nuclear signal is significant (Fig. 2 in Liu *et al*. [418]). Further, work by David *et al*. [364] and my work here, show that the anti-puromycin antibody can label nuclear puromycin-bearing peptides, although access to the nucleolus (if labelling truly reflects specific, and not non-specific, binding) may depend on whether permeabilization is carried out before or after fixation. [Note the robust nucleolar

staining observed by Liu *et al*. [418] using O-propargyl-puromycin, involves fixation with methanol, which is known to permeabilize effectively and precipitate proteins [425].]

3. On extended incubation with puromycin, it is observed that the peri-nuclear space in the cytoplasm has higher signal (from puromycin-bearing peptides) than the nucleus, irrespective of whether puromycin is detected using an anti-puromycin antibody (Supplementary Fig. 3, in Schmidt *et al*. [417]) or O-propargyl-puromycin is detected by click chemistry (Fig. 2 in Liu *et al*. [418]). Any interpretation claiming biased detection of puromycin-bearing peptides (by anti-puromycin antibody) at the plasma membrane is mis-placed, as it is impossible to conclude this from the work of Schmidt *et al*. [417] (Supplementary Fig. 3) and cannot be concluded from this work or David *et al*.'s [364] work.
4. Cycloheximide pre-treatment before a puromycin pulse only partially inhibits release of puromycylated peptides from translating ribosomes [383,426], and slows translation. In other words, cycloheximide should generally reduce the signal seen in both nucleus and cytoplasm. And after puromycin pulses of more than 1 min, the pool of free puromycylated peptides [426] becomes large enough to obscure the underlying translation sites (**Figs 3.6** and **3.7**). Only with puromycin pulses of ~5 s do most nascent puromycylated peptides remain at translation sites (**Figs 3.5**, **3.6B**, **3.8**, and **3.9**).

In conclusion, I suggest that concerns regarding the use of puromycin to localize nascent peptides [365] are unfounded, provided that short pulses of ~5 s are used.

I have observed in published work that the distribution newly-synthesized peptides can differ depending on the detection method. I now discuss these cases.

1. A probe might favour certain cellular compartments. For example, when examining the distribution of Aha-bearing peptides in live cells, the copper used in ‘Cu(I)-catalyzed azid-alkyne cycloaddition’ (CuAAC) [427] is cytotoxic; an alternative is to use ‘strain promoted azide-alkyne cycloaddition’ (SPAAC) [428-430]. SPAAC uses fluorescently-labelled ‘strained cyclooctynes’ as probes, but these are known to have azide-independent binding-affinity for certain cellular structures – in particular those rich in single peptidyl-cysteines (the reduced form of cystines formed by disulphide linkage of cysteines in proteins). [The chemistry involves a thiol-yne addition [431].] Such cysteines will be present in the reducing environment of the eukaryotic cytoplasm [432], and this explains the bias towards cytoplasmic staining by ‘strained cyclooctynes’ [429,433-435]. A second example is provided by the fluorophore – BODIPY (boron-dipyrromethene/4,4-difluoro-5,7-dimethyl-4-bora-3a,4a-diaza-s-indacene), which has been used as either BODIPY-alkyne [414] or BODIPY-cyclooctyne [435] to localize Aha-tagged peptides, and as BODIPY-lysine-tRNA [362] to detect BODIPY-lysine-tagged (nascent) peptides. BODIPY is a known lipophilic dye and is commonly used to detect neutral lipids and lipid droplets (LD), which are present in high concentration in the eukaryotic cytoplasm [436,437]. As a result, BODIPY-tagged probes tend to label the cytoplasm preferentially (as is evident in Figure 1 of Beatty and Tirrall [414], where BODIPY-alkyne – but not other fluorescent-alkynes – preferentially labels the cytoplasm in a way that is inert to competition by methionine or inhibition by anisomycin).
2. Attaching a tag like biotin that is found in a particular compartment of the cell. Biotin-lysine-tRNA has been used to localize biotin-lysine-bearing peptides [362]; however it is known that endogenous biotin is common in the cytosol and

mitochondria [438-440], and this underlies the high cytoplasmic and mitochondrial staining by anti-biotin antibodies or fluorescently-labelled avidin/streptavidin.

4.9 Substantial translation occurs in the nucleus

It is widely accepted that most translation in eukaryotic cells occurs at the RER [366], and that no translation occurs in the nucleus [363]. My work shows that substantial translation occurs in the nucleus. What might be the reason for the discrepancy? I believe there is a simple explanation consistent with all the data. This explanation should be viewed against the background that ~95% newly-made RNA in the nucleus never reaches the cytoplasm as mRNA [441] (much of it is intronic RNA that is degraded co-transcriptionally [442-445]), and >30% of newly-made protein is also destroyed soon after it is made (as many proteins misfold) [446-448]. Although such synthesis and destruction might appear wasteful to us, it nevertheless constitutes an integral part of cell metabolism.

I incubate cells with tagged precursors (i.e., Aha and puromycin) for a shorter period (i.e., 5 s) than used hitherto, and I find that these tags are incorporated into (nascent) peptides – in both the cytoplasm and nucleus (**Figs 3.4** and **3.5**) – that have half-lives of ~50 s (**Fig. 3.3E**). As a ribosome polymerizes ~5 amino acids per second [382], and so completes a typical human protein with ~400 amino acid residues [449,450] in ~80 seconds, this means that incorporated precursors are predominantly found in nascent peptides (i.e., those still associated with the ribosome). Thus, I can class labelled-peptides into two groups:

1. The rapidly-degrading proteins/peptides (similar to RDPs [451]; with a half-life of ~50 s). After labelling for longer than one minute, this significant majority of

peptides will have left the synthetic site (**Fig. 3.6D**), and/or been degraded (**Fig. 3.3E**); then, longer pulses will lead to almost no accumulation of such tagged precursors in newly-made proteins.

2. The stable/slowly-degrading proteins (similar to SDPs [451]). This minority (as seems plausible; **Section 3.2.2**) of proteins (presumably made in the cytoplasm, although I have not disapproved their nuclear synthesis) do fold correctly and persist for many minutes (the half-life of a typical human protein is between 30-120 min [452]). Then longer pulses will inevitably lead to an accumulation of tagged precursors in such newly-made (but no longer nascent) proteins (**Figs 3.2**, **3.3** and **3.6D**).

Given that, after 2 min of Aha-pulse, the nuclear signal is approximately equal to that of cytoplasmic signal (**Fig. 3.2D**), and both signals decay during a chase with similar kinetics, the amount of remaining peptides (at a definite time-point during the chase) is similar in both compartments. However, I have not done any experiments to resolve what fraction of these proteins is rapidly degraded, or are more stable. Therefore, I cannot see what fraction of a typical protein in the cell was made in the nucleus.

This complex overlap in sub-populations of newly-made (including nascent) peptides gives us a limited scope for observing nascent peptides. The way I overcome this limitation is to use short pulses rather than longer ones. In other words, a 5-s pulse provides a better snapshot of where peptides are being made than longer pulses.

4.10 Translating nuclear ribosomes proofread nascent transcripts

We have seen that protein/peptide turnover in both the cytoplasm (due to presence of cytoplasmic peptidases [453]) and the nucleus is rapid (the half-lives of newly-made peptides in both compartments are similar – about 50 s; **Fig. 3.3E**); this makes overall

protein synthesis inefficient [454]. But might the (inefficient) nuclear protein synthesis play some role? The obvious answer (in addition to the proposed role in generating MHC class I antigens [455]) is that it is involved in proof-reading the newly-made message, prior to export to the cytoplasm and mass production of protein [456,381]. Proof-reading involves the (nuclear) NMD pathway which recognizes a premature termination codon (PTC) in a mis-spliced mRNA, and triggers RNA degradation [376-378]. As a translating ribosome is the only known mechanism for detecting a PTC [374,375], it follows that the nuclear peptide synthesis seen results from the action of proof-reading (nuclear) ribosomes.

I tested whether a PTC in an artificial *Cd2* gene triggered RNA degradation *via* NMD. Levels of *Cd2* RNA were monitored using RNA FISH with intronic probes; direct introduction of a PTC into the gene, or addition of SSA [457] to inhibit splicing and so prevent the removal of stop codons in introns, reduced levels of nascent CD2 RNA (**Fig. 3.13C**). As adding cycloheximide and anisomycin reversed this effect (**Fig. 3.13C**), a translating ribosome was involved. Further, this process occurred in the nucleus as the RNA FISH signal was nuclear (**Fig. 3.13B**), as expected when targeting an intron [458-460]. All these results are consistent with a translating nuclear ribosome detecting the PTC and triggering NMD [374,375].

4.11 Ribosomes can translate introns

If a ribosome were to translate an unspliced transcript, it would soon terminate at a stop codon in an intron; this would generate a truncated peptide that would probably be toxic to the cell. Therefore, the cell must find some way of preventing a ribosome from translating an intron [369]. This could be achieved in a number of ways including the exclusion of active ribosomes from the nucleus, the positioning of the ribosome within a

transcription factory so that it could only act on the spliced message, and by marking the intron in some way that would prevent ribosome action. Therefore, I was curious to see if an intron could be translated by a ribosome. Therefore, I transfected a vector encoding the haemagglutinin (HA) tag [461] within a *Cd2* intron into cells, and tried to detect the HA tag by indirect immuno-labelling. Only very low levels were seen (**Fig. 3.14A**); this is expected, as splicing should remove the intron encoding the tag. However, when splicing was prevented using SSA [457], HA was detected (**Fig. 3.14B**). Clearly, ribosomes can translate intronic RNA. Therefore, I suggest that nuclear ribosomes are normally prevented from translating introns because they are so positioned within a transcription factory that they only act on the maturing transcript once it has been spliced (**Fig. 4.1**).

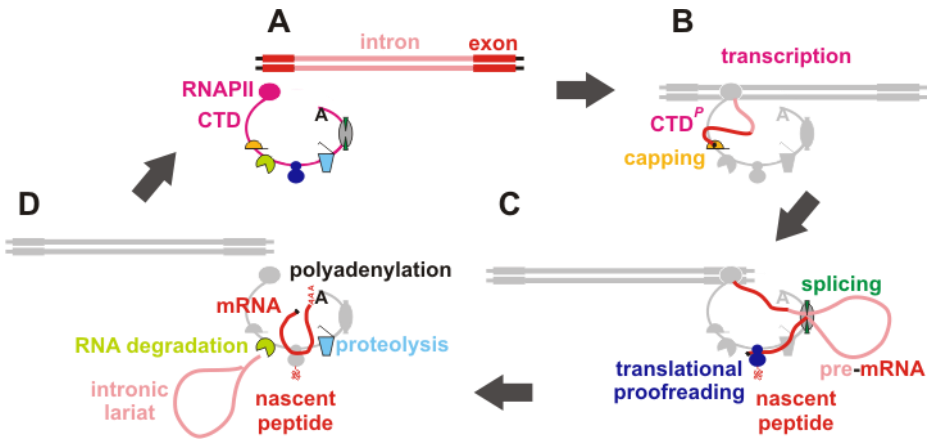


Fig. 4.1. A model for transcript processing. The CTD of RNAPII is known to regulate transcript processing co-transcriptionally. Adapted with permission of [Journal of Cell Science], from (Iborra *et al.* [456], © 2004); permission conveyed through Copyright Clearance Center, Inc.

A. The CTD potentially associates with sites involved in capping [462], RNA degradation [463-465], translational proofreading [362,381], proteolysis [362], splicing and polyadenylation [466].

B. Transcription began as the template (DNA with exons and introns) bound to the polymerizing complex and was reeled in as the transcript was extruded; the CTD is now hyper-phosphorylated, and a cap has been added.

C. The transcript (pre-mRNA) continues to be extruded through a splicing site as the ribosome/NMD machinery begins proofreading the now-spliced message (and so does not read introns that contain many termination codons).

D. Once introns are removed (lariat; subjected to RNA degradation), the transcript is cleaved, polyadenylated, and exported to the cytoplasm (now called mRNA); but if errors are detected, the faulty transcript and peptide produced during proofreading are degraded by nucleases and proteasomes.

4.12 References

353. **Miller OL Jr, Hamkalo BA, Thomas CA Jr.** Visualization of bacterial genes in action. *Science*. 1970;**169**(3943):392-395.

354. **Yanofsky C.** Attenuation in the control of expression of bacterial operons. *Nature.* 1981;289(5800):751-758.

355. **Richardson JP.** Preventing the synthesis of unused transcripts by Rho factor. *Cell.* 1991;64(6):1047-1049.

356. **French SL, Santangelo TJ, Beyer AL, Reeve JN.** Transcription and translation are coupled in Archaea. *Mol Biol Evol*. 2007;**24**(4):893-895.

357. **Proshkin S, Rahmouni AR, Mironov A, Nudler E.** Cooperation between translating ribosomes and RNA polymerase in transcription elongation. *Science.* 2010;**328**(5977):504-508.

358. **Allfrey VG.** Amino acid incorporation by isolated thymus nuclei. I. The role of desoxyribonucleic acid in protein synthesis. *Proc Natl Acad Sci U S A.* 1954;**40**(10):881-885.

359. **Goldstein L.** On the question of protein synthesis by cell nuclei. *Adv. Cell Biol.* 1970;**1**:187-210.

360. **Kuehl L.** Nuclear protein synthesis. 1974; In: Busch H. (ed) *The cell nucleus.* Academic Press, New York pp 345-375.

361. **Goidl JA, Allen WR.** Does protein synthesis occurs within the nucleus? *Trends Biochem Sci.* 1978;**3**:N225-N228.

362. **Iborra FJ, Jackson DA, Cook PR.** Coupled transcription and translation within nuclie of mammalian cells. *Science.* 2001;**293**(5532):1139-1142.

363. **Dahlberg JE, Lund E.** Does protein synthesis occur in the nucleus? *Curr Opin Cell Biol.* 2004;**16**(3):335-338.

364. **David A, Dolan BP, Hickman HD, Knowlton JJ, Clavarino G, Pierre P, Bennink JR, Yewdell JW.** Nuclear translation visualized by ribosome-bound nascent chain puromycylation. *J Cell Biol.* 2012;**197**(1):45-57.

365. **Dahlberg J, Lund E.** Nuclear translation or nuclear peptidyl transferase? *Nucleus*. 2012;**3**(4):1-2.

366. **Allfrey V, Daly MM, Mirsky AE.** Synthesis of protein in the pancreas. II. The role of ribonucleoprotein in protein synthesis. *J Gen Physiol.* 1953;**37**(2):157-175.

367. **Allfrey VG, Mirsky AE, Osawa S.** Protein synthesis in isolated cell nuclei. *J Gen Physiol*. 1957;**40**(3):451-490.

368. **Gilbert W.** Why genes in pieces? *Nature.* 1978;**271**(5645):501.

369. **Martin W, Koonin EV.** Introns and the origin of nucleus–cytosol compartmentalization. *Nature*. 2006;**440**(7080):41-45.

370. **Moss T, Stefanovsky VY.** At the center of eukaryotic life. *Cell*. 2002;**109**(5):545-548.

371. **Leary DJ, Huang S.** Regulation of ribosome biogenesis within the nucleolus. *FEBS Lett*. 2001;**509**(2):145-150.

372. **Chang Y-F, Imam JS, Wilkinson MF.** The nonsense-mediated decay RNA surveillance pathway. *Annu Rev Biochem.* 2007;**76**:51-74.

373. **Brogna S, Wen J.** Nonsense-mediated mRNA decay (NMD) mechanisms. *Nat Struct Mol Biol.* 2009;**16**(2):107-113.

374. **Carter MS, Doscow J, Morris P, Li S, Nhim RP, Sandstedt S, Wilkinson MF.** 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.* 1995;**270**(48):28995-29003.

375. **Stalder L, Mühlemann O.** Transcriptional silencing of nonsense codon-containing immunoglobulin micro genes requires translation of its mRNA. *J Biol Chem.* 2007;**282**(22):16079-16085.

376. **Carter MS, Li S, Wilkinson MF.** A splicing dependent regulatory mechanism that detects translation signals. *EMBO J.* 1996;**15**(21):5965-5975.

377. **Bühler M, Wilkinson MF, Mühlemann O.** Intranuclear degradation of nonsense codon-containing mRNA. *EMBO Rep.* 2002;3(7):646-651.

378. **Maquat LE.** Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. *Nat Rev Mol Cell Biol.* 2004;**5**(2):89-99.

379. **Custódio N, Carmo-Fonseca M, Geraghty F, Pereira HS, Grosveld F, Antoniou M.** Inefficient processing impairs release of RNA from the site of transcription. *EMBO J*. 1999;**18**(10):2855-2866.

380. **de Turris V, Nicholson P, Orozco RZ, Singer RH, Mühlemann O.** Cotranscriptional effect of a premature termination codon revealed by live-cell imaging. *RNA.* 2011;**17**(12):2094-2107.

381. **Wilkinson MF, Shyu AB.** RNA surveillance by nuclear scanning? *Nat Cell Biol.* 2002;**4**(6):E144-E147.

382. **Nielsen PJ, McConkey EH.** Evidence for control of protein synthesis in HeLa cells *via* the elongation rate. *J Cell Physiol*. 1980;**104**(3):269-281.

383. **Williamson AR, Schweet R.** Role of the genetic message in polyribosome function. *J Mol Biol*. 1965;**11**(2):358-372.

384. **Siuta-Mangano P, Lane MD.** 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*. 1981;**256**(5):2094-2097.

385. **Komarnitsky P, Cho EJ, Buratowski S.** Different phosphorylated forms of RNA polymerase II and associated mRNA processing factors during transcription. *Genes Dev*. 2000;**14**(19):2452-2460.

386. **Pombo A, Jackson DA, Hollinshead M, Wang Z, Roeder RG, Cook PR.** Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. *EMBO J*. 1999;**18**(8):2241-2253.

387. **Eskiw CH, Rapp A, Carter DR, Cook PR.** RNA polymerase II activity is located on the surface of protein-rich transcription factories. *J Cell Sci*. 2008;**121**(12):1999-2007.

388. **Papantonis A, Larkin JD, Wada Y, Ohta Y, Ihara S, Kodama T, Cook PR.** Active RNA polymerases: mobile or immobile molecular machines? *PLoS Biol*. 2010;**8**(7):e1000419.

389. **Mason DW, Williams RF.** Kinetics of antibody reactions and the analysis of cell surface antigens. 1986; In: Weir DM *et al.* (eds) *Handbook of experimental immunology*, vol 1. Immunochemistry. Blackwell Scientific Publications, Oxford pp **38**:1-34.

390. **Görisch SM, Wachsmuth M, Ittrich C, Bacher CP, Rippe K, Lichter P.** Nuclear body movement is determined by chromatin accessibility and dynamics. *Proc Natl Acad Sci U S A*. 2004;**101**(36):13221-13226.

391. **Dross N, Spriet C, Zwerger M, Müller G, Waldeck W, Langowski J.** Mapping eGFP oligomer mobility in living cell nuclei. *PLoS One*. 2009;**4**(4):e5041.

392. **Brody Y, Neufeld N, Bieberstein N, Causse SZ, Böhnlein EM, Neugebauer KM, Darzacq X, Shav-Tal Y.** The *in vivo* kinetics of RNA polymerase II elongation during co-transcriptional splicing. *PLoS Biol*. 2011;**9**(1):e1000573.

393. **Brody Y, Shav-Tal Y.** Transcription and splicing: when the twain meet. *Transcription*. 2011;**2**(5):216-220.

394. **Vargas DY, Shah K, Batish M, Levandoski M, Sinha S, Marras SA, Schedl P, Tyagi S.** Single-molecule imaging of transcriptionally coupled and uncoupled splicing. *Cell*. 2011;**147**(5):1054-1065.

395. **Bradbury AF, Smyth DG.** Peptide amidation. *Trends Biochem Sci*. 1991;**16**(3):112-115.

396. **Eipper BA, Stoffers DA, Mains RE.** The biosynthesis of neuropeptides: peptide α-amidation. *Annu Rev Neurosci*. 1992;**15**:57-85.

397. **Eipper BA, Milgram SL, Husten EJ, Yun HY, Mains RE.** Peptidylglycine α-amidating monooxygenase: a multifunctional protein with catalytic, processing, and routing domains. *Protein Sci*. 1993;**2**(4):489-497.

398. **Folk JE, Chung SI.** Transglutaminases. *Methods Enzymol*. 1985;**113**:358-375.

399. **Rice RH, Green H.** Relation of protein synthesis and transglutaminase activity to formation of the cross-linked envelope during terminal differentiation of the cultured human epidermal keratinocyte. *J Cell Biol*. 1978;**76**(3):705-711.

400. **Lipp P, Thomas D, Berridge MJ, Bootman MD.** Nuclear calcium signalling by individual cytoplasmic calcium puffs. *EMBO J*. 1997;**16**(23):7166-73.

401. **He H, Kong SK, Chan KT.** Identification of source of calcium in HeLa cells by femtosecond laser excitation. *J Biomed Opt*. 2010;**15**(5):057010.

402. **Kaji A, Kaji H, Novelli GD.** Soluble amino acid-incorporating system. I. Preparation of the system and nature of the reaction. *J Biol Chem*. 1965;**240**(3):1185-1191.

403. **Abramochkin G, Shrader TE.** Aminoacyl-tRNA recognition by the leucyl/phenylalanyl-tRNA-protein transferase. *J Biol Chem*. 1996;**271**(37):22901-22907.

404. **Yarmolinsky MB, Haba GL.** Inhibition by puromycin of amino acid incorporation into protein. *Proc Natl Acad Sci U S A*. 1959;**45**(12):1721-1729.

405. **Nathans D.** Puromycin inhibition of protein synthesis: incorporation of puromycin into peptide chains. *Proc Natl Acad Sci U S A*. 1964;**51**(4):585-592.

406. **Nissen P, Hansen J, Ban N, Moore PB, Steitz TA.** The structural basis of ribosome activity in peptide bond synthesis. *Science*. 2000;**289**(5481):920-930.

407. **Suto K, Shimizu Y, Watanabe K, Ueda T, Fukai S, Nureki O, Tomita K.** Crystal structures of leucyl/phenylalanyl-tRNA-protein transferase and its complex with an aminoacyl-tRNA analog. *EMBO J*. 2006;**25**(24):5942-5950.

408. **Finking R, Marahiel MA.** Biosynthesis of nonribosomal peptides. *Annu Rev Microbiol*. 2004;**58**:453-488.

409. **Fernandez J, Yaman I, Huang C, Liu H, Lopez AB, Komar AA, Caprara MG, Merrick WC, Snider MD, Kaufman RJ, Lamers WH, Hatzoglou M.** Ribosome stalling regulates IRES-mediated translation in eukaryotes, a parallel to prokaryotic attenuation. *Mol Cell*. 2005;**17**(3):405-416.

410. **Yamasaki S, Anderson P.** Reprogramming mRNA translation during stress. *Curr Opin Cell Biol*. 2008;**20**(2):222-226.

411. **Pederson T, Politz JC.** The nucleolus and the four ribonucleoproteins of translation. *J Cell Biol*. 2000;**148**(6):1091-1095.

412. **Reid DW, Nicchitta CV.** The enduring enigma of nuclear translation. *J Cell Biol*. 2012;**197**(1):7-9.

413. **Beatty KE, Liu JC, Xie F, Dieterich DC, Schuman EM, Wang Q, Tirrell DA.** Fluorescence visualization of newly synthesized proteins in mammalian cells. *Angew Chem Int Ed Engl*. 2006;**45**(44):7364-7367.

414. **Beatty KE, Tirrell DA.** Two-color labeling of temporally defined protein populations in mammalian cells. *Bioorg Med Chem Lett*. 2008;**18**(22):5995-5999.

415. **Dieterich DC, Hodas JJ, Gouzer G, Shadrin IY, Ngo JT, Triller A, Tirrell DA, Schuman EM.** *In situ* visualization and dynamics of newly synthesized proteins in rat hippocampal neurons. *Nat Neurosci*. 2010;**13**(7):897-905.

416. **Gambetti P, Hirt L, Stieber A, Shafer B.** Distribution of puromycin peptides in mouse entorhinal cortex. *Exp Neurol*. 1972;**34**(2):223-228.

417. **Schmidt EK, Clavarino G, Ceppi M, Pierre P.** SUnSET, a nonradioactive method to monitor protein synthesis. *Nat Methods*. 2009;**6**(4):275-277.

418. **Liu J, Xu Y, Stoleru D, Salic A.** Imaging protein synthesis in cells and tissues with an alkyne analog of puromycin. *Proc Natl Acad Sci U S A*. 2012;**109**(2):413-418.

419. **Ong SE.** The expanding field of SILAC. *Anal Bioanal Chem*. 2012;**404**(4):967-976.

420. **Schultze B, Maurer W.** Nuclear and cytoplasmic protein synthesis in various cell types from rats and mice. 1967; In: Goldstein L. (ed) *The control of nuclear activity.* Prentice Hall, New Jersey: pp 319.

421. **Goodman CA, Pierre P, Hornberger TA.** Imaging of protein synthesis with puromycin. *Proc Natl Acad Sci U S A*. 2012;**109**(17):E989.

422. **Salic A.** Reply to Goodman *et al*.: Imaging protein synthesis with puromycin and the subcellular localization of puromycin–polypeptide conjugates. *Proc Natl Acad Sci U S A*. 2012;**109**(17):E990.

423. **Dieterich DC, Link AJ, Graumann J, Tirrell DA, Schuman EM.** Selective identification of newly synthesized proteins in mammalian cells using bioorthogonal noncanonical amino acid tagging (BONCAT). *Proc Natl Acad Sci U S A*. 2006;**103**(25):9482-9487.

424. **David A, Netzer N, Strader MB, Das SR, Chen CY, Gibbs J, Pierre P, Bennink JR, Yewdell JW.** RNA binding targets aminoacyl-tRNA synthetases to translating ribosomes. *J Biol Chem*. 2011;**286**(23):20688-20700.

425. **Fischer AH, Jacobson KA, Rose J, Zeller R.** Fixation and permeabilization of cells and tissues. *CSH Protoc*. 2008;**3**(5):10.1101/pdb.top36.

426. **Obrig TG, Culp WJ, McKeehan WL, Hardesty B.** The mechanism by which cycloheximide and related glutarimide antibiotics inhibit peptide synthesis on reticulocyte ribosomes. *J Biol Chem*. 1971;**246**(1):174-181.

427. **Hein JE, Fokin VV.** Copper-catalyzed azide-alkyne cycloaddition (CuAAC) and beyond: new reactivity of copper(I) acetylides. *Chem Soc Rev*. 2010;**39**(4):1302-1315.

428. **Agard NJ, Prescher JA, Bertozzi CR.** A strain-promoted [3 + 2] azide-alkyne cycloaddition for covalent modification of biomolecules in living systems. *J Am Chem Soc*. 2004;**126**(46):15046-15047.

429. **Beatty KE, Fisk JD, Smart BP, Lu YY, Szychowski J, Hangauer MJ, Baskin JM, Bertozzi CR, Tirrell DA.** Live-cell imaging of cellular proteins by a strain-promoted azide-alkyne cycloaddition. *Chembiochem*. 2010;**11**(15):2092-2095.

430. **Jewett JC, Bertozzi CR.** Cu-free click cycloaddition reactions in chemical biology. *Chem Soc Rev*. 2010;**39**(4):1272-1279.

431. **van Geel R, Pruijn GJ, van Delft FL, Boelens WC.** Preventing thiol-yne addition improves the specificity of strain-promoted azide-alkyne cycloaddition. *Bioconjug Chem*. 2012;**23**(3):392-398.

432. **Rietsch A, Beckwith J.** The genetics of disulfide bond metabolism. *Annu Rev Genet*. 1998;**32**:163-184.

433. **Jewett JC, Sletten EM, Bertozzi CR.** Rapid Cu-free click chemistry with readily synthesized biarylazacyclooctynones. *J Am Chem Soc*. 2010;**132**(11):3688-3690.

434. **Ngo JT, Tirrell DA.** Noncanonical amino acids in the interrogation of cellular protein synthesis. *Acc Chem Res*. 2011;**44**(9):677-685.

435. **Beatty KE, Szychowski J, Fisk JD, Tirrell DA.** A BODIPY-cyclooctyne for protein imaging in live cells. *Chembiochem*. 2011;**12**(14):2137-2139.

436. **Cocchiaro JL, Kumar Y, Fischer ER, Hackstadt T, Valdivia RH.** Cytoplasmic lipid droplets are translocated into the lumen of the *Chlamydia trachomatis* parasitophorous vacuole. *Proc Natl Acad Sci U S A*. 2008;**105**(27):9379-9384.

437. **Listenberger LL, Brown DA.** Fluorescent detection of lipid droplets and associated proteins. *Curr Protoc Cell Biol*. 2007;Chapter 24:Unit 24.2.

438. **Kirkeby S, Moe D, Bøg-Hansen TC, van Noorden CJ.** Biotin carboxylases in mitochondria and the cytosol from skeletal and cardiac muscle as detected by avidin binding. *Histochemistry*. 1993;**100**(6):415-421.

439. **Hollinshead M, Sanderson J, Vaux DJ.** Anti-biotin antibodies offer superior organelle-specific labeling of mitochondria over avidin or streptavidin. *J Histochem Cytochem*. 1997;**45**(8):1053-1057.

440. **Coene ED, Shaw MK, Vaux DJ.** Anti-biotin antibodies offer superior organelle-specific labelling of mitochondria over avidin or streptavidin. *Methods Mol Biol*. 2008;**418**:157-170.

441. **Jackson DA, Pombo A, Iborra F.** The balance sheet for transcription: an analysis of nuclear RNA metabolism in mammalian cells. *FASEB J*. 2000;**14**(2):242-254.

442. **Moore MJ.** Nuclear RNA turnover. *Cell*. 2002;**108**(4):431-434.

443. **Danin-Kreiselman M, Lee CY, Chanfreau G.** RNAse III-mediated degradation of unspliced pre-mRNAs and lariat introns. *Mol Cell*. 2003;**11**(5):1279-89.

444. **Dye MJ, Gromak N, Haussecker D, West S, Proudfoot NJ.** Turnover and function of noncoding RNA polymerase II transcripts. *Cold Spring Harb Symp Quant Biol*. 2006;**71**:275-284.

445. **Wada Y, Ohta Y, Xu M, Tsutsumi S, Minami T, Inoue K, Komura D, Kitakami J, Oshida N, Papantonis A, Izumi A, Kobayashi M, Meguro H, Kanki Y, Mimura I, Yamamoto K, Mataki C, Hamakubo T, Shirahige K, Aburatani H, Kimura H, Kodama T, Cook PR, Ihara S.** A wave of nascent transcription on activated human genes. *Proc Natl Acad Sci U S A*. 2009;**106**(43):18357-61.

446. **Wheatley DN, Giddings MR, Inglis MS.** Kinetics of degradation of "short-" and "long-lived" proteins in cultured mammalian cells. *Cell Biol Int Rep*. 1980;**4**(12):1081-1090.

447. **Schubert U, Antón LC, Gibbs J, Norbury CC, Yewdell JW, Bennink JR.** Rapid degradation of a large fraction of newly synthesized proteins by proteasomes. *Nature*. 2000;**404**(6779):770-774.

448. **Qian SB, Princiotta MF, Bennink JR, Yewdell JW.** Characterization of rapidly degraded polypeptides in mammalian cells reveals a novel layer of nascent protein quality control. *J Biol Chem*. 2006;**281**(1):392-400.

449. **Scherer S.** *A Short Guide to the Human Genome*. 2008; Cold Spring Harbor Laboratory Press, New York.

450. **Brocchieri L, Karlin S.** Protein length in eukaryotic and prokaryotic proteomes. *Nucleic Acids Res*. 2005;**33**(10):3390-3400.

451. **Yewdell JW, Nicchitta CV.** The DRiP hypothesis decennial: support, controversy, refinement and extension. *Trends Immunol*. 2006;**27**(8):368-373.

452. **Yen HC, Xu Q, Chou DM, Zhao Z, Elledge SJ.** Global protein stability profiling in mammalian cells. *Science*. 2008;322(5903):918-923.

453. **Reits E, Griekspoor A, Neijssen J, Groothuis T, Jalink K, van Veelen P, Janssen H, Calafat J, Drijfhout JW, Neefjes J.** Peptide diffusion, protection, and degradation in nuclear and cytoplasmic compartments before antigen presentation by MHC class I. *Immunity*. 2003;**18**(1):97-108.

454. **Dolan BP, Bennink JR, Yewdell JW.** Translating DRiPs: progress in understanding viral and cellular sources of MHC class I peptide ligands. *Cell Mol Life Sci*. 2011;**68**(9):1481-1489.

455. **Dolan BP, Knowlton JJ, David A, Bennink JR, Yewdell JW.** RNA polymerase II inhibitors dissociate antigenic peptide generation from normal viral protein synthesis: a role for nuclear translation in defective ribosomal product synthesis? *J Immunol.* 2010;**185**(11):6728-6733.

456. **Iborra FJ, Escargueil AE, Kwek KY, Akoulitchev A, Cook PR.** Molecular cross-talk between the transcription, translation and nonsense-mediated decay machineries. *J Cell Sci.* 2004;**117**(6):899-906.

457. **Kaida D, Motoyoshi H, Tashiro E, Nojima T, Hagiwara M, Ishigami K, Watanabe H, Kitahara T, Yoshida T, Nakajima H, Tani T, Horinouchi S, Yoshida M.** Spliceostatin A targets SF3b and inhibits both splicing and nuclear retention of pre-mRNA. *Nat Chem Biol*. 2007;**3**(9):576-583.

458. **Zhang G, Taneja KL, Singer RH, Green MR.** Localization of pre-mRNA splicing in mammalian nuclei. *Nature*. 1994;**372**(6508):809-812.

459. **Brody Y, Neufeld N, Bieberstein N, Causse SZ, Böhnlein EM, Neugebauer KM, Darzacq X, Shav-Tal Y.** The *in vivo* kinetics of RNA polymerase II elongation during co-transcriptional splicing. *PLoS Biol*. 2011;**9**(1):e1000573.

460. **Schmidt U, Basyuk E, Robert MC, Yoshida M, Villemin JP, Auboeuf D, Aitken S, Bertrand E.** Real-time imaging of cotranscriptional splicing reveals a kinetic model that reduces noise: implications for alternative splicing regulation. *J Cell Biol*. 2011;**193**(5):819-829.

461. **Wilson IA, Niman HL, Houghten RA, Cherenson AR, Connolly ML, Lerner RA.** The structure of an antigenic determinant in a protein. *Cell*. 1984;**37**(3):767-778.

462. **Fong N, Bentley DL.** Capping, splicing, and 3' processing are independently stimulated by RNA polymerase II: different functions for different segments of the CTD. *Genes Dev*. 2001;**15**(14):1783-1795.

463. **Andrulis ED, Werner J, Nazarian A, Erdjument-Bromage H, Tempst P, Lis JT.** The RNA processing exosome is linked to elongating RNA polymerase II in Drosophila. *Nature*. 2002;**420**(6917):837-841.

464. **Libri D, Dower K, Boulay J, Thomsen R, Rosbash M, Jensen TH.** Interactions between mRNA export commitment, 3'-end quality control, and nuclear degradation. *Mol Cell Biol*. 2002;**22**(23):8254-8266.

465. **Lykke-Andersen J.** Identification of a human decapping complex associated with hUpf proteins in nonsense-mediated decay. *Mol Cell Biol*. 2002;**22**(23):8114-8121.

466. **Maniatis T, Reed R.** An extensive network of coupling among gene expression machines. *Nature*. 2002;**416**(6880):499-506.

Chapter 5

Isolating transcription factories

5.1 Introduction

Mammalian nuclei contain three types of RNA polymerases that are now classified according to the genes they transcribe – RNA polymerases I (RNAPI), II (RNAPII), and III (RNAPIII) [467]. RNAPI makes 45S ribosomal RNA (rRNA), RNAPII transcribes most protein-coding genes to generate messenger RNAs (mRNAs) and some small RNA genes, while RNAPIII produces transfer RNA (tRNA), 5S rRNA, and various other small RNAs (including 7SK small nuclear RNA). The first is active in the nucleolus, and the other two in the nucleoplasm.

Due to the vital role played by these polymerases, there has been a huge effort to evaluate their structure and function. As a result, the structures of the core enzymes of RNAPI, RNAPII, and RNAPIII have been solved, and their functions have been monitored *in vitro* [468]. However, the activity of each polymerase depends on many additional transcription factors [469], while RNAPII also associates with large complexes involved in – for example – capping, splicing, polyadenylation and quality control [470-472]. As we have seen, a number of these mega-complexes cluster to form factories where they transcribe two or more templates [473]. However, the existence of such transcription factories has been controversial, largely because they have not yet been isolated [474].

Two factors have made isolation of transcription factories – and active RNA polymerases – difficult:

1) Only a minority of the RNA polymerases in the cell are active at any moment; the majority form an inactive – and obscuring – pool [475,476]. These inactive forms are easy to isolate, and – as a result – most studies have been performed using them.
2) The active fraction is tightly attached to the nuclear sub-structure [476,477], and this inevitably makes it difficult to isolate.

Therefore, we first permeabilize cells in a 'physiological buffer' [478] so that much of the inactive (soluble) fraction is lost. After isolating nuclei, most chromatin is then detached using DNase, and the nuclei washed. The active fraction – which is still bound to the sub-structure – pellets with the chromatin-depleted nuclei. As caspases are known to deconstruct nuclei *in vivo* [479,480], we now use them to release active polymerases from the sub-structure (core subunits of the three polymerases lack sites recognized by the caspases selected for use, with the exception of RPB9). Three kinds of mega-complexes are then released from the sub-structure, and these were now partially resolved in 2D 'native gels' [481] (**Fig. 5.1.**). Finally, the regions of the gels containing the mega-complexes were excised from the gels, and their proteomes analyzed by mass spectrometry.

Each mega-complex contained the core of a factory, and each complex is named after the type of polymerase it contains. Thus, complex I contained 3 subunits unique to RNAPI, and complex II contains 5 subunits unique to RNAPII (both also contained some shared subunits). Each complex also possessed a characteristic set of proteins. For example, 83% of the proteins in complex I were also present in the proteome of isolated nucleoli [482], complex II contained various general transcription factors (e.g., ATB1, TFIIH subunit 1, CTCF) and specific regulators (e.g., SAFB1 and B2, SF1, CAF1 subunit B) while complex III contained proteins such as Lupus La and EXOSC6. All three

complexes shared many proteins involved in DNA and RNA metabolism (e.g., RNA helicase A and DDX1, hnRNP A2/B1, U2AF65).

Each complex also contained the expected nascent transcripts (detected using RT PCR using intronic probes). For example, relative to complex II, complex I and complex III contained at least 30-fold more nascent 45S rRNA and '7SK RNA and tRNA-leu (CAA)', respectively. Each complex also remained associated with residual DNA fragments of the expected type (inevitably not all DNA is removed by treatment with DNase I); for example, complex I was richest in (residual) ribosomal DNA, complex II in DNA encoding three RNAPII transcripts (*RPS6*, *ARHGAP5*, and *MIR191*), and complex III in the *RN7SK* and *tRNA-leuCAA* genes.

My contribution to this study was to use conventional immuno-fluorescence, the '*in situ* proximity ligation assay (*in situ* PLA)' [483], and 'antibody blocking' [484,485] to confirm – for example – that selected proteins found only in complex II were found close to active RNAPII in fixed cells.

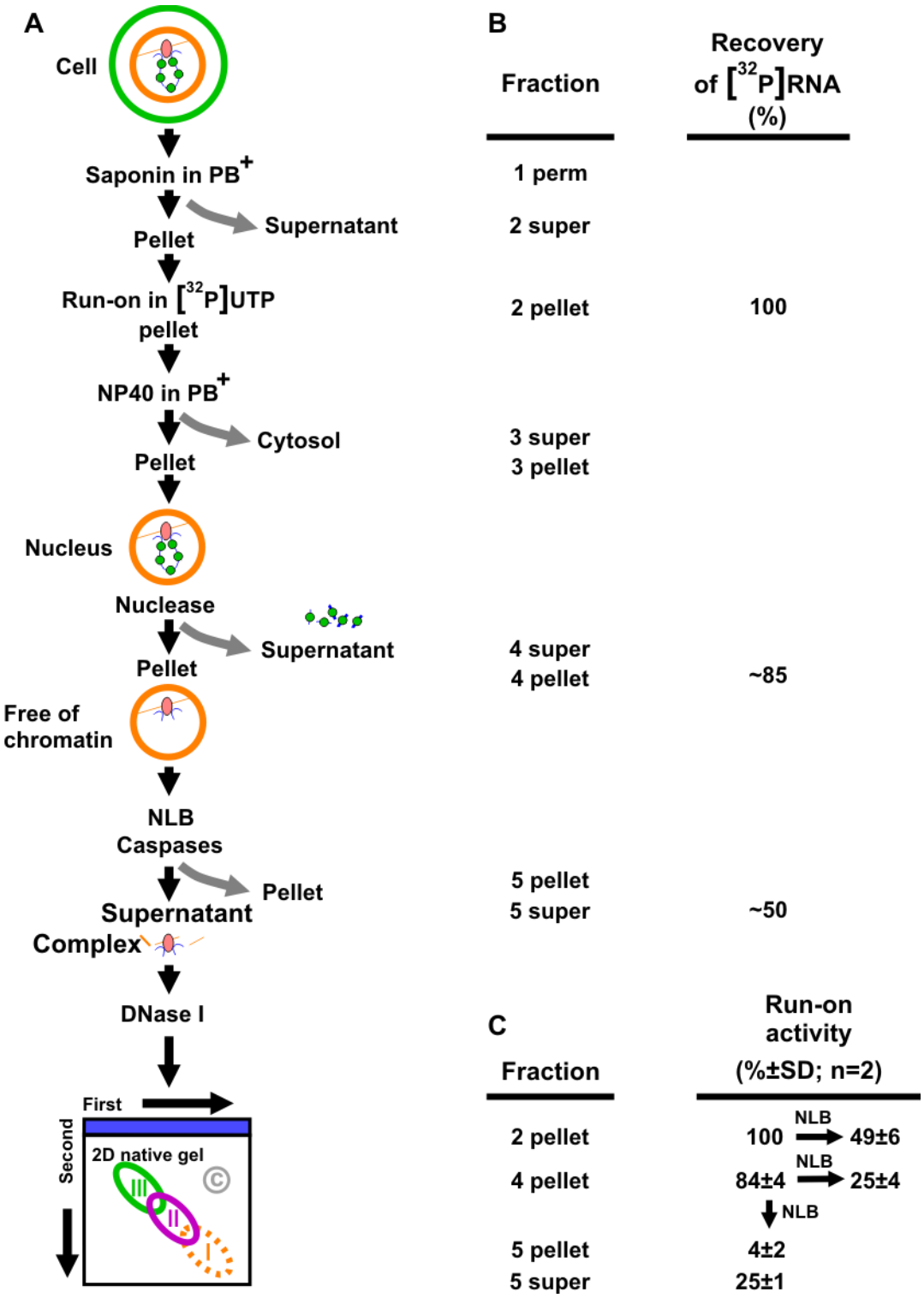


Fig. 5.1. Purification procedure. Adapted from Melnik *et al*. [486].

A. Strategy. Cartoon shows a chromatin loop with nucleosomes (green circle) tethered to a polymerizing complex (oval) attached to the substructure (brown). The cells are permeabilized and in some cases a 'run-on' is performed in [^{32}P]UTP so that nascent RNA can be tracked. The nuclei are then washed with NP-40, most of the chromatin is detached with a nuclease (here, DNase I), the chromatin-depleted nuclei are

resuspended in nuclear lysis buffer (NLB) and polymerizing complexes are released from the substructure with caspases. After pelleting, chromatin associated with polymerizing complexes in the supernatant is degraded with DNase I, and the complexes are partially resolved in two-dimensional (2D) gels (using blue native and native gels in the first and second dimensions, respectively); rough positions of complexes (and a control region, labelled 'C') are shown. Finally, different regions are excised, and their content is analyzed by mass spectrometry.

B. Recovery of [^{32}P] RNA, after including a run-on. Fractions correspond to those at the same level in **A**.

C. Run-on activity assayed later during fractionation (as in **A**, but without run-on at beginning). Different fractions, with names as in **A**, were allowed to extend transcripts by <40 nucleotides in [^{32}P]UTP, and the amount of [^{32}P]RNA per cell was determined by scintillation counting. Fractions '2pellet' and '4pellet' were also resuspended in NLB before run-ons were performed; results indicate that NLB reduces incorporation to half or less. Despite this, '5super' has 25% of the run-on activity of permeabilized cells ('2pellet'), which is equivalent to half of the original (after correction for the effects of NLB).

5.2 Materials and methods

5.2.1 Conventional immuno-fluorescence colocalization

HeLa cells growing on coverslips were processed for conventional immuno-fluorescence (**Section 2.3.2**). Antigens were indirectly immuno-labelled with the rabbit antibody targeting phospho-serine 2 in the C-terminal domain of the largest subunit of polymerase II, and the goat anti-CTCF, followed with a secondary donkey anti-rabbit IgG tagged with Cy3 and chicken anti-goat IgG tagged with Alexa488. Images were acquired with the confocal microscope.

5.2.2 *in situ* PLA

HeLa cells growing on coverslips were processed for *in situ* PLA (**Section 2.3.2**). Antigens indirectly immuno-labelled with various primary antibodies: (i) a mouse monoclonal against RPC32, a subunit of RNA polymerase III, (ii) a rabbit polyclonal

targeting phospho-serine 2 in the C-terminal domain of the largest subunit of RNA polymerase II, (iii) goat polyclonal antibodies against CTCF or EXOSC6, and (iv) a normal goat IgG. This was followed with the secondary antibodies (covalently attached to oligonucleotides), which were either 'Duolink II PLA probe anti-mouse PLUS' or 'Duolink II PLA probe anti-rabbit PLUS' applied with 'Duolink II PLA probe anti-goat MINUS'.

5.2.3 Antibody-blocking assay

HeLa cells growing on coverslips were processed for antibody-blocking assay (**Section 2.3.2**). Antigens indirectly immuno-labelled with various primary antibodies: (i) mouse monoclonal antibodies directed against Sp3 and ATRX, (ii) goat polyclonal antibodies against CTCF, EXOSC6, DDX1, hnRNP A2/B1, Lupus La, U2AF65, and (iii) normal mouse or goat IgG. The detection antibodies were either the mouse monoclonal antibody against RPC32 or the rabbit polyclonal antibody against phospho-serine 2 in the C-terminal domain of the largest subunit of RNA polymerase II. Fluorescently-tagged antibodies were a donkey polyclonal raised against either mouse IgG (tagged with Cy3) or rabbit IgG (tagged with Cy3).

5.3 Results

5.3.1 Conventional immuno-fluorescence colocalization

I begin with an example using conventional immuno-fluorescence to highlight the problems inherent in this approach. **Figure 5.2** shows a representative confocal section of a HeLa cell nucleus (stained blue with DAPI). The 'active' form of RNAPII is immuno-labelled red (using an antibody against the phospho-serine 2 in the C-terminal domain of the largest subunit of RNAPII); this antibody is not expected to detect the majority of

RNAPII which is present in the soluble pool. CTCF – which is found in complex II with RNAPII, is immuno-labelled green. It is evident that the localization pattern of both RNAPII and CTCF within the nucleus is complicated, and determining whether they colocalize significantly or purely as a result of random chance is difficult. A line scan through the middle of this image shows that many peaks of RNAPII (red) overlap peaks of CTCF (green); however, many red and green peaks do not overlap. Perhaps the best statistic of the degree of colocalization in such complex images is that developed by van Steensel *et al*. [487,488]. The test involves separating the red and green channels, linearly shifting the red channel over the green channel in the *x*-direction by one pixel at a time (up to twenty pixels on either side), and plotting the cross-correlation function (CCF) represented by Pearson's correlation coefficient (R). A bell-shaped curve peaking about the median *y*-axis ($\Delta x = 0$) then indicates some colocalization between red and green signals (the higher the peak, the greater the degree of colocalization). If the curve dips at $\Delta x = 0$, then the two patterns are mutually exclusive. And if the curve remains flat at $\Delta x = 0$, results are consistent with a random distribution of signals in one channel relative to the other. The curve obtained with this representative image is almost flat, indicative of only the slightest colocalization of RNAPII and CTCF.

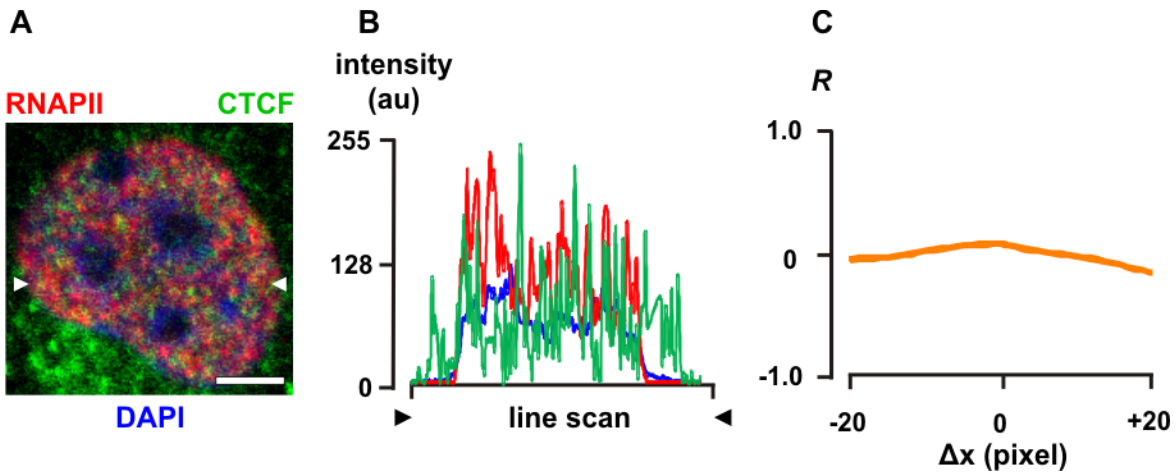


Fig. 5.2. Immuno-fluorescence colocalization assay.

HeLa cells were fixed, RNAPII (detected using an antibody recognizing phospho-serine 2 in the C-terminal repeat in the largest catalytic subunit, followed by a Cy3-tagged secondary) and CTCF (indirectly immuno-labelled with Alexa488) were detected by immuno-fluorescence. Optical sections of DAPI counterstained cells were imaged using a confocal microscope. Adapted from Melnik *et al*. [486].

A. Here, the merged image illustrates a single equatorial confocal section and the complex distributions of RNAPII (red) and CTCF (green) in a HeLa nucleus stained with DAPI (blue). Bar: 5 μm; arrowheads mark position of line scan.

B. The line scan illustrates overlap between red and green signals; it is difficult to establish the degree of co-localization (if any) as so many pixels contain signals.

C. This graph illustrates one approach [487] that can be used to establish the cross-correlation function (CCF) of the red and green components in the image (determined by shifting the red component by ± 20 pixels in the x axis, and plotting Pearson's correlation coefficient, *R*, against Δx. A peak at $\Delta x = 0$ would indicate colocalization between red and green foci, but the peak height is small and the peak width broad, thus remaining ambiguous in suggesting colocalization of RNAPII and CTCF.

Several inter-related reasons make the use of this approach problematic here. First, a minority of CTCF is likely to be found inside a factory, as ≥ 90% of ~10 other transcription factors are in the soluble pool and not in factories [476]. Second, markers like transcription factors are distributed throughout nuclei to yield immuno-fluorescence images in which most pixels contain signal above background (even in a single confocal section); then, two such markers inevitably overlap [489]. Third, the light microscope has a resolution of ~200 nm at best (compared to a ~90-nm factory [490,491]).

5.3.2 The proximity ligation assay (*in situ* PLA)

In situ PLA has been little used to evaluate proximity of protein molecules in cells as it is such a demanding technique; however, a kit is now commercially available. I used it to test the proximity of (i) 'active' RNAPII (using an antibody against the phospho-

serine 2 in the C-terminal domain of the largest subunit of RNAPII) and (ii) RNAPIII (using an antibody against its RPC32 subunit) relative to candidate proteins from complex II (i.e., CTCF) and complex III (i.e., EXOSC6). **Fig. 5.3** shows representative images (2D projections of a *z*-stack of confocal sections) of nuclei stained with DAPI (blue) and rolling circle products marking the localizing proteins labelled red. The *in situ* PLA for CTCF (a complex II protein) with RNAPII (the best marker for a polymerase II factory) shows significantly more red foci than with RNAPIII. The bar graph summarises quantitatively the results of the *in situ* PLA. CTCF has significantly higher association with RNAPII than with RNAPIII; moreover, RNAPIII has significantly higher association with EXOSC6 (a protein found exclusively in complex III) than with CTCF or a non-specific immunoglobulin G (IgG). These results confirm the results obtained by mass-spectrometry; proteins co-associating in a complex are indeed associated in fixed cells with the appropriate polymerase.

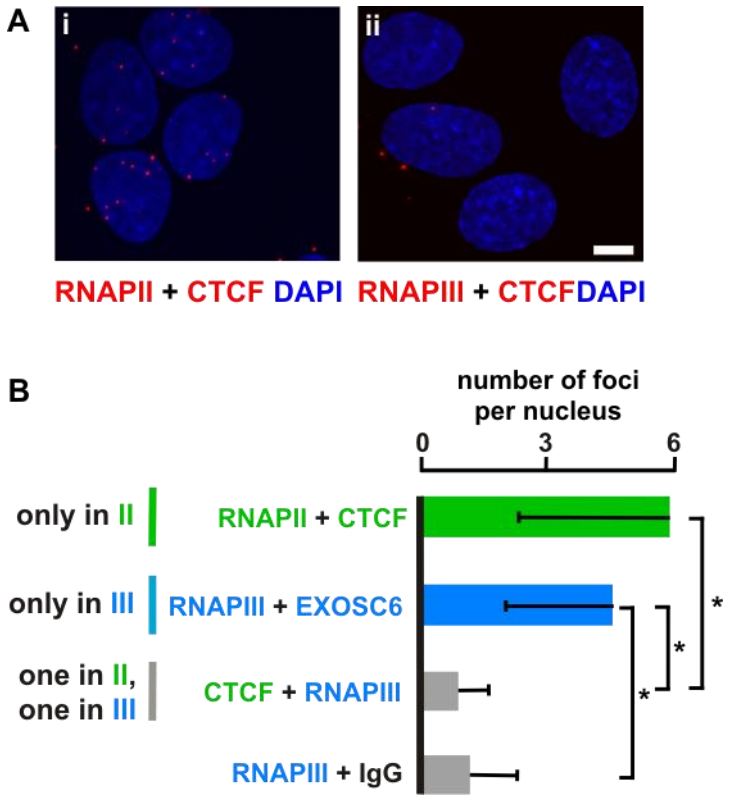


Fig. 5.3. *In situ* proximity ligation assay (*in situ* PLA).

HeLa cells were fixed, RNAPII (detected using an antibody recognizing phospho-serine 2 in the C-terminal repeat in the largest catalytic subunit), RNAPIII (detected using an antibody recognizing the RPC32 subunit), CTCF and EXOSC6 were indirectly immuno-labelled in pairs (RNAPII + CTCF, RNAPIII + EXOSC6, RNAPIII + CTCF, and RNAPIII + IgG) using secondaries conjugated to distinct oligonucleotides, followed by padlock probe hybridization and rolling circle amplification (RCA). *Z*-stacks of DAPI-stained cells were imaged using a confocal microscope. In the 2D projections of these *z*-stacks here, foci (red) mark sites where the two targets lay close together. Adapted from Melnik *et al*. [486].

A. Example.

i. Antibodies targeting two components seen in the same complex (the active form of RNAPII and CTCF, which are both in complex II) yield nuclear foci, whilst

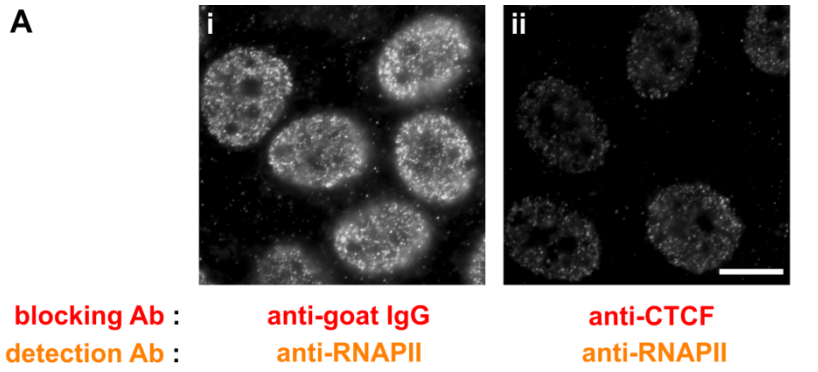
ii. antibodies targeting two proteins seen in different complexes (the RNAPIII subunit RPC32 and CTCF) yield background numbers of foci. Bar: 10 μm.

B. Quantitative results. Antibody pairs against RNAPII + CTCF (only in complex II), and RNAPIII + EXOSC6 (only in complex III) yield significantly more nuclear foci than those given by CTCF + RNAPIII (in different complexes) and RNAPIII and a control immunoglobulin ($P < 0.0001$; n = 27-39 cells; two-tailed Student's *t* test).

5.3.3 Antibody blocking

I further evaluate some associations by assaying if antibody-binding against one of the partners blocks the detection of the other partner (**Fig. 5.4**). An antibody against CTCF (a protein found exclusively in complex II), blocks access of an antibody against the 'active' form of RNAPII, reducing the signal obtained by indirect immuno-fluorescence (non-specific blocking is assessed using a non-specific IgG). Such blocking suggests that CTCF is closely associated with RNAPII. The bar graph summarises quantitatively the results of the antibody-blocking assay that proteins exclusively associated with complex II (CTCF, Sp3, and ATRX), significantly block the detection of RNAPII; similarly, proteins exclusively associated with complex III (Lupus La and

EXOSC6), significantly block the detection of RNAPIII (detected using an antibody against its RPC32 subunit) as compared to blocking RNAPII; again, proteins associated with all the three complexes, significantly block the detection of both RNAPII and RNAPIII. These results again validate the results obtained by mass spectrometry; proteins found together in the same complex are indeed associated with the representative polymerases in cells.



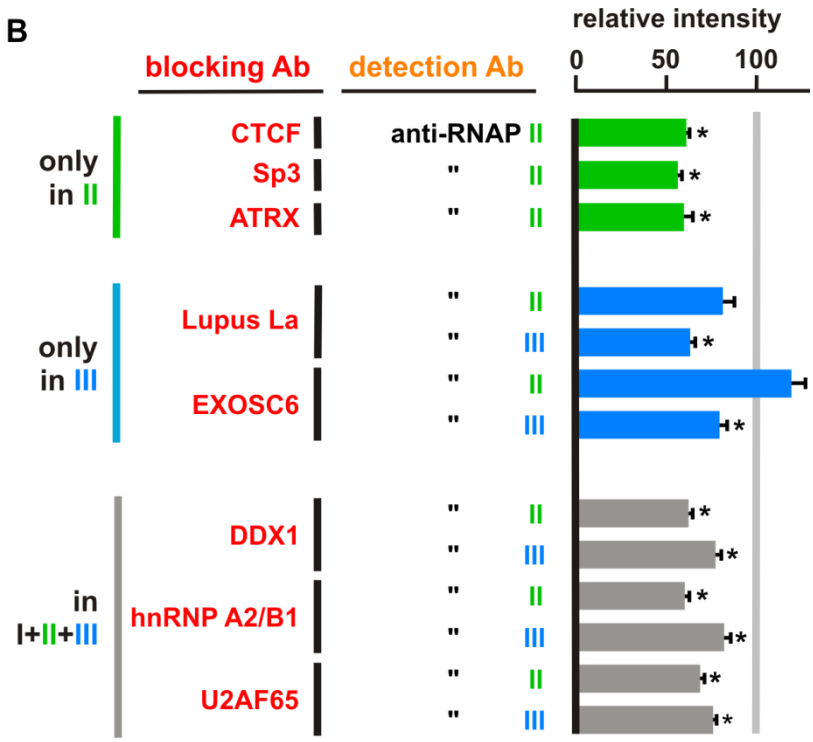


Fig. 5.4. Antibody blocking assay. Adapted from Melnik *et al*. [486].

A. Example of antibody blocking. HeLa cells were fixed, and incubated with two primary antibodies – one targets the active form of RNAPII (i.e., phospho-serine 2 in the C-terminal domain of the catalytic subunit) and the other was either

i. a non-blocking control or

ii. an anti-CTCF.

After incubation with a secondary antibody tagged with Cy3 that targets only the anti-polymerase, images were collected. Active RNAPII is seen in factories throughout the nucleoplasm, but the anti-CTCF reduces signal (**Aii**); such blocking indicates that CTCF lies within a few nm of RNAPII. Bar: 10 μm.

B. Colocalization revealed by antibody blocking. Fluorescent intensities over 5 nuclei in images like those in (**Aii**) were expressed relative to like those in (**Ai**). Antibodies targeting three proteins seen by mass spectrometry only in complex II (CTCF, Sp3, ATRX) block access of anti-RNAPII. Similarly, antibodies targeting two proteins seen only in complex III (Lupus La antigen, EXOSC6) block access of anti-RNAPIII (detected using an anti-RPC32 subunit and a Cy3-conjugated secondary); however, there is some blocking of access of anti-RNAPII by anti-La which (P = 0.02; n = 5 cells; two-tailed Student's t test). Antibodies targeting proteins seen in all three complexes (DDX1, hnRNP A2/B1, U2AF65), block access of antibodies targeting both RNAPII and RNAPIII. Clearly, these pairs of epitopes lie sufficiently close together (i.e., within ≤ 10 nanometers) that the unlabelled antibody can block access of the labelled one. These results confirm those obtained by mass spectrometry and in **Fig. 5.3.** *: difference relative to the value given by a non-blocking control antibody was significant ($P \leq 0.0002$; n = 5 cells; two-tailed Student's t test).

5.4 Discussion

In this study, we described a method for isolating large fragments of transcription factories as overlapping complexes on 2D gels. The three complexes obtained – complexes I, II, and III were named after the polymerases they contained. Mass spectrometry then revealed that each contained a characteristic set of proteins. As part of this collaborative study, I used immuno-fluorescence and its variants to validate some co-associations seen.

Traditionally, ‘immuno-fluorescence colocalization analysis’ has been a technique of choice to verify association of candidate proteins in cells. This has proved very useful when the candidate proteins are found in a few large sites in the cell. Then, when one of the candidate proteins is labelled red and another green, a yellow focus in the resulting image indicates colocalization of the candidate proteins. However, this approach has several disadvantages when protein targets (diameter ~5 nm) [492] are found in many small sites like transcription factories (diameter ~90 nm). First, the diffraction limit of light limits resolution of a microscope to (at best) ~200 nm in the *xy*-axis and ~500 nm in the *z*-axis [493]. Then, the red and green fluors contributing to a yellow focus can lie hundreds of nanometers apart. Second, assessing the ‘degree of colocalization’ in a complex image like that in **Figure 5.2** is difficult. Third, it is often the case that the majority of the nuclear proteins of interest here (e.g., polymerases, transcription factors) are also found outside factories; for example, only about a quarter of the RNAPII in a cell is actively involved in transcription at any given time) [476]. Therefore, I used the following two higher resolution techniques.

1. The *in situ* PLA can detect candidate proteins lying ≤40 nm apart. **Figure 5.3A** demonstrates the quality of PLA signals, which are observed as diffraction-limited fluorescent spots. I observed spots indicative of colocalization of CTCF (exclusive to complex II) with RNAPII, and of EXOSC6 (exclusive to complex III) with RNAPIII – compared to the fewer spots seen with CTCF (or a non-specific IgG) and RNAPIII. This again confirmed the results obtained by mass spectrometry. However, the number of spots observed was very low, raising doubts about the overall detection efficiency of the assay. This limitation might have been expected given its dependency on so many steps after the initial immuno-detection of candidate proteins, including (i) hybridization of the padlock probes with two

different oligonucleotides (attached to the primary antibodies), (ii) ligation of the appropriate ends, and (iii) rolling circle amplification by the φ29 DNA polymerase.

2. I also used the higher-resolution (~10 nm) antibody blocking technique (**Fig. 2.2**). This has the advantage of using traditional immuno-detection protocol without using further molecular processes as in *in situ* PLA, and it depends on the fluorescent detection of only one of the candidate proteins. As evident from **Fig. 5.4**, candidate proteins found exclusively in complex II (CTCF, Sp3, ATRX) blocked the fluorescent detection of RNAPII, those exclusively associated with complex III (Lupus La, EXOSC6) blocked access to RNAPIII (but not RNAPII), and still others associated with all three complexes (DDX1, hnRNP A2/B1, U2AF65) blocked detection of both RNAPII and RNAPIII. Thus, this assay also confirms the results obtained by mass spectrometry.

In conclusion, we partially purified three mega-complexes, each containing one of the three eukaryotic RNA polymerases. All three complexes have masses larger than the largest protein marker commercially available – 8 MDa [494]. Mass spectrometry revealed that each possessed a characteristic proteome, as well as many shared components. Of particular interest here is whether any proteins in complex II play roles in translation. Inspection of the ~500 proteins uncovered in complex II showed that 12 were ribosomal proteins; these could simply be contaminants. This compares with the 14 ribosomal proteins detected in complex I (many ribosomal proteins are expected to be found in complex I, because polymerase I factories are embedded in the nucleolus – the site of ribosome assembly). However, and significantly, complex II contained 19 initiation factors (e.g., subunits of eIF1A, eIF2, eIF3, eIF4A, eIF6), whilst complex I

contained only 2 (eIF4A-3 and eIF6). This is consistent with the ribosomes in complex II (but not complex I) being active [495].

All the experimental results described in this chapter have been published (Melnik *et al.*; [486]). The biochemical purification of factories performed in this study, along with other recent evidence indicating that active RNAPs are immobilized in factories [491], provide strong proof for the existence of transcription factories [496].

5.5 References

467. **Roeder RG.** The eukaryotic transcriptional machinery: complexities and mechanisms unforeseen. *Nat Med*. 2003;**9**(10):1239-1244.

468. **Cramer P, Armache KJ, Baumli S, Benkert S, Brueckner F, Buchen C, Damsma GE, Dengl S, Geiger SR, Jasiak AJ, Jawhari A, Jennebach S, Kamenski T, Kettenberger H, Kuhn CD, Lehmann E, Leike K, Sydow JF, Vannini A.** Structure of eukaryotic RNA polymerases. *Annu Rev Biophys*. 2008;**37**:337-352.

469. **Sikorski TW, Buratowski S.** The basal initiation machinery: beyond the general transcription factors. *Curr Opin Cell Biol*. 2009;**21**(3):344-351.

470. **Maniatis T, Reed R.** An extensive network of coupling among gene expression machines. *Nature*. 2002;**416**(6880):499-506.

471. **Zhong XY, Wang P, Han J, Rosenfeld MG, Fu XD.** SR proteins in vertical integration of gene expression from transcription to RNA processing to translation. *Mol Cell*. 2009;**35**(1):1-10.

472. **Moore MJ, Proudfoot NJ.** Pre-mRNA processing reaches back to transcription and ahead to translation. *Cell*. 2009;**136**(4):688-700.

473. **Cook PR.** A model for all genomes: the role of transcription factories. *J Mol Biol*. 2010;**395**(1):1-10.

474. **Sutherland H, Bickmore WA.** Transcription factories: gene expression in unions? *Nat Rev Genet*. 2009;**10**(7):457-466.

475. **Jackson DA, Iborra FJ, Manders EM, Cook PR.** Numbers and organization of RNA polymerases, nascent transcripts and transcription units in HeLa nuclei. *Mol Biol Cell*. 1998;**9**(6):1523-1536.

476. **Kimura H, Tao Y, Roeder RG, Cook PR.** Quantitation of RNA polymerase II and its transcription factors in an HeLa cell: little soluble holoenzyme but significant amounts of polymerases attached to the nuclear substructure. *Mol Cell Biol*. 1999;**19**(8):5383-5392.

477. **Jackson DA, Cook PR.** Transcription occurs at a nucleoskeleton. *EMBO J*. 1985;**4**(4):919–925.

478. **Pombo A, Jackson DA, Hollinshead M, Wang Z, Roeder RG, and Cook PR.** Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. *EMBO J.* 1999;***18***:2241-2253.

479. **Croft DR, Coleman ML, Li S, Robertson D, Sullivan T, Stewart CL, Olson MF.** Actin-myosin–based contraction is responsible for apoptotic nuclear disintegration. *J Cell Biol*. 2005;**168**(2):245-255.

480. **Crawford ED, Wells JA.** Caspase substrates and cellular remodeling. *Annu Rev Biochem*. 2011;**80**:1055-1087.

481. **Nováková Z, Man P, Novák P, Hozák P, Hodný Z.** Separation of nuclear protein complexes by blue native polyacrylamide gel electrophoresis. *Electrophoresis*. 2006;**27**(7):1277-1287.

482. **Ahmad Y, Boisvert FM, Gregor P, Cobley A, Lamond AI.** NOPdb: Nucleolar Proteome Database - 2008 update. *Nucleic Acids Res*. 2009;**37**:D181-184.

483. **Söderberg O, Leuchowius KJ, Gullberg M, Jarvius M, Weibrecht I, Larsson LG, Landegren U.** Characterizing proteins and their interactions in cells and tissues using the *in situ* proximity ligation assay. *Methods*. 2008;**45**(3):227-232.

484. **Mason DW, Williams RF.** Kinetics of antibody reactions and the analysis of cell surface antigens. 1986; In: Weir DM *et al.* (eds) *Handbook of experimental*

immunology, vol 1. Immunochemistry. Blackwell Scientific Publications, Oxford pp **38**:1-34.

485. **Pombo A, Jackson DA, Hollinshead M, Wang Z, Roeder RG, Cook PR.** Regional specialization in human nuclei: visualization of discrete sites of transcription by RNA polymerase III. *EMBO J*. 1999;**18**(8):2241-2253.

486. **Melnik S, Deng B, Papantonis A, Baboo S, Carr IM, Cook PR.** The proteomes of transcription factories containing RNA polymerases I, II or III. *Nat Methods*. 2011;**8**(11):963-968.

487. **van Steensel B, van Binnendijk EP, Hornsby CD, van der Voort HT, Krozowski ZS, de Kloet ER, van Driel R.** Partial colocalization of glucocorticoid and mineralocorticoid receptors in discrete compartments in nuclei of rat hippocampus neurons. *J Cell Sci*. 1996;**109** (4):787-792.

488. **Rasband, W.S.** ImageJ, http://rsb.info.nih.gov/ij/ U.S. National Institutes of Health, Bethesda, Maryland, USA (1997-2007).

489. **Pombo A, Cook PR.** The localization of sites containing nascent RNA and splicing factors. *Exp Cell Res*. 1996;**229**(2):201-203.

490. **Eskiw CH, Rapp A, Carter DR, Cook PR.** RNA polymerase II activity is located on the surface of protein-rich transcription factories. *J Cell Sci*. 2008;**121**(12):1999-2007.

491. **Papantonis A, Larkin JD, Wada Y, Ohta Y, Ihara S, Kodama T, Cook PR.** Active RNA polymerases: mobile or immobile molecular machines? *PLoS Biol*. 2010;**8**(7):e1000419.

492. **Moran U, Phillips R, Milo R.** SnapShot: Key numbers in biology. *Cell*. 2010;**141**(7):1262-1262.e1.

493. **Abbe E.** *Archiv für Mikroskopische Anat Entwicklungsmech.*1873;**9***:*413–468.

494. **De Ioannes P, Moltedo B, Oliva H, Pacheco R, Faunes F, De Ioannes AE, Becker MI.** Hemocyanin of the molluscan *Concholepas concholepas* exhibits an unusual heterodecameric array of subunits. *J Biol Chem*. 2004;**279**(25):26134-26142.

495. **Iborra FJ, Jackson DA, Cook PR.** The case for nuclear translation. *J Cell Sci*. 2004;**117**(24):5713-5720.

496. **Edelman LB, Fraser P.** Transcription factories: genetic programming in three dimensions. *Curr Opin Genet Dev*. 2012;**22**(2):110-114.

Chapter 6

Specialized 'NFκB' transcription factories

6.1 Introduction

There is growing evidence to support the concept that genes diffuse to 'transcription factories' to be transcribed [497], and that some factories specialize in transcribing specific subsets of genes [498-500]. The nucleolus provides the prototypic example of such a dedicated factory: it contains RNA polymerase I that specializes in transcribing genes encoding rRNA to the exclusion of others [501]. Specialization is known to depend upon various gene promoters which fire when specifically bound to one of the three RNA polymerases; such specialization has been shown using RNA fluorescence *in situ* hybridization (RNA FISH) [502], or partial purification of such specialized factories [503]. All known protein coding genes are transcribed by RNA polymerase II, and the first experiments concluding that factories containing this polymerase involved PTF ('proximal sequence element/PSE' binding transcription factor). This transcription factor was found to be locally concentrated in 'PTF domains' within the nucleus, and these domains brought specific segments of certain chromosomes together to facilitate their transcription [504]. Genes encoding interleukins, when transcribed, are also shown to be anchored by SATB1 (special AT-rich sequence binding protein 1) [505] at T_H2 cytokine-specific transcription factories in thymocytes [506,507]. Similarly, genes encoding all subunits of cytochrome *c* oxidase and glutamatergic neurochemicals are co-regulated by NRF1 (nuclear respiratory factor 1), and transcribed in special factories in neurons [508,509]. Erythroid specific genes also get together in specialized transcription factories [510] with the LCR (locus control region) [511,512],

and KLF1 (Kruppel-like factor 1) [513,514]. Numerous additional experiments performed in the past decade [515], and which use techniques like ChIP (chromatin immuno-precipitation) [516,517] and 3C (chromatin conformation capture) [518] coupled to high-throughput sequencing [519], are now extending such analyses genome wide. These approaches include 3C and it adaptations [520] such as ACT (associated chromosome trap) [521], 4C (3C-on-chip [522], or circular 3C [523]), e4C (enhanced ChIP 3C-on-chip) [513], 5C (3C carbon copy) [524], 6C (combined 3C-ChIP-cloning) [525], Hi-C [526] and ChIA-PET (chromatin interaction analysis by paired-end tag sequencing) [527-529]. All the results obtained with such approaches hint that co-expressed genes tend to be found together in 3D nuclear space, which is consistent with such genes being co-transcribed in specialized transcription factories.

A collaboration with Tatsuhiko Kodama's research group (University of Tokyo; see Wada *et al*. [530]) revealed that the powerful cytokine, TNFα (tumour necrosis factor alpha), signals through NFκB (nuclear factor kappa B) to induce a subset of human genes in HUVECs (human umbilical vein endothelial cells) to fire quickly (within 10 min) and synchronously. One of these genes was *SAMD4A* – a very long gene of 221 kbp. To create a temporal 'high-resolution map' of one transcription cycle of this long gene Wada *et al.* used a 25-mer oligonucleotide tiling microarrays covering its length (**Fig. 6.1**). HUVECs were starved to synchronize them in the G0 phase of the cell cycle, TNFα added, cells collected thereafter every 7.5 min, total RNA isolated, and the RNA hybridized to the tiling microarray bearing probes complementary to *SAMD4A*. In **Figure 6.1**, the height of the needles reflects the intensity of signal detected by each probe (red – intronic; yellow – exonic). A wave of red needles (intronic signal) sweeps down from 'Start' (at 7.5 min) to 'End' (between 75-90 min); this signal is generated by the 'pioneering' polymerase as it initiates (after ~7.5 min) and only reaches the terminus of

this long gene more than one hour later. Probes covering the first few kbps from the transcription start site (*tss*) also yield a signal throughout the period from 7.5-180 min (green rectangle in **Fig. 6.1**), suggesting that other polymerases continuously initiate and subsequently abort transcription in this region.

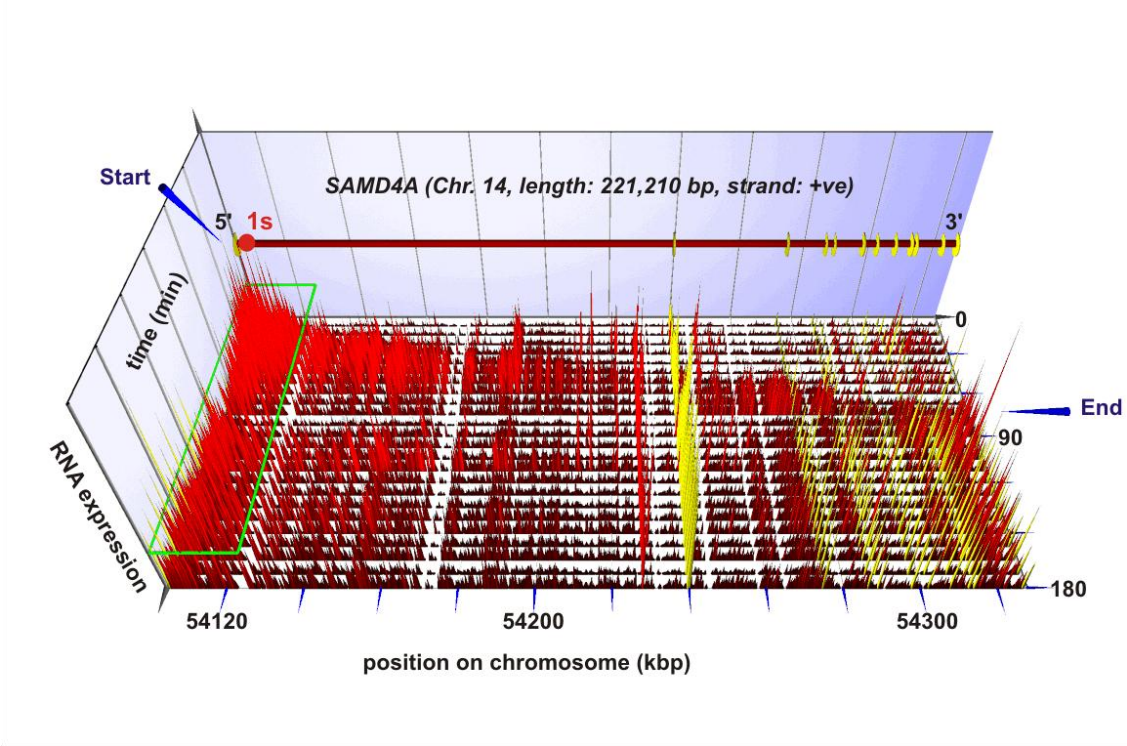


Fig. 6.1. A transcription wave visualized using microarrays.

HUVECs were stimulated with TNFα, samples collected every 7.5 min for 3 h, and total nuclear RNA purified and hybridized to a tiling microarray bearing 25-mers complementary to *SAMD4A*. The vertical axis gives intensity of signal detected by intronic and exonic probes (marked by red and yellow needles, respectively). Gene length and genomic location are shown at top, probe positions within the gene from left to right; and time after stimulation from top to bottom. Purple arrowheads indicate the 'Start' and 'End' of the first wave of transcription that sweeps down the gene; the green rectangle marks the position of probes continuously yielding signal between 7.5–180 min, indicative of recurrent initiation; and the red sphere '1s' indicates the target region for RNA FISH probes used in my experiments. Adapted from Wada *et al*. [530].

Papantonis *et al*. utilized this system to prove that RNA polymerases, when actively transcribing, are fixed in factories [531]. A 3C experiment was done using two responding genes: a short 11-kbp gene – *TNFAIP* (on chromosome 14) – was used as a reference point or 'anchor', and *SAMD4A* – which lies ~50 kbp away from *TNFAIP* on chromosome 14 was used as the 'test gene'. TNFα induces both these genes to fire simultaneously, and while it takes a pioneering polymerase ~75 min to transcribe *SAMD4A*, the much shorter *TNFAIP* is transcribed repeatedly throughout this period. If polymerases track along their templates, then we would expect these two genes rarely to lie next to each other in 3D nuclear space, simply because they lie so far apart on the genetic map (**Fig. 6.2**, **left**). However, if RNA polymerases are fixed in factories when active, and if genes responding to TNFα tend to be transcribed in specialized 'NFκB' factories, we would expect the promoters to come together soon after stimulation; subsequently, the short reference gene should lie next to just the segment of *SAMD4A* being transcribed at that particular moment [530] (**Fig. 6.2**, **right**). Results were consistent with the latter possibility. In other words, active polymerases are fixed in factories.

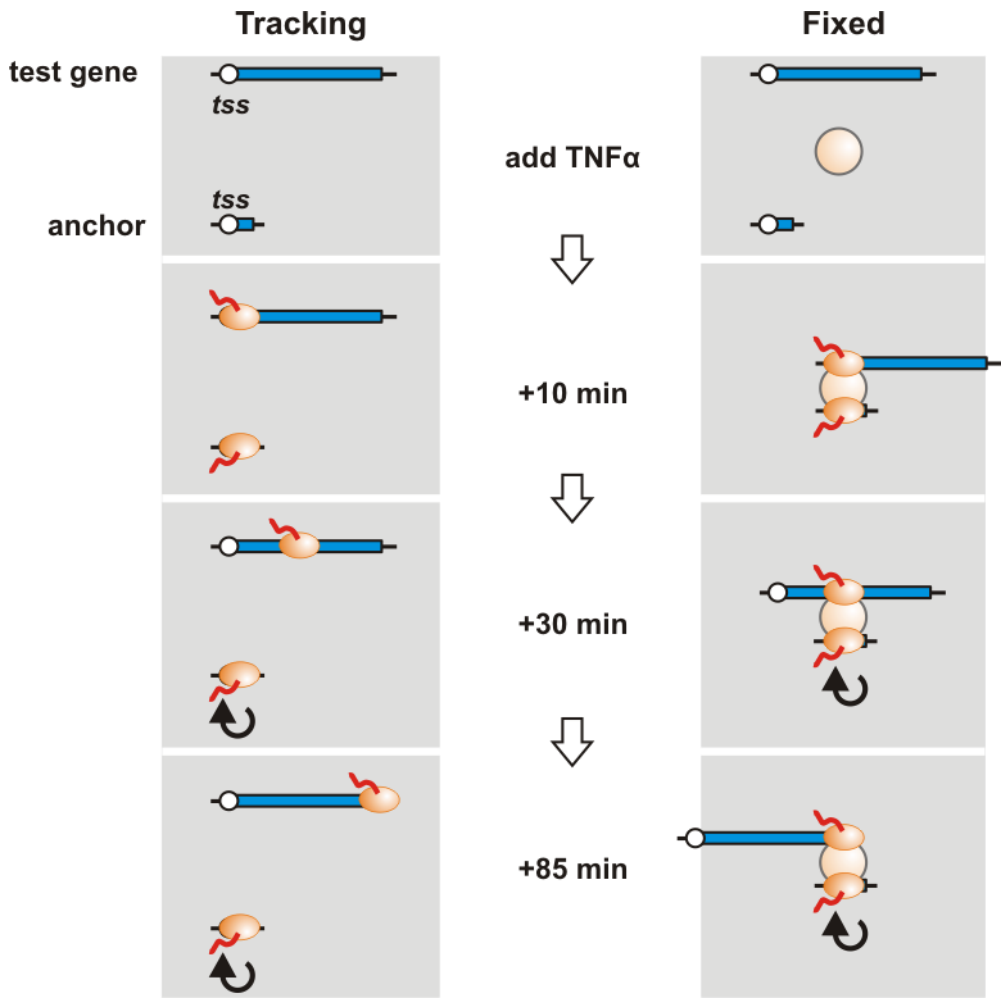


Fig. 6.2. Distinguishing between tracking and fixed RNA polymerases.

Before adding TNFα, both the 'anchor' and 'test gene' are not transcribed. Assuming they lie far apart on the same chromosome, they are unlikely to yield detectable 3C products. Ten min. after adding TNFα, RNA polymerases (ovals) initiate at the *tss* (transcription start site) of both genes, producing nascent transcripts (red). If active polymerases track, it remains unlikely that any part of the two genes will contact each other. However, if the two genes diffuse to one factory (sphere) and are then transcribed by transiently-immobilized polymerases, the two promoters will lie close together. After 30 min, the pioneering polymerase on the short gene has terminated and been replaced by others that continuously transcribe it, while the pioneering polymerase on the long gene has only transcribed one-third of the way into the long gene. If polymerases track, the two genes are still unlikely to be together. But if polymerases are immobilized in a factory, the parts of the two genes currently being transcribed will lie together and yield a 3C product. After 85 min, the pioneering polymerase reaches the terminus of the long gene. If polymerases

track, the two genes will still be apart; if immobilized, the terminus should now contact the short gene. Adapted from Papantonis *et al*. [531].

Results of these experiments also suggested that genes responding to TNFα might be co-transcribed in specialized factories that contained high concentrations of NFκB. We evaluated this hypothesis using a combination of 3C and its variants, ChIP, RNA FISH coupled to high-precision microscopy (~22 nm resolution), and immuno-fluorescence. I used immuno-fluorescence to analyze the temporal dynamics of phosphorylated p65 ($p65^P$ – the 'active' subunit of NFκB). NFκB is normally concentrated in the cytoplasm, but addition of TNFα (indirectly) induces phosphorylation of its p65 subunit, which is then imported into the nucleus. We then imagine that this influx stabilizes binding of responding genes to 'NFκB factories', followed by the transcription of those responding genes. Then, we might also expect nascent *SAMD4A* RNA (detected by RNA FISH) to colocalize with $p65^P$ (detected by immuno-fluorescence) in a factory (**Fig. 6.3**). With these approaches, I (i) confirmed that stimulation induced both phosphorylation of p65 and import into the nucleus, and (ii) demonstrated that nascent transcripts copied from a responding gene did indeed colocalize with $p65^P$ in factories.

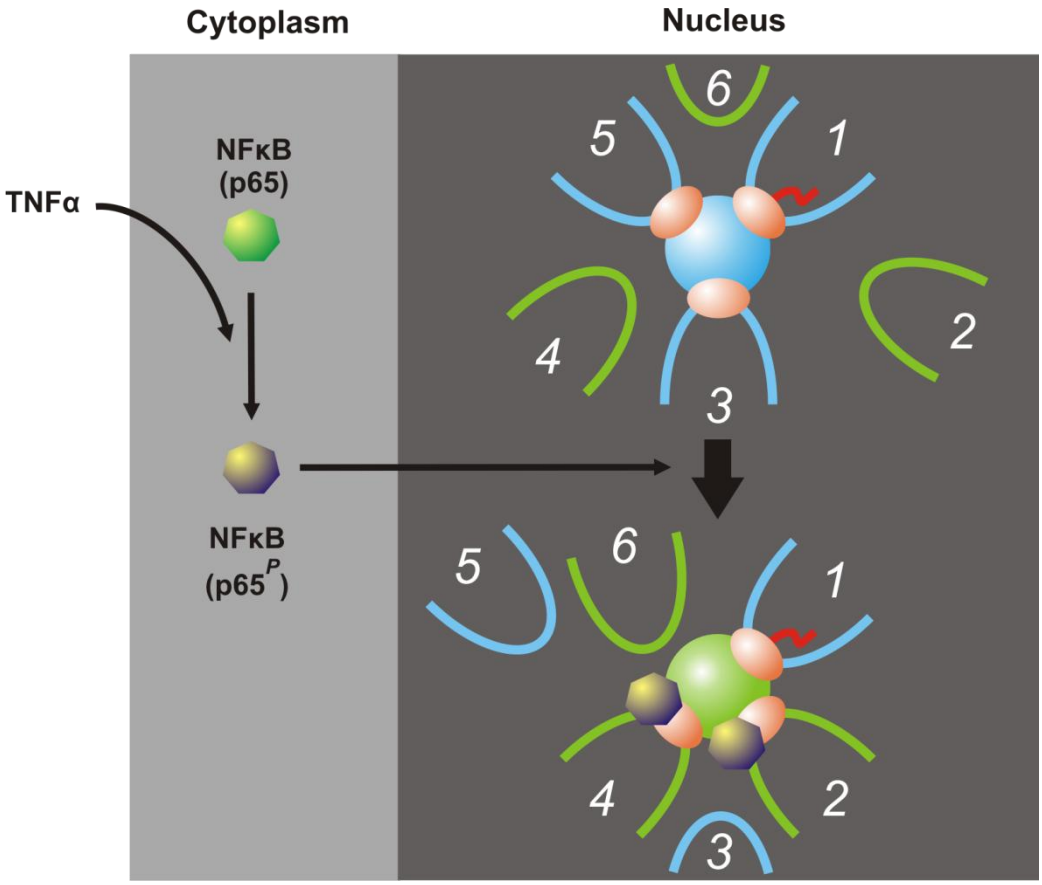


Fig. 6.3. The hypothesis.

NFκB (green) is usually cytoplasmic, and genes *1*, *3*, and *5* are being transcribed in a factory (blue sphere) whilst TNFα-responsive genes *2*, *4* and *6* are unattached and inactive. TNFα induces phosphorylation of the p65 subunit of NFκB (now purple), import into the nucleus, binding to responsive promoters and/or the factory, and – once relevant promoters have diffused through the nucleoplasm and collided with the factory – transcription of responsive genes in what has become a 'specialized factory' (green sphere). As a result, *2* now lies near other responsive NFκB binding genes. Gene *1* is still attached and transcribed, but may later be replaced by responsive gene *6*. I wished to show that TNFα did induce phosphorylation of p65 and import into the nucleus in HUVECs, and that nascent RNA copied from a responsive gene did indeed colocalize with the $p65^{P}$ in factories. Adapted from Papantonis *et al*. [532].

6.2 Materials and methods

6.2.1 Introduction

As described in **Section 6.1**, I performed two different kinds of experiments in this study:

1. Using immuno-fluorescence, I confirmed that $p65^{P}$ entered into the nuclei of HUVECs at the expected time.
2. Using immuno-fluorescence coupled with RNA FISH, I localized nascent RNA molecules relative to sites containing $p65^{P}$ (as we expected the transcription factor to lie close to nascent transcripts copied from its target genes).

6.2.2 Immuno-fluorescence

HUVECs growing on coverslips were starved overnight, treated with TNFα. In some cases 10 μM BAY 11-7082 – an inhibitor of phosphorylation of IκBα, and so (indirectly) of p65 [533] – was also present. Cells were then processed for conventional immuno-fluorescence (**Section 2.3.2**). Phosphorylated (Ser536) p65 was detected by a rabbit monoclonal antibody, followed by an Alexa 488 conjugated donkey anti-rabbit IgG, and mean nuclear fluorescence calculated (background subtracted; measured as the minimum intensity in the area).

6.2.3 Immuno-fluorescence combined with RNA FISH using intronic probes

HUVECs were starved overnight, treated with TNFα and processed for simultaneous RNA FISH with immuno-fluorescence (**Section 2.3.5**). RNA FISH was performed with 25 ng (for *SAMD4A* and *EXT1*) or 10 ng (for *EDN1*) of the relevant probes in hybridization mix. Phosphorylated (Ser536) p65 was detected by a rabbit

monoclonal antibody, followed by an Alexa 488 conjugated donkey anti-rabbit IgG. For the panel in **Figure 6.5** illustrating non-colocalization of nascent *EDN1* RNA and $p65^P$ (detected using Alexa 488 and Cy3, respectively), the red channel is pseudo-coloured green, and the green channel pseudo-coloured red (to facilitate comparison with the other panels). A (red) FISH focus, which is defined as >4 contiguous (90-nm) pixels that contain signal above background (defined as the average intensity of a 50 pixel line-scan across the focus), and is typically 10±3 pixels in size, is deemed to colocalize with a (green) $p65^P$ focus if ≥33% of red pixels overlapped green pixels. *P* values (two-tailed) from Fisher's exact test were calculated using GraphPad (http://www.graphpad.com); they were considered significant when <0.005.

6.3 Results

NFκB is normally bound to IκBα and sequestered in the cytoplasm; on stimulation with TNFα (i) IκBα is phosphorylated [534], and (ii) the p65/RelA subunit of NFκB is phosphorylated at Ser536 (amongst others) [535], leading to NFκB's dissociation from IκBα and import into the nucleus. Therefore, I analyzed the cytokine-induced entry of $p65^P$ into HUVEC nuclei, and **Figure 6.4** illustrates the results.

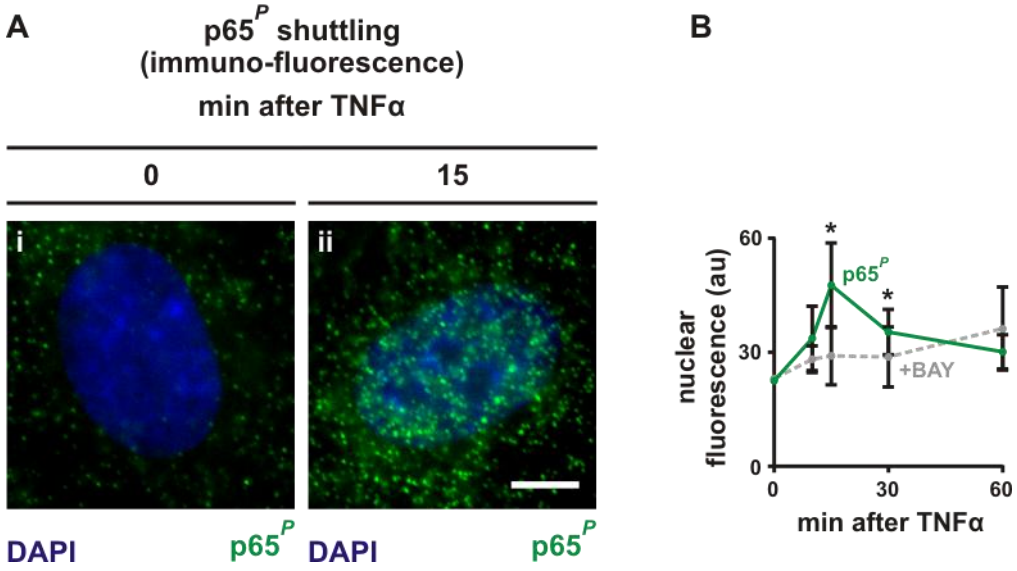


Fig. 6.4. $p65^P$ shuttling after TNFα induction. Adapted from Papantonis *et al*. [532].

HUVECs were grown for 1 h ± BAY 11-7082, TNFα added for 0-60 min, and the phosphorylated p65 subunit of NFκB ($p65^P$) detected by immuno-fluorescence (detected by an antibody recognizing Ser536 phosphorylation on p65, followed by an Alexa488-conjugated secondary).

A. the nucleus (DAPI-stained) contains more $p65^P$ 15 min after induction.

B. the intensity of nuclear fluorescence (arbitrary units, au) is significantly higher (*) after 15 and 30 min (the highest concentration of $p65^P$ seen, coincides with the highest signal seen close to the *SAMD4A* promoter between 7.5-15 min in **Fig. 6.1.**), compared to 0 min or in BAY-treated cells ($P \leq 0.005$; $n = 20$ cells; unpaired two-tailed Student's *t*-test). Bar: 5 μm.

Before stimulation, immunofluorescence reveals that HUVECs contain little $p65^P$ (green) in either the nucleus or cytoplasm (**Fig. 6.4Ai**); however, soon after stimulation with TNFα, significant amounts appear within the nucleus (**Fig. 6.4Aii**). A kinetic analysis reveals that nuclear levels peak after 15 min, and that an inhibitor of IκBα phosphorylation (and so p65 phosphorylation) – BAY 11-7082 [533] – abolishes this (**Fig. 6.4B**). This peak coincides with the maximal rate of initiation on *SAMD4A* as revealed using tiling microarrays (**Fig. 6.1.**; Wada *et al*. [530]). These results confirm that TNFα induces NFκB to enter the nucleus in the expected way; they also indicate that the 'active' form of the transcription factor becomes locally concentrated in discrete nuclear foci – presumably factories specializing in transcribing responsive genes.

Although we presume that the $p65^P$ foci detected in the nucleus are 'NFκB factories', the number of such foci/factories seen is only a lower limit – as it is unlikely that all target molecules are detected in any immuno-labelling experiment. Moreover, $p65^P$ may not always localize just in transcription factories; for example, it is frequently bound to *Alu* repeats [536]. Despite these reservations, these results are nevertheless consistent with the notion that on stimulation $p65^P$ floods into the nucleus to become concentrated in 'NFκB factories'.

If the $p65^P$ foci seen in **Figure 6.4A** represent factories specializing in transcribing genes responding to TNFα, we might also expect those foci to contain nascent RNA encoded by those responsive genes. Therefore, I combined the immuno-fluorescence detection of $p65^P$ with the RNA FISH detection of nascent *SAMD4A* RNA (using probes targeting region 1s, which lies ~1.5 kbp into the first intron) and *EXT1* RNA (using probes targeting a region which lies ~2 kbp into the first intron). Transcripts encoded by *EDN1*, a constitutive and non-responsive gene, serve as a control (as they would be expected to be produced in a different kind of factory). **Figure 6.5** illustrates some representative images, in which a discrete red focus marks either (nascent) *SAMD4A* or *EXT1* RNA. 62% *SAMD4A* transcripts colocalize with $p65^P$ foci, which is much higher than the 16% seen with *EDN1* RNA. This indicates that nascent RNA copied from the TNFα responsive gene is much more likely than that from the control gene to colocalize with the $p65^P$ foci.

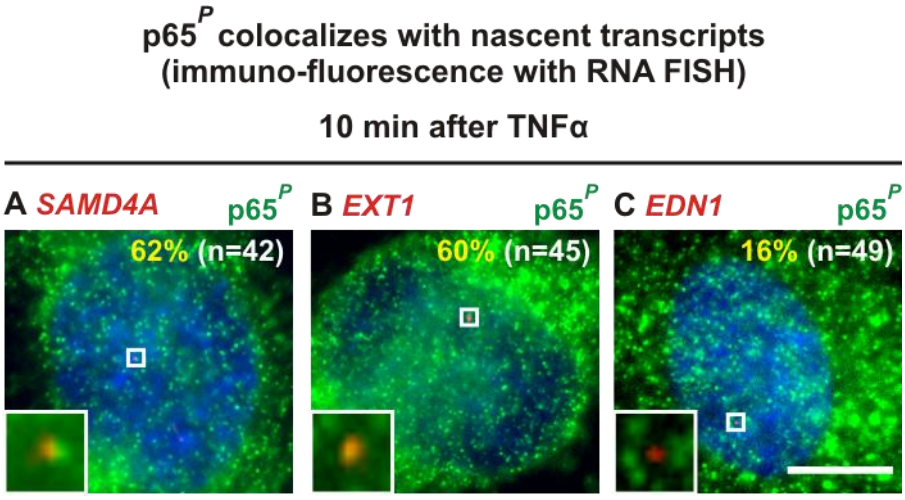


Fig. 6.5. After stimulation, nascent *SAMD4A* and *EXT1* transcripts colocalize with NFκB ($p65^P$) in factories.

HUVECs were grown, TNFα added for 10 min, and nascent RNA copied from *SAMD4A*, *EXT1*, and *EDN1* (a constitutively-expressed gene used as a control), detected by RNA FISH (50-mer probes tagged with either Alexa555 for *SAMD4A* and *EXT1*, or Alexa488 for *EDN1*). The phosphorylated p65 subunit of NFκB

(p65^{P}) detected by immuno-fluorescence (detected by an antibody recognizing Ser536 phosphorylation on p65, followed by either an Alexa488-conjugated secondary when combined with RNA FISH for *SAMD4A* and *EXT1*, or Cy3-conjugated secondary for *EDN1*). In **C**, the *EDN1* RNA FISH is pseudo-coloured red and the p65^{P} immuno-fluorescence is pseudo-coloured green. A single red focus in each nucleus marks nascent RNA copied from one allele; the insets show nascent *SAMD4A* and *EXT1* transcripts colocalize with p65^{P} foci (green). (**A**) 62% *SAMD4A* foci (*n*=42 cells) and (**B**) 60% *EXT1* foci (*n*=45 cells) colocalize with p65^{P} foci, significantly more (*P*=0.004 and 0.003, respectively; two-tailed Fisher's exact test) than the (**C**) 16% found with *EDN1* foci (*n*=49 cells). Bar: 5 μm. Adapted from Papantonis *et al*. [532].

6.4 Discussion

TNFα is a potent cytokine that triggers the NFκB signalling cascade which drives the inflammatory response [537]. NFκB (a complex of p65/RelA and p50 subunits) is normally sequestered in the cytoplasm by IκBα, but stimulation with TNFα induces multiple phosphorylations within the NFκB/IκBα complex, degradation of IκBα, and release of phosphorylated NFκB. The now-active form of NFκB (containing p65^{P}) is imported into the nucleus where it binds to cognate *cis*-elements and induces transcription of TNFα responsive genes. Some responsive genes fire within minutes of adding the cytokine, including *SAMD4A* (on human chromosome 14) and *EXT1* (on human chromosome 8) – two genes used as reference points in this study. If the traditional model for transcription applies, we would then expect RNA polymerase II to first diffuse to the promoter/NFκB complex, and – after initiating – track down the template. In this case, we would have no reason to expect TNFα-responsive genes on different chromosomes to come together in 3D nuclear space, either before or after induction. However, if such responsive genes from different chromosomes do come together on induction, this would be consistent with them being transcribed in specialized 'NFκB factories' (**Fig. 6.3**).

To evaluate whether TNFα responsive genes co-associate with phosphorylated p65 in the 3D space when actively transcribed, I first determined the dynamics of import of $p65^P$ into HUVEC nuclei (**Fig. 6.4**). As expected, there is essentially no $p65^P$ in the cell at 0 min, but within 15 min of TNFα induction it is found in discrete nuclear foci. Subsequently, levels decrease as it is gradually exported (and/or dephosphorylated) by some or all of the various known feedback inhibitory mechanisms involved in the signalling pathway [538]. BAY 11-7082, an inhibitor of phosphorylation of IκBα, eliminates the appearance and import of $p65^P$.

I next used simultaneous intronic-RNA FISH with immuno-fluorescence to localize nascent *SAMD4A* and *EXT1* transcripts relative to the $p65^P$ foci. 10 min after stimulation, both nascent transcripts colocalized with the foci at a higher frequency than those copied from a constitutively-active gene – *EDN1* (**Fig. 6.5**). This result is consistent with the two responsive genes being transcribed in 'NFκB factories'.

Papantonis *et al*. [532] went on to use 3C and its derivatives to show that *SAMD4A* and *EXT1* tended to contact other responsive genes after – but not before – cytokine stimulation. Moreover, analysis of the genome-wide interactome obtained after 'pulling-down' RNA polymerase II and ChIA-PET confirmed that TNFα induced contacts between responsive genes.

The above experiments provide strong circumstantial evidence that TNFα responsive genes come together in specialized 'NFκB factories' when transcribed. Preliminary experiments further showed that another cytokine – $TGF\beta_1$ (transforming growth factor β_1 [539]) – utilizes the same strategy, as it signals through the SMAD group of transcription factors, and induces $TGF\beta_1$ responding genes to co-associate. This points

to the generalization that many – perhaps all – cytokines will signal through factories that specialize in transcribing responding genes.

These results beg answers to many questions, including:

1. Do responsive genes already lie close to factories before induction? Our current working model is that factories are not pre-determined to specialize in one particular way; rather we imagine that they are 'naïve', and frequently visited by various genes. However, when a specific stimulus is given, the concentration of a particular transcription factor might rise, and this increase might stabilize binding of a responsive gene to the 'naïve' factory (and so increase the rate of transcriptional initiation). As a result, the local concentration of that transcription factor in and around the factory will increase, and this will stabilize binding/initiation of additional responsive genes. As this virtuous cycle continues, the naïve factory will develop into a specialized one rich in that transcription factor.
2. How many of 'NFκB' or 'TGFβ_1' factories might exist in one nucleus? Given that contacts between TNFα responsive genes are rare, we imagine that each HUVEC nucleus will contain many tens of 'NFκB factories' after stimulation. This is consistent with many $p65^P$ foci being seen in **Figure 6.5**.
3. How many different types of specialized factories might be present in one nucleus? Given that we have been easily able to detect factories transcribing genes responding to both TNFα and TGFβ_1, we again speculate that many types of specialized factories will exist.

All the experimental results described in this chapter have been published (Papantonis *et al*. [532]).

6.5 References

497. **Edelman LB, Fraser P.** Transcription factories: genetic programming in three dimensions. *Curr Opin Genet Dev*. 2012;**22**(2):110-114.

498. **Chakalova L, Fraser P.** Organization of transcription. *Cold Spring Harb Perspect Biol*. 2010;**2**(9):a000729.

499. **Cook PR.** A model for all genomes: the role of transcription factories. *J Mol Biol*. 2010;**395**(1):1-10.

500. **Bartlett J, Blagojevic J, Carter D, Eskiw C, Fromaget M, Job C, Shamsher M, Trindade IF, Xu M, Cook PR.** Specialized transcription factories. *Biochem Soc Symp*. 2006;**73**:67-75.

501. **Scheer U, Benavente R.** Functional and dynamic aspects of mammalian nucleolus. *Bioessays*. 1990;**12**(1):14-21.

502. **Xu M, Cook PR.** Similar active genes cluster in specialized transcription factories. *J Cell Biol*. 2008;**181**(4):615-623.

503. **Melnik S, Deng B, Papantonis A, Baboo S, Carr IM, Cook PR.** The proteomes of transcription factories containing RNA polymerases I, II or III. *Nat Methods*. 2011;**8**(11):963-968.

504. **Pombo A, Cuello P, Schul W, Yoon JB, Roeder RG, Cook PR, Murphy S.** Regional and temporal specialization in the nucleus: a transcriptionally-active nuclear domain rich in PTF, Oct1 and PIKA antigens associates with specific chromosomes early in the cell cycle. *EMBO J*. 1998;**17**(6):1768-1778.

505. **Cai S, Lee CC, Kohwi-Shigematsu T.** SATB1 packages densely looped, transcriptionally active chromatin for coordinated expression of cytokine genes. *Nat Genet*. 2006;**38**(11):1278-1288.

506. **Spilianakis CG, Flavell RA.** Long-range intrachromosomal interactions in the T helper type 2 cytokine locus. *Nat Immunol*. 2004;**5**(10):1017-27.

507. **Spilianakis CG, Lalioti MD, Town T, Lee GR, Flavell RA.** Interchromosomal associations between alternatively expressed loci. *Nature*. 2005;**435**(7042):637-645.

508. **Dhar SS, Ongwijitwat S, Wong-Riley MT.** Chromosome conformation capture of all 13 genomic loci in the transcriptional regulation of the multisubunit bigenomic cytochrome *c* oxidase in neurons. *J Biol Chem*. 2009;**284**(28):18644-18650.

509. **Dhar SS, Wong-Riley MT.** Chromosome conformation capture of transcriptional interactions between cytochrome c oxidase genes and genes of glutamatergic synaptic transmission in neurons. *J Neurochem*. 2010;**115**(3):676-683.

510. **Osborne CS, Chakalova L, Brown KE, Carter D, Horton A, Debrand E, Goyenechea B, Mitchell JA, Lopes S, Reik W, Fraser P.** Active genes dynamically colocalize to shared sites of ongoing transcription. *Nat Genet*. 2004;**36**(10):1065-1071.

511. **Tolhuis B, Palstra RJ, Splinter E, Grosveld F, de Laat W.** Looping and interaction between hypersensitive sites in the active β-globin locus. *Mol Cell*. 2002;**10**(6):1453-1465.

512. **Carter D, Chakalova L, Osborne CS, Dai YF, Fraser P.** Long-range chromatin regulatory interactions *in vivo*. *Nat Genet*. 2002;**32**(4):623-626.

513. **Schoenfelder S, Sexton T, Chakalova L, Cope NF, Horton A, Andrews S, Kurukuti S, Mitchell JA, Umlauf D, Dimitrova DS, Eskiw CH, Luo Y, Wei CL, Ruan Y, Bieker JJ, Fraser P.** Preferential associations between co-

regulated genes reveal a transcriptional interactome in erythroid cells. *Nat Genet*. 2010;**42**(1):53-61.

514. **Eskiw CH, Fraser P.** Ultrastructural study of transcription factories in mouse erythroblasts. *J Cell Sci*. 2011;**124**(21):3676-3683.

515. **Xu M.** Specialised transcription factories. 2007. D.Phil. Thesis: University of Oxford.

516. **Gilmour DS, Lis JT.** Detecting protein-DNA interactions in vivo: distribution of RNA polymerase on specific bacterial genes. *Proc Natl Acad Sci U S A*. 1984;**81**(14):4275-4279.

517. **Nelson JD, Denisenko O, Bomsztyk K.** Protocol for the fast chromatin immunoprecipitation (ChIP) method. *Nat Protoc*. 2006;**1**(1):179-185.

518. **Dekker J, Rippe K, Dekker M, Kleckner N.** Capturing chromosome conformation. *Science*. 2002;**295**(5558):1306-1311.

519. **Metzker ML.** Sequencing technologies — the next generation. *Nat Rev Genet*. 2010;**11**(1):31-46.

520. **van Steensel B, Dekker J.** Genomics tools for unraveling chromosome architecture. *Nat Biotechnol*. 2010;**28**(10):1089-1095.

521. **Ling JQ, Li T, Hu JF, Vu TH, Chen HL, Qiu XW, Cherry AM, Hoffman AR.** CTCF mediates interchromosomal colocalization between Igf2/H19 and Wsb1/Nf1. *Science*. 2006;**312**(5771):269-272.

522. **Simonis M, Klous P, Splinter E, Moshkin Y, Willemsen R, de Wit E, van Steensel B, de Laat W.** Nuclear organization of active and inactive chromatin domains uncovered by chromosome conformation capture–on-chip (4C). *Nat Genet*. 2006;**38**(11):1348-1354.

523. **Zhao Z, Tavoosidana G, Sjölinder M, Göndör A, Mariano P, Wang S, Kanduri C, Lezcano M, Sandhu KS, Singh U, Pant V, Tiwari V, Kurukuti S, Ohlsson R.** Circular chromosome conformation capture (4C) uncovers extensive networks of epigenetically regulated intra- and interchromosomal interactions. *Nat Genet*. 2006;**38**(11):1341-1347.

524. **Dostie J, Richmond TA, Arnaout RA, Selzer RR, Lee WL, Honan TA, Rubio ED, Krumm A, Lamb J, Nusbaum C, Green RD, Dekker J.** Chromosome Conformation Capture Carbon Copy (5C): A massively parallel solution for mapping interactions between genomic elements. *Genome Res*. 2006;**16**(10):1299-1309.

525. **Tiwari VK, Cope L, McGarvey KM, Ohm JE, Baylin SB.** A novel 6C assay uncovers polycomb-mediated higher order chromatin conformations. *Genome Res*. 2008;**18**(7):1171-1179.

526. **Lieberman-Aiden E, van Berkum NL, Williams L, Imakaev M, Ragoczy T, Telling A, Amit I, Lajoie BR, Sabo PJ, Dorschner MO, Sandstrom R, Bernstein B, Bender MA, Groudine M, Gnirke A, Stamatoyannopoulos J, Mirny LA, Lander ES, Dekker J.** Comprehensive mapping of long-range interactions reveals folding principles of the human genome. *Science*. 2009;**326**(5950):289-293.

527. **Fullwood MJ, Liu MH, Pan YF, Liu J, Xu H, Mohamed YB, Orlov YL, Velkov S, Ho A, Mei PH, Chew EG, Huang PY, Welboren WJ, Han Y, Ooi HS, Ariyaratne PN, Vega VB, Luo Y, Tan PY, Choy PY, Wansa KD, Zhao B, Lim KS, Leow SC, Yow JS, Joseph R, Li H, Desai KV, Thomsen JS, Lee YK, Karuturi RK, Herve T, Bourque G, Stunnenberg HG, Ruan X, Cacheux-Rataboul V, Sung WK, Liu ET, Wei CL, Cheung E, Ruan Y.** An

oestrogen-receptor-α-bound human chromatin interactome. *Nature*. 2009;**462**(7269):58-64.

528. **Fullwood MJ, Han Y, Wei CL, Ruan X, Ruan Y.** Chromatin interaction analysis using paired-end tag sequencing. *Curr Protoc Mol Biol*. 2010;Chapter 21:Unit 21.15.1-25.

529. **Li G, Ruan X, Auerbach RK, Sandhu KS, Zheng M, Wang P, Poh HM, Goh Y, Lim J, Zhang J, Sim HS, Peh SQ, Mulawadi FH, Ong CT, Orlov YL, Hong S, Zhang Z, Landt S, Raha D, Euskirchen G, Wei CL, Ge W, Wang H, Davis C, Fisher-Aylor KI, Mortazavi A, Gerstein M, Gingeras T, Wold B, Sun Y, Fullwood MJ, Cheung E, Liu E, Sung WK, Snyder M, Ruan Y.** Extensive promoter-centered chromatin interactions provide a topological basis for transcription regulation. *Cell*. 2012;**148**(1-2):84-98.

530. **Wada Y, Ohta Y, Xu M, Tsutsumi S, Minami T, Inoue K, Komura D, Kitakami J, Oshida N, Papantonis A, Izumi A, Kobayashi M, Meguro H, Kanki Y, Mimura I, Yamamoto K, Mataki C, Hamakubo T, Shirahige K, Aburatani H, Kimura H, Kodama T, Cook PR, Ihara S.** A wave of nascent transcription on activated human genes. *Proc Natl Acad Sci U S A*. 2009;**106**(43):18357-18361.

531. **Papantonis A, Larkin JD, Wada Y, Ohta Y, Ihara S, Kodama T, Cook PR.** Active RNA polymerases: mobile or immobile molecular machines? *PLoS Biol*. 2010;**8**(7):e1000419.

532. **Papantonis A, Kohro T, Baboo S, Larkin JD, Deng B, Short P, Tsutsumi S, Taylor S, Kobayashi M, Li G, Poh MH, Ruan X, Aburatani H, Ruan Y, Kodama T, Wada Y, Cook PR.** TNFα signals through specialized factories

where responsive coding and miRNA genes are transcribed. *EMBO J*. 2012;**31**(23): 4404-4414.

533. **Pierce JW, Schoenleber R, Jesmok G, Best J, Moore SA, Collins T, Gerritsen ME.** Novel inhibitors of cytokine-induced IkBα phosphorylation and endothelial cell adhesion molecule expression show anti-inflammatory effects *in vivo*. *J Biol Chem*. 1997;**272**(34):21096-21103.

534. **Beg AA, Finco TS, Nantermet PV, Baldwin AS Jr.** Tumor necrosis factor and interleukin-1 lead to phosphorylation and loss of IκBα: a mechanism for NF-κB activation. *Mol Cell Biol*. 1993;**13**(6):3301-3310.

535. **Sakurai H, Chiba H, Miyoshi H, Sugita T, Toriumi W.** IkB kinases phosphorylate NF-kB p65 subunit on serine 536 in the transactivation domain. *J Biol Chem*. 1999;**274**(43):30353-30356.

536. **Antonaki A, Demetriades C, Polyzos A, Banos A, Vatsellas G, Lavigne MD, Apostolou E, Mantouvalou E, Papadopoulou D, Mosialos G, Thanos D.** Genomic analysis reveals a novel NF-κB binding site in *Alu* repetitive elements. *J Biol Chem*. 2011;**286**(44):38768-38782.

537. **Hayden MS, Ghosh S.** NF-κB, the first quarter-century: remarkable progress and outstanding questions. *Genes Dev*. 2012;**26**(3):203-234.

538. **Ruland J.** Return to homeostasis: downregulation of NF-κB responses. *Nat Immunol*. 2011;**12**(8):709-14.

539. **Massagué J.** TGFβ in cancer. *Cell*. 2008;**134**(2):215-230.