Universidade de Lisboa

Faculdade de Ciências

Departamento de Biologia Vegetal

Translational control by an upstream open reading frame in the human erythropoietin transcript

Cristina Maria Botelho da Rocha Barbosa

Doutoramento em Biologia

(Biologia Molecular)

2013

Universidade de Lisboa

Faculdade de Ciências

Departamento de Biologia Vegetal

Translational control by an upstream open reading frame in the human erythropoietin transcript

Cristina Maria Botelho da Rocha Barbosa

Tese orientada pela Doutora Luísa Romão Loison (Instituto Nacional de Saúde Dr. Ricardo Jorge) e pela Professora Doutora Rita Zilhão (Faculdade de Ciências da Universidade de Lisboa), especialmente elaborada para a obtenção do grau de doutor em Biologia (Biologia Molecular)

2013

As opiniões expressas nesta publicação são da exclusiva responsabilidade da sua autora.

You will never know until you try it!

Prefácio

O trabalho de investigação descrito na presente tese de Doutoramento foi realizado na Unidade de Investigação e Desenvolvimento do Departamento de Genética Humana do Instituto Nacional de Saúde Dr. Ricardo Jorge, sob a orientação da Doutora Luísa Romão Loison e co-orientação da Professora Doutora Rita Zilhão, membro da Faculdade de Ciências da Universidade de Lisboa.

Este estudo teve como objetivo principal identificar e caracterizar o modo de regulação da expressão génica do transcrito da eritropoietina humana por uma pequena grelha de leitura a montante da grelha de leitura principal. Foi principalmente importante estudar a sua relevância biológica.

Em conformidade com o disposto no nº 5 do artigo 41º do Regulamento dos Estudos Pós-Graduados da Universidade de Lisboa, deliberação nº 93/2006, publicado em Diário da República, 2º série – Nº 209 – 30 de Outubro de 2006, esta dissertação apresenta-se em língua inglesa e inclui um resumo em português com mais de 1200 palavras (ver Resumo).

Durante a elaboração desta tese tirou-se proveito dos resultados obtidos para publicação numa revista de circulação internacional com arbitragem científica, estando a minha contribuição pessoal devidamente indicada:

Barbosa C and Romão L. Translation of the human erythropoietin transcript is regulated by an upstream open reading frame in response to hypoxia. (under review)

No âmbito do trabalho realizado para a obtenção desta dissertação foi publicado um artigo de revisão numa revista de circulação internacional com arbitragem científica:

Barbosa C, Peixeiro I and Romão L. (2013) Gene expression regulation by upstream open reading frames and human disease. PloS Genetics 9(8): e1003529. doi:10.1371/journal.pgen.1003529.

Durante a elaboração desta tese contribuí para outros projetos em curso no laboratório,

cujos resultados foram publicados em revistas de circulação internacional com arbitragem científica, estando a minha contribuição devidamente indicada:

Peixeiro I, Inácio A, **Barbosa C**, Silva AL, Liebhaber SA and Romão L. (2012) Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations. Nucleic Acids Research 40, 1160–1173. doi:10.1093/nar/gkr820;

Martins R, Proença D, Silva B, **Barbosa C**, Silva AL, Faustino P and Romão L. (2012) Alternative Polyadenylation and Nonsense-Mediated Decay Coordinately Regulate the Human HFE mRNA Levels. PLoS ONE *7*, e35461. doi: 10.1371/journal.pone.0035461

Pereira F, Kong J, Silva AL, Teixeira A, **Barbosa C**, Liebhaber SA and Romão L. Resistance to NMD via the “AUG-proximity effect” reflects specific features of mRNA sequence and structure. Nucleic Acids Research (under review)

O projeto que deu origem à primeira publicação indicada (Peixeiro et al., 2012) foi o ponto de partida para o capítulo IV da presente tese.

Este trabalho foi financiado pela Fundação para a Ciência e a Tecnologia (FCT) na forma de uma Bolsa de Doutoramento com a Referência SFRH/BD/63581/2009, através do Programa de Financiamento Plurianual do Center for Biodiversity, Functional and Integrative Genomics (BioFIG; PEst-OE/BIA/UI4046/2011) e pelo projeto com a referência PTDC/BIM-MED/0352/2012.

Aproveito o presente espaço para agradecer a diversas pessoas essenciais ao desenvolvimento desta tese.
Não poderia começar sem agradecer à minha orientadora, Doutora Luísa Romão, por me dar a oportunidade de fazer uma tese de mestrado iniciando o presente projeto e depois permitindo-me continuar e fazer crescer este meu “bebé”. Foi ao seu lado que foi possível para mim crescer em diversos aspetos que não só o científico e a sua amizade e compreensão nunca passaram despercebidos.
Ao Doutor João Lavinha, na qualidade de responsável da Unidade de I&D do Departamento de Genética Humana do Instituto Nacional de Saúde Dr. Ricardo Jorge, o

meu agradecimento por me ter acolhido nesta instituição. Deixo igualmente o meu apreço pelo seu envolvimento no desenrolar destes anos e sua contribuição para discussões no mínimo estimulantes.

À Professora Doutora Rita Zilhão, na qualidade de orientadora interna, agradeço a sua disponibilidade e incessante interesse na correta evolução da tese.

Os meus atuais e ex-colegas foram peças essenciais para a manutenção de um espírito de perseverança, de interesse científico e de boa disposição no dia-a-dia, sem os quais o concluir deste trabalho teria sido impossível. Desta forma deixo um caloroso agradecimento a: Alexandre Teixeira, Ana Luísa Silva, Ana Morgado, Ana Ramos, Andreia Coelho, Ângela Inácio, Bruno Silva, Cláudia Onofre, Francisco Pereira, Rafaela Lacerda e Rute Martins. O nome da Isabel Peixeiro foi deixado de fora de propósito, visto que não seria justo, após a ligação formada entre nós, esta não ter um agradecimento especial.

A Isabel foi um modelo a nível científico e pessoal. Mais do que os seus ensinamentos a nível prático ela mostrou-me como não ter medo de avançar mesmo quando não há mais ninguém ao nosso lado e como manter o espírito crítico. Para além disso, presenciar a sua gravidez e ver o Tiago crescer foi dos momentos mais marcantes e orgulhosos para mim.

Ao Peter, Paulo e suas “onconetes”, também um agradecimento por proporcionarem um excelente sentido de equipa no instituto, por tornarem o ambiente muito divertido e por todas as ajudas a nível prático. A todos da Unidade de Genética Molecular e da Unidade de Tecnologia e Inovação um muito obrigada pelo apoio e disponibilidade. Sem esquecer um carinhoso obrigada ao Zé Manuel.

Aos meus amigos agradeço a paciência pelas minhas ausências e por alguns momentos de frustração, mas principalmente agradeço os nossos momentos juntos e palavras de alento que sem dúvida contribuíram para me manter sã e consciente de que há outras coisas importantes na vida. Aqui fica a lista sem especial ordem: Joana Cruz, Lara, Marta Perfeita, MaC, Thomas, Rui, Sara Parreira, Teresa Matos, Tolas, Inês Ulrica, Fábio Santos, Rita Ferreira e Filipa Nunes.

Nos últimos dois anos entrou na minha vida uma nova família. Aos pais do Paulo, Florbela e Humberto, aos seus irmãos e respetivas mulheres, Miguel e Cárita, e Pedro e Susana, e às suas sobrinhas, Beatriz e Carolina, um muito sincero obrigada por tudo!

Para os meus pais vai o maior agradecimento possível. Foi o seu amor, os seus

ensinamentos e os valores que me transmitiram que tornaram tudo possível. Muitas das suas palavras encorajaram-me diariamente levando-me a avançar e a sorrir apesar de tudo.

Por fim, ao Paulo. Não há palavras para descrever como o seu apoio e a sua visão do mundo, tão diferente da minha, me ajudou a alcançar mais a todos os níveis. Agradeço-lhe por ter virado a minha vida ao contrário, por me lembrar constantemente das minhas prioridades e por me ajudar a concluir o presente trabalho.

Acknowledgments

I would like to use this blank space to thank several people essential to the development of this thesis.

First of all, I wish to use this opportunity to thank my supervisor, Doctor Luísa Romão, for having given me the opportunity to carry out my Master's thesis, during which, I started this project, then allowing me to proceed to my PhD thesis and watching this baby growing up. It was at her side that it was possible for me to grow up in many ways, other than only scientifically speaking. Her friendship and understanding will never go unnoticed.

To Doctor João Lavinha, as head of the R&D Unit of the *Departamento de Genética Humana* of *Instituto Nacional de Saúde Dr. Ricardo Jorge*, my acknoledge for having welcomed me in this institution. I also appreciate his participation and very stimulating contribution to scientific discussions throughout these years.

To Professor Doctor Rita Zilhão, my internal supervisor, I thank for her availability and interest in the proper evolution of my thesis

My current and former lab mates were essential for the maintenance of a spirit of perseverance, of scientific interest and willingness day by day. Without them the conclusion of this work would have been impossible. Thus, I am most thankful to: Alexandre Teixeira, Ana Luísa Silva, Ana Morgado, Ana Ramos, Andreia Coelho, Ângela Inácio, Bruno Silva, Cláudia Onofre, Francisco Pereira, Rafaela Lacerda and Rute Martins. The name Isabel Peixeiro was left out on purpose, since it would not be fair on her if I would not endorse her a special acknowledge after the bond we have created.

Isabel was a scientific and personal role model. She taught many things at technical level, but more than that, she shown me how to carry on even when there is nobody on our side and how to maintain critical thinking. Besides, witnessing her pregnancy and watching Tiago growing up, were the most memorable and proud moments.

To Peter, Paulo and their “onconetes”, I thank for providing an excellent team spirit at the institute, for making the working environment a lot of funnier and for all the help at practical level. I thank everyone in the Molecular Genetics and the Technology and Innovation Units for all the support and availability. I also wish to address my warm thank you to Zé Manuel.

To my friends I thank their tolerance towards my absence and their support during moments of frustration, but mostly I thank them for our moments together and their words of encouragement, which undoubtedly contributed to keep me sane and aware that there are other important things in life. Here is a list with no particular order: Joana Cruz, Lara, Marta Perfeita, MaC, Thomas, Rui, Sara Parreira, Teresa Matos, Tola, Inês Ulrica, Fábio Santos, Rita Ferreira and Filipa Nunes.

Over the past two years, a new family came into my life: to Paulo's parents, Florbela and Humberto, his brothers and their wives, Miguel and Cárita, and Pedro and Susana, and his nieces, Carolina and Beatriz, a very sincere thank you for everything!

The greatest acknowledge of all is to my parents. It was their love, their teachings and their values that made it all possible. Their words encourage me everyday to move forward and smile even in the worst moments.

At last, I wish to thank Paulo. I became speechless when time is come to describe how his support and vision of the world, so different from my own, helped me to go further more at all levels. Thank you for turning my life upside down, for constantly reminding me of my priorities and for helping me to complete this work.

Resumo

Os estudos da regulação da expressão génica têm revelado elevada complexidade e diversidade de processos responsáveis por uma correta definição das características dos organismos e por um aumento de versatilidade e adaptação dos mesmos. Apesar da regulação transcricional ter sido realçada devido à sua importância para o controlo da expressão génica, a regulação pós-transcricional tem demonstrado ser capaz de contribuir para este controlo com uma multiplicidade de mecanismos que permitem uma modulação da expressão de uma forma mais rápida e versátil (Mata et al., 2005; Mignone et al., 2002; Sonenberg and Hinnebusch, 2009).

Pequenas grelhas de leitura a montante da grelha de leitura principal (uORFs – *upstream open reading frames*) são um exemplo de elementos que atuam em *cis,* envolvidos na regulação pós-transcricional. As uORFs encontram-se na região 5' líder do transcrito, parecem estar envolvidas na inibição da tradução da ORF (ORF - *open reading frame*) principal, e estão presentes principalmente em proto-oncogenes, e em genes envolvidos no crescimento e diferenciação celular (Kozak, 1987; Morris, 1995; Morris and Geballe, 2000; Spriggs et al., 2010). Os últimos estudos apontam para que cerca de 49% do transcritoma humano contenha uORFs (Calvo et al., 2009). Se o codão de iniciação (AUG) da uORF for reconhecido pela maquinaria de tradução é evidente o constrangimento que esta causa ao reconhecimento e tradução da grelha de leitura mais a jusante, funcionando assim como um regulador negativo da expressão génica.

Na presente tese, o objectivo foi estudar o funcionamento, os mecanismos associados e a relevância biológica da uORF presente no transcrito da eritropoietina humana (EPO).

EPO é uma hormona glicoproteica envolvida na estimulação da eritropoiese, i.e., na produção de eritrócitos e na sobrevivência dos seus precursores. A EPO é uma proteína constituída por 193 aminoácidos, com 30,4 kDa (Bunn, 1990; Krantz, 1991), codificados por um transcrito de 1340 nucleótidos, cuja região 5' líder é composta por 181 nucleótidos, onde está localizada uma uORF de 14 codões.

Ao longo do estudo da EPO foram-lhe reconhecidas outras funções não--hematopoiéticas, nomeadamente, como resultado das suas atividades de estimulação da proliferação, diferenciação e atividade antiapoptótica, a EPO foi reconhecida como

cardio e neuroprotetora (Digicaylioglu and Lipton, 2001; Gassmann and Soliz, 2009; Maiese et al., 2008).

De forma análoga ao que aconteceu com as suas funções, também o reconhecimento dos tecidos onde é produzida foi alargado. Inicialmente foi atribuída a sua produção e excreção ao fígado, na vida fetal, e ao rim, na vida adulta (Dame et al., 1998; Paliege et al., 2010). No entanto, o seu mRNA é expresso numa multiplicidade de outros órgãos tais como: células cerebrais, coração ou pulmões (Dame et al., 2001; Fandrey and Bunn, 1993; Ghezzi and Brines, 2004; Hoch et al., 2011).

Sendo assim, a EPO é uma proteína multifacetada e fundamental para uma diversidade de processos biológicos, o que revela a necessidade de uma regulação fina da expressão desta proteína. De facto, são vários os mecanismos responsáveis pela correta e coordenada produção da EPO. Um dos mais bem e frequentemente estudados é o aumento da transcrição da EPO como resposta à hipóxia. Neste processo está envolvido um factor de transcrição induzido pela hipóxia (HIF – *hypoxia inducible factor*).

Na presente tese demonstramos como a regulação pela uORF da EPO funciona como mais um nível desta já complexa estrutura de controlo de expressão da EPO. Os nossos resultados demonstram que a uORF da EPO é extremamente conservada ao longo da evolução. A sua conservação é observada na presença, tamanho, região intercistrónica, sequência nucleotídica e sequência peptídica, o que nos indica que haja uma funcionalidade associada à sua existência. De facto, os resultados obtidos revelam que a esta uORF é funcional, sendo reconhecida pela maquinaria de tradução em todas as linhas celulares humanas estudadas: linhas celulares de rim fetal (HEK293), hepatócitos de adulto (HepG2) e rim adulto (REPC), que foram selecionadas precisamente por corresponderem aos locais com maior produção e secreção da EPO.

Para além da preservação da sua função em todos os tecidos analisados verificamos também a manutenção dos vários mecanismos associados à função da uORF da EPO. Mais especificamente, os nossos resultados demostram que tanto o *leaky scanning* no AUG da uORF da EPO, como a reiniciação da tradução estão envolvidos no reconhecimento e expressão da ORF principal. Adicionalmente, esta uORF funciona de uma forma independente do péptido que codifica, não promovendo o bloqueio da maquinaria de tradução nem, devido ao seu pequeno tamanho, sendo capaz de induzir a rápida degradação do respetivo transcrito (NMD – *nonsense-mediated mRNA decay*).

Em seguida demonstrámos que a região a 3' não traduzida (3'UTR – *3' untranslated region*), descrita como envolvida no controlo da estabilidade do transcrito (McGary et al., 1997; Rondon et al., 1991), é responsável pelo aumento da expressão da ORF principal nas três linhas celulares em estudo. No entanto, apenas na linha celular REPC, este facto corresponde a um aumento dos níveis de mRNA, mantendo-se estes inalterados nas células HEK293 e HepG2. Adicionalmente, demonstrámos que esta região tem uma função independente da uORF da EPO, mantendo esta última um impacto negativo na expressão da ORF principal, mesmo na presença da 3'UTR. Estes mecanismos verificaram-se em todas as linhas celulares em estudo revelando uma manutenção do funcionamento da uORF em todos os tecidos em que há expressão.

Tendo em conta que os exemplos de uORFs descritos demonstram a sua capacidade de alterar a sua repressão em resposta a condições de stresse (Chen et al., 2010; Mouton-Liger et al., 2012; Pentecost et al., 2005), decidimos verificar se a uORF da EPO tem essa função. Para tal induzimos nas células HEK293, HepG2 e REPC hipóxia química e privação de nutrientes e verificámos que apenas nas células REPC, sob efeito de hipóxia, a uORF é de facto menos repressiva, permitindo um aumento de expressão da ORF principal. Deste modo, verificámos que a regulação da expressão mediada pela uORF da EPO é específica de tecido e de estímulo.

Na tentativa de perceber qual o mecanismo subjacente verificámos que, apesar da complexa estrutura secundária da região 5' líder do transcrito da EPO, esta não apresentava sequências internas de entrada do ribossoma (IRES – *internal ribosome entry sites*) em condições normais, nem em condições de hipóxia, não sendo este processo o responsável pela diminuição do impacto negativo da uORF. No entanto, demonstrámos que ocorre uma maior percentagem de *leaky scanning* nestas condições, ou seja, que o AUG da uORF da EPO está a ser menos reconhecido e que este efeito está diretamente relacionado com a fosforilação do factor de iniciação eucariótico (eIF – eucaryotic initiation factor) 2α, tal como já foi descrito anteriormente para outras uORFs (Palam et al., 2011; Zhou et al., 2008a). Esta resposta da uORF da EPO está relacionada com a regulação da sua expressão em condições de hipóxia no rim e com as suas funções hematopoiéticas, apresentando-se, deste modo, como um novo mecanismo de regulação para além dos já descritos.

Como foi referido anteriormente, a EPO é uma proteína multifacetada com um elevado potencial neuroprotetor e cuja expressão foi observada em células cerebrais. Tendo isto em consideração, decidimos estudar o efeito da uORF da EPO numa linha celular de fibroblastos do cérebro (SW1088). O nosso primeiro objetivo foi verificar se a uORF da EPO mantém a sua funcionalidade também nesta linha celular e se os mecanismos de ação são preservados. Os nossos resultados evidenciam que a uORF é funcional, inibindo a tradução da ORF principal na mesma ordem de grandeza observada nas linhas celulares anteriormente referidas. Adicionalmente, verificámos que, também nesta linha celular, tanto o mecanismo de *leaky scanning* no AUG da uORF como a reiniciação da tradução são responsáveis pela tradução da ORF principal. Concomitantemente, a uORF funciona de forma independente da sequência peptídica, e tem um efeito independente da presença da 3'UTR que, tal como nas linhas celulares HEK239 e HepG2, é capaz de aumentar os níveis de proteína. Consequentemente, tal levou-nos a estudar a resposta da uORF a situações de stresse. Para tal, induzimos isquemia química nas células SW1088. Os resultados foram surpreendentes visto que a capacidade de tradução da ORF principal aumentou grandemente quando as células foram expostas ao estímulo, apontando para um alívio do efeito repressor da uORF. No entanto, este efeito resultou da diminuição dos níveis de mRNA mantendo-se os níveis de proteína inalterados.

Adicionalmente, fomos estudar as características e fatores envolvidos na reiniciação da tradução após a leitura da uORF da EPO. Na presente dissertação, demonstrámos que o tamanho da uORF determina a capacidade de reiniciação, tal como era esperado. Verificámos ainda que a depleção das subunidades *h*, *f* e *e* do complexo eIF3 diminui a capacidade de reiniciação, mas o mesmo não se verifica com a depleção das subunidades *a* e *c* do eIF3. Assim, é possível concluir que o complexo proteico que constitui o eIF3 está diretamente implicado na eficiência de reiniciação através de subunidades específicas.

Em conclusão, o trabalho desenvolvido na presente dissertação demonstrou a existência de um novo mecanismo de regulação da expressão da EPO, dissecou os mecanismos dessa regulação da tradução e revelou a sua implicação na resposta a diferentes condições de stresse, indicando a sua relevância biológica. Os nossos resultados contribuíram também para elucidar a base molecular adjacente ao mecanismo de reiniciação.

Palavras-chave

Expressão génica; tradução; controlo traducional; grelha de leitura a montante da grelha de leitura principal (*upstream open reading frame* – uORF); eritropoietina (EPO); factor de iniciação eucariótico 2α (eukaryotic initiation factor 2α – eIF2α); factor de iniciação eucariótico 3 (eukaryotic initiation factor 3 – eIF3)

Abstract

Functional upstream open reading frames (uORFs) are *cis*-acting regulatory elements of gene expression that repress translation of the main ORF in normal conditions. Under stress conditions, they are able to alleviate their repressive effect as a response to the environmental change. Also, they are evolutionarily conserved and are present in about 49% of the human transcriptome.

Human erythropoietin (EPO) is a hormone largely known for its hematopoietic and non-hematopoietic activities, such as cardio and neuroprotection. EPO is produced mainly in fetal liver, and in the adult kidney, but also in several other organs, such as the brain. EPO gene expression is highly regulated at many levels and in response to stress conditions, being the activation of EPO transcription in response to hypoxia one of the best studied parameters. Here, we report that EPO expression is also regulated by a 14-codon uORF within the 5' leader sequence of the transcript. Indeed, we show that *EPO* uORF represses translation of the main ORF in the cell lines derived from organs known to be the major sites of production for this protein: embryonic kidney - HEK293, adult liver - HepG2, and adult kidney - REPC cells. Although both leaky scanning and translation reinitation are responsible for the low levels of EPO AUG recognition under normal conditions, in REPC cells under hypoxia the uAUG is less recognized, which accounts for an increase in the expression of the main ORF. Furthermore, we show that this derepression is related to the phosphorylation of eukaryotic initiation factor 2α (eIF2α) that occurs during hypoxia. In addition, we proved that *EPO* uORF is functional in neuronal cells (cell line SW1088) and that the mechanisms related to the uORF repression are also preserved. However, during chemical ischemia, EPO synthesis is increased. Surprisingly, the mRNA levels are decreased, indicating a distinct regulation mechanism from the one observed in response to hypoxia in REPC cells.

Trying to extend the current knowledge about the mechanistic basis of reinitiation, and using the *EPO* uORF as experimental model, we further shown that the uORF length controls reinitiation. In addition, we demonstrated that the reinitiation event is dependent on eIF3h, f and e subunits, but independent of eIF3a and c subunits.

Together, these findings provide a thorough characterization of the mechanisms involved in *EPO* uORF activity and uncover the importance of this element in the regulation of EPO expression under stress conditions both in renal and neuronal cells.

Keywords

Gene expression; translation; translational control; upstream open reading frame (uORF); human erythropoietin (EPO); eukaryotic initiation factor (eIF) 2α; eIF3

Abbreviations

4E-BP	eukaryotic translation initiation factor 4E-binding protein
A	adenosine
AD	Alzheimer's disease
AdoMetDC	S-adenosylmethionine decarboxylase
AIDS	acquired immunodeficiency syndrome
ARNT	aryl hydrocarbon receptor nuclear translocator
ATF4	activating transcription factor 4
ATP	adenosine triphosphate
A-site	aminoacyl-site
BACE1	β-site amyloid precursor protein-cleaving enzyme 1
bp	base pairs
C	cytidine
CAT1	cationic amino acid transporter 1
CDDO	2-cyano-3,12-dioxooleana-1,9-dien-28-oic acid
cDNA	mRNA-complementary DNA
CFTR	cystic fibrosis transmembrane conductance regulator
CHOP	CCAAT/enhancer-binding protein homologous protein
CI	chemical ischemia
CITED2	Cbp/p300-interacting transactivator with Glu/Asp-rich carboxy-terminal domain 2
CPT1C	carnitine palmitoyltransferase 1C
CREB	cAMP-response element binding protein
C/EBP	CCAAT/enhancer binding protein
C-terminal	carboxyl-terminal
dFBS	dialyzed fetal bovine serum
DMEM	Dulbecco's modified Eagle medium
DMSO	dimethyl sulfoxide
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
dNTP	deoxynucleoside triphosphate
DRD3	human dopamine D3 receptor
eEF	eukaryotic elongation factor
eIF	eukaryotic translation initiation factor
EJC	exon junction complex
EPO	erythropoietin
EPOR	erythropoietin receptor
ER	endoplasmic reticulum
ERBP	EPO RNA binding protein
eRF	eukaryotic translation release factor

ERK	extracellular signal-regulated kinase
E-site	exit-site
FLuc	firefly luciferase
FXII	human clotting factor XII
G	guanosine
GADD34	growth arrest DNA-inducible gene 34
GEF	guanine nucleotide exchange factor
GCH1	guanosine triphosphate cyclohydrolase 1
GCN2	general control non-derepressible-2 kinase
GDP	guanosine diphosphate
GPS1	G protein pathway suppressor 1
GRB2	growth factor receptor bound protein 2
GTP	guanosine triphosphate
HDAC1	histone deacetylase 1
HAMP	hepcidin
HIF	hypoxia inducible factor
HR	human hairless homolog
HRE	hypoxia responsive element
HRI	heme-regulated inhibitor kinase
hsp70	heat-shock protein 70
IFRD1	interferon-related development regulator 1
Ig	immunoglobulin
IRES	internal ribosome entry site
JAK	Janus kinase
KCNJ11	potassium inwardly-rectifying channel, subfamily J, member 11
KIE	kidney inducible element
LDLR	low-density lipoprotein receptor gene
Luc	luciferase
m7G	7-methylguanosine
MAPK	mitogen-activated protein kinases
MEK	MAPK/ERK kinases
Met	methionine
Met-tRNAi	methionine-loaded initiator tRNA
mRNA	messenger ribonucleic acid
mRNP	messenger ribonucleoprotein particle
mTOR	mammalian target of rapamycin
NDST	N-deacetylase/N-sulfotransferase
NK-kB	nuclear factor-kappa B
NMD	nonsense-mediated mRNA decay
nt	nucleotide

N-terminal	amino-terminus
ORF	open reading frame
PABP	poly(A)-binding protein
PABPC1	poly(A)-binding protein cytoplasmic 1
PAGE	polyacrilamide gel electrophoresis
PCBP	poly(C)-binding protein
PCR	polymerase chain reaction
PERK	PKR-like endoplasmic reticulum kinase
PEX7	peroxisomal biogenesis factor 7
PI3K	phosphatidylinositol-3 kinase
PKB	protein kinase B
PKC	protein kinase C
PKR	double-stranded RNA-activated kinase
Poly(A)	poly-adenilate
POMC	proopiomelanocortin
Pre-mRNA	messenger ribonucleic acid precursor
PTPRJ	receptor-like protein-tyrosine phosphatase J
P-site	peptidyl-site
PTC	premature translation termination codon
PVDF	polyvinylidene difluoride
rhEPO	recombinant human EPO
RLuc	*Renilla* luciferase
RNA	ribonucleic acid
RNAi	RNA interference
RNase	ribonuclease
RPMI	Roswell Park Memorial Institute
RT	reverse transcription
RT-qPCR	reverse transcription-quantitative PCR
SDS	sodium dodecyl sulphate
SH2	Src homology-2
SHC	Src homology-2 domain containing transforming protein
siRNA	short interfering RNA
SMG5	suppressor of morphological defects on genitalia 5
SOS	son of sevenlees
STAT	signal transducer and activator of transcription
T	thimidine
TGFβ3	transforming growth factor-β3
TIE2	endothelial cell tyrosine kinase receptor
Tp	thapsigargin
TPO	thrombopoietin
TRB3	tribbles homolog 3

tRNA	transfer ribonucleic acid
U	uridine
uAUG	upstream AUG codon
uORF	upstream open reading frame
UTR	untranslated region
VEGF-A	vascular endothelial growth factor A
VHL	von Hippel-Lindau tumor supressor
WT	wild type

Table of contents

Prefácio v

Acknowledgments ix

Resumo xi

Palavras-chave xv

Abstract xvi

Keywords xvii

Abbreviations xviii

Table of contents xxii

CHAPTER I – General Introduction 27

I.1. mRNA translation: mechanisms and control 28

I.1.1. Translation initiation 29

I.1.2. Translation elongation 30

I.1.3. Translation termination and recycling 31

I.1.4. Mechanisms of mRNA translational control 32

I.1.4.1. Global control of protein synthesis 32

I.1.4.2. Specific control of protein synthesis 33

I.2. Upstream open reading frames (uORFs) 34

I.2.1. uORFs as translational regulatory elements 35

I.2.2. uORFs and mRNA decay 40

I.2.2.1. Nonsense-mediated mRNA decay (NMD) 40

I.2.2.2. Example of uORFs that trigger NMD 42

I.2.3. uORFs and the eIF3 complex 43

I.2.4. uORFs and the cellular response to stress conditions 45

I.2.5. uORFs and human disease 52

I.3. Human Erythropoietin (EPO) 63

I.3.1. EPO signaling pathways 64

I.3.2. Transcriptional regulation of the *EPO* gene 67

I.3.3. Post-transcriptional regulation of the *EPO* transcript 69

I.3.4. EPO as a therapeutic target 70

I.4. Aims 72

CHAPTER II – Translation of the human erythropoietin transcript is regulated by an upstream open reading frame in response to hypoxia 75

Author's note 76

II.1. Abstract 77

II.2. Introduction 77

II.3. Materials and Methods 82

II.3.1. Plasmid constructs 82

II.3.2. Cell culture and plasmid transfection 84

II.3.3. siRNA transfection 85

II.3.4. SDS-PAGE and Western blotting 85

II.3.5. Luminometry assay 85

II.3.6. RNA isolation 86

II.3.7. Reverse transcription-quantitative PCR (RT-qPCR) 86

II.3.8. Statistical analysis 86

II.4. Results 88

II.4.1. The human *EPO* 5' leader sequence comprises a conserved uORF 88

II.4.2. The *EPO* uORF represses translation of a downstream main ORF 89

II.4.3. Both translation reinitiation and uAUG leaky scanning are involved in the translational initiation at the main AUG codon 92

II.4.5. Translational repression exerted by the *EPO* uORF is peptide sequence-independent 94

II.4.5. The 3'UTR of the *EPO* mRNA has no impact on the inhibitory effect of the uORF 95

II.4.6. The *EPO* uORF does not trigger nonsense-mediated mRNA decay 98

II.4.7. EPO is regulated at the translational level in response to hypoxia, but not to nutrient deprivation, specifically in renal cells 99

II.4.8. *EPO* translational derepression in response to hypoxia in REPC cells is not mediated by an internal ribosome entry site 102

II.4.9. *EPO* translational derepression in response to hypoxia is mediated by leaky scanning of ribosomes through the inhibitory uORF 105

II.4.10. Hypoxia-induced phosphorylation of eIF2α is required for *EPO* translational regulation 107

II.5. Discussion 109

II.6. Acknowledgements 116

CHAPTER III – The role of the erythropoietin upstream open reading frame in the human neuronal tissue 117
Author's note 118
III.1. Abstract 119
III.2. Introduction 119
III.3. Materials and Methods 122
III.3.1. Plasmid constructs 122
III.3.2. Cell culture and plasmid transfection 122
III.3.3. Luminometry assay 123
III.3.4. RNA isolation 123
III.3.5. Reverse transcription-quantitative PCR (RT-qPCR) 123
III.3.6. Statistical analysis 124
III.4. Results 124
III.4.1. *EPO* uORF represses translation in neuronal cells 124
III.4.2. The mechanism by which the main ORF is recognized is maintained in liver, kidney and neuronal cells 126
III.4.3. In neuronal cells, the translational machinery is not blocked by the *EPO* uORF-encoded peptide 128
III.4.4. In neuronal cells, *EPO* 3'UTR has no impact on the inhibitory effect of the uORF 129
III.4.5. The repressive effect of the *EPO* uORF is inhibited during chemical ischemia 132
III.5. Discussion 133
III.6. Acknowledgements 136
CHAPTER IV – The translation reinitiation mechanism of the human erythropoietin transcript 137
Author's note 138
IV.1. Abstract 139
IV.2. Introduction 139
IV.3. Materials and Methods 142
IV.3.1. Plasmid constructs 142
IV.3.2. Cell culture, plasmid and siRNA transfection 143
IV.3.3. RNA isolation 143
IV.3.4. Semi-quantitative RT-PCR 143
IV.3.5. Dual luciferase assay 144

IV.3.8. SDS-PAGE and Western blotting 144
IV.4. Results 145
IV.4.1. The size of *EPO* uORF influences translation reinitiation efficiency 145
IV.4.2. eIF3h, f and e affect the efficiency of translation reinitiation 147
IV.4.3. eIF3a and c do not affect the efficiency of translation reinitiation 149
IV.5. Discussion 150
IV.6. Acknowledgements 153
CHAPTER V – General Discussion 155
V.1. General Discussion and Future Perspectives 156
CHAPTER VI – References 161

Figures

Figure I.1. The canonical translation initiation process. 30
Figure I.2. Mechanisms of uORF-mediated translational control. 37
Figure I.3. Features that modulate the uORF impact. 39
Figure I.4. A model for NMD-resistance of AUG-proximal nonsense-mutated mRNAs. 41
Figure I.5. uORFs response to stress conditions. 47
Figure I.6. EPO signalling pathways. 65
Figure II.1. The 5' leader sequence of the *EPO* transcript includes a highly conserved uORF. 88
Figure II.2. The *EPO* uORF represses translation of the downstream main ORF. 90
Figure II.3. Both translation reinitiation and uAUG leaky scanning are involved in the translational initiation at the main AUG codon. 93
Figure II.4. Translational repression exerted by the *EPO* uORF is peptide sequence-independent. 95
Figure II.5. The 3'UTR of the *EPO* mRNA enhances the inhibitory effect of the uORF in REPC cells. 97
Figure II.6. The human *EPO* transcript is resistant to nonsense-mediated mRNA decay. 99
Figure II.7. The *EPO* uORF responds to hypoxia but not to nutrient starvation, specifically in REPC cells. 101

Figure II.8. EPO translational derepression in response to hypoxia in REPC cells is not mediated by an internal ribosome entry site (IRES). 104
Figure II.9. *EPO* translational derepression in response to hypoxia of REPC cells is mediated by leaky scanning of ribosomes through the inhibitory uORF. 106
Figure II.10. Hypoxia induces phosphorylation of eIF2α, which is required for *EPO* translational regulation in REPC cells. 108
Figure III.1. The *EPO* uORF represses translation of the downstream main ORF in neuronal cells. 125
Figure III.2. Both translation reinitiation and uAUG leaky scanning are involved in the translational initiation at the main AUG codon. 127
Figure III.3. In neuronal cells, the translational repression exerted by the *EPO* uORF is peptide sequence-independent. 129
Figure III.4. In neuronal cells, the 3'UTR of the *EPO* mRNA has no influence in the inhibitory effect of the uORF. 131
Figure III.5. EPO relative protein levels are enhanced in SW1088 cells in response to chemical ischemia. 133
Figure IV.1. The size of the uORF influences the translation reinitiation efficiency. 146
Figure IV.2. Depletion of eIF3h, f, and e, alters the reinitiation efficiency after the translation of the 14-codon uORF. 148
Figure IV.3. Depletion of eIF3a and c does not affect the reinitiation efficiency. 150

Tables

Table I.1. Examples of human genes encoding mRNAs that, under stress conditions, evade global repression of translation and are upregulated due to the presence of uORFs 50
Table I.2. Examples of human diseases associated with polymorphisms or mutations that introduce/eliminate uORFs or modify the encoded uORF peptide 59
Table II.1. DNA oligonucleotides used in the current work. 87
Table IV.1. Sequences of the siRNAs used in the current work. 143
Table IV.2. DNA oligonucleotides used in the current work. 144

CHAPTER I – General Introduction

I.1. mRNA translation: mechanisms and control

Eukaryotic gene expression is a complex sequence of biochemical processes cells use to produce specific gene products, either RNAs or proteins.

The messenger RNA (mRNA) precursor, originated in the nucleus from the DNA, undergoes splicing, 5' capping, 3' polyadenylation and in some cases RNA editing, generating the mature mRNA. Then, the mature mRNA is transported into the cytoplasm where it is translated, stored or even degraded. In the course of these events, individual transcripts associate with particular proteins forming messenger ribonucleoprotein particles (mRNPs), which are able to dictate the fate of the transcript (Fasken and Corbett, 2005). The formed mRNPs can influence the fate of the transcript by altering its cellular localization, translation and decay, in response to a network of cellular signals (Moore, 2005).

It is essential to tight regulate all the events taking part in this intricate process in order to ensure the quality and fidelity of gene expression, thus allowing homeostasis of the organisms.

Most studies done during the second half of the twentieth century put their emphasis on the recognition of several regulatory mechanisms at transcriptional level. These mechanisms operate at the earliest point of gene expression and are able to modulate the downstream outcome of mRNA synthesis and therefore protein expression. However, there are an increasing number of studies in post-transcriptional control mechanisms that illustrate how regulation of gene expression at this level presents more rapid and reversible responses, allowing cells to adapt to changes in the surrounding environment by altering the patterns of gene expression. The mRNA quality control is ensured by a number of surveillance mechanisms that act at different steps of mRNA biogenesis, and, in particularly, at the translation stage (Gebauer and Hentze, 2004; Silva and Romão, 2009; Sonenberg and Hinnebusch, 2009).

Translation is a complex, fine tuned process that can be divided into four stages – initiation, elongation, termination and ribosome recycling – each of which requiring a particular set of conditions and factors.

I.1.1. Translation initiation

Translation initiation is the rate-limiting step and, in eukaryotic cells, requires the participation of several eukaryotic initiation factors (eIFs) (Figure I.1.) (Livingstone et al., 2010). Canonical translation initiation is mediated by the recruitment of the cap-binding protein complex, eukaryotic initiation factor 4F (eIF4F), which comprises eIF4E, eIF4G and eIF4A, to the mRNA 5' end (Sonenberg and Hinnebusch, 2009). eIF4E is the factor that recognizes the m^7G cap. eIF4G has a binding site for eIF4E and the poly(A)-binding protein (PABP), which in turn is bound to the poly(A) tail, resulting in mRNA circularization (Morino et al., 2000; Sonenberg and Hinnebusch, 2009). The unwinding of the 5' leader sequence by the ATP dependent helicase eIF4A, enables binding of the 40S ribosomal subunit (Gebauer and Hentze, 2004). The association of eIF1, eIF1A and eIF3 to the 40S subunit facilitates the binding of the ternary complex eIF2-GTP-Met-$tRNA_i^{Met}$ (Sonenberg and Hinnebusch, 2009). The resulting 43S preinitiation complex can land next to the cap and scans in a 5' to 3' direction until it recognizes an AUG codon base-pairing with Met-$tRNA_i^{Met}$ (Kozak, 1999; Sonenberg and Hinnebusch, 2009). eIF3 is also involved in recruiting the 43S preinitiation complex to the mRNA and interacts with eIF4G, at least in mammals (Hinnebusch, 2006). Upon recognition of the start codon, eIF5 stimulates GTP hydrolysis, resulting in the release of eIF2-GDP and probably other 40S-bound initiation factors. eIF1 allows scanning 43S complexes to discriminate against codon-anticodon mismatches and preventing premature eIF5-induced hydrolysis of eIF2-GTP and Pi release (Holcik and Pestova, 2007). eIF1A also regulates start codon selection promoting continued scanning at non-AUG codons or by arresting scanning and promoting eIF1 release at AUG codons (Sonenberg and Hinnebusch, 2009). After the release of eIF2-GDP and other eIFs, eIF5B catalyzes the recruitment of the 60S subunit to form an 80S ribosome, and elongation can start (Gebauer and Hentze, 2004; Pisarev et al., 2007; Sonenberg and Hinnebusch, 2009). Since eIF3 binds mainly to the solvent side of the 40S subunit, its dissociation is not essential for subunit joining and may be delayed (Szamecz et al., 2008; Valásek et al., 2002).

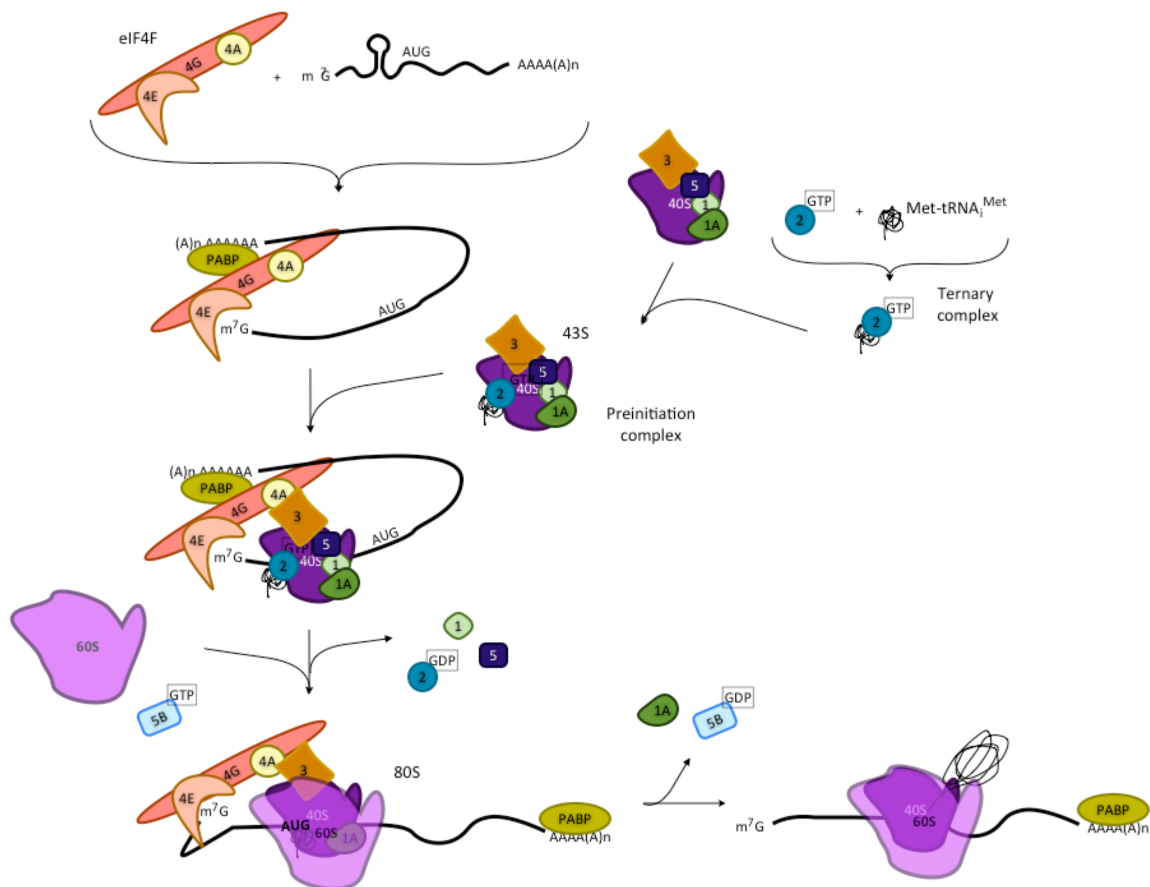


Figure I.1. The canonical translation initiation process.
The eIF4F, that comprises eIF4E, 4A and 4G, is recruited to the mRNA 5' end. This complex interacts with PABP, through eIF4G, presumably circularizing the mRNA. The association of eIF1, 1A, 3 and 5 to the 40S subunit facilitates the binding of the ternary complex, comprising eIF2-GTP-Met-tRNA$_i^{Met}$. The resulting 43S preinitiation complex can land next to the cap and scans the mRNA in a 5' to 3' direction. After recognition of the AUG initiation codon, eIF1 is displaced and eIF5 mediates the hydrolysis of eIF2-bound GTP. Joining of the 60S ribosomal subunit will cause the displacement eIF2-GDP and other eIFs mediated by eIF5B and the assembly of 80S elongation-competent ribosomes induces the release of eIF1A and eIF5B [adapted from (Sonenberg and Hinnebusch, 2009)].

I.1.2. Translation elongation

The elongation stage is the sequential addition of amino acids organized according to the nucleotide sequence, to the growing polypeptide chain (Abbott and Proud, 2004). The ribosome presents three tRNA-binding sites: the A- (aminoacyl) site, that receives the incoming aminoacyl-tRNA for the newly encountered mRNA codon, the P- (peptidyl)

site, which holds the tRNA with the nascent peptide chain and the E- (exit) site that retains the deacylated tRNA prior to its release (Proud, 1994). The eukaryotic elongation factor (eEF) 1A is a key factor that recruits the tRNAs to the ribosomal A-site, upon hydrolysis of GTP. The regeneration of active eEF1A-GTP complexes is mediated by eEF1B.

The correct codon-anticodon base pairing between the mRNA and the tRNA is needed so that the tRNA enters the next stage of elongation, which involves a conformational change of the large ribosomal subunit and the GTP hydrolysis so that the aminoacyl tRNA enter the A-site. Then, the translocation of peptidyl-tRNA from A- to P- and of deacylated tRNA from P- to E- sites is then promoted by eEF2 in a GTP dependent manner. Also there is a movement of the ribosome relative to the mRNA by exactly three nucleotides, which places the next codon in A-site allowing the addition of the next amino acid to the growing protein chain (Abbott and Proud, 2004; Kapp and Lorsch, 2004; Pisarev et al., 2007).

I.1.3. Translation termination and recycling

Termination of translation occurs when the ribosomal A-site reaches one of the three possible termination codons (UAA, UAG or UGA). The eukaryotic translation release factor (eRF) 1 determines the termination of translation by inducing the hydrolysis of the ester bond of the P-site peptidyl-tRNA, releasing the new polypeptide (Frolova et al., 2000; Kisselev et al., 2003). Other factor involved in the termination event is eRF3. This factor has GTPase activity and interacts with eRF1 forming a stable complex. Its function is to ensure a fast and efficient hydrolysis of the peptidyl-tRNA by eRF1 (Alkalaeva et al., 2006). After the polypeptide release the ribosomes in post-termination complexes have to be dissociate. How this happens for eRF1 and 3 is still unknown. However, it has been proposed a role for the eIFs 3, 1 and 1A in the dissociation of these complexes into the 60S ribosomal subunit, mRNA, tRNA and 40S subunit associated with the eIFs mentioned, that can be recycled for other translational events (Pisarev et al., 2007).

In particular, the recycled 40S subunit can undergo new rounds of initiation on the same mRNA. This is possible due to the circularization of the mRNA meditated by the interaction between eIF4G and PABP, as mentioned before. Also, in this stage the eIF3

has a major role since it interacts with the eIF4G and is associated with the recycled 40S subunit (Hinnebusch, 2006; LeFebvre et al., 2006; Pisarev et al., 2007). In this matter, some post-termination events, such as reinitiation and mRNA decay pathways, can be influenced by ribosomal recycling and mainly by eIF3.

I.1.4. Mechanisms of mRNA translational control

Post-transcriptional regulation of gene expression is extremely important to establish the cellular levels of proteins, since they may not correlate to the corresponding mRNA levels. Also, it has been increasingly recognized as a key mechanism by which cells and organisms can rapidly change their gene expression patterns in response to internal or external stimuli. Emerging examples illustrate that expression of all genes is regulated at multiple post-transcriptional steps including mRNA processing, nuclear export and localization, stability, and translation of mature mRNA molecules. Translation itself is regulated by a diverse collection of mechanisms that act mainly at the initiation step, but also during elongation and termination and even after termination.

The translational control can be exerted by two general modes: by global control, that impacts the translation of most mRNAs in the cell and that is exerted mostly at translation initiation; and by mRNA-specific control, where the translation of a specific, or a defined group of mRNAs, is modulated without affecting general protein biosynthesis or the translational status of the cellular transcriptome as a whole.

I.1.4.1. Global control of protein synthesis

The global control of protein synthesis is generally achieved by changes in the initiation stage of translation by altering the phosphorylation state of initiation factors or the regulators that interact with them. One of the most commonly used mechanisms for inhibiting global translation is by phosphorylation of the initiation factor eIF2 (Hinnebusch et al., 2007). In order to be recycled, eIF2 is recharged with GTP by the guanine nucleotide exchange factor (GEF) eIF2B. eIF2 consists of three subunits: α, β and γ. When eIF2 is phosphorylated on serine 51 of its α subunit, it becomes a competitive inhibitor of eIF2B, preventing eIF2 recycling and reducing translation initiation rates by

lowering the ternary complex concentration (Hinnebusch et al., 2007). In mammalian cells, phosphorylation of eIF2α on serine 51 is a major mechanism that regulates initiation of translation in response to various cellular stresses, including virus infection, nutrient deprivation, iron deficiency, and accumulation of unfolded proteins in the endoplasmic reticulum (ER) (Hinnebusch et al., 2007). Depending on the specific cellular stress, eIF2α is phosphorylated by at least 4 different kinases: double-stranded RNA-activated kinase (PKR), which is stimulated by viral infection; general control non-derepressible 2 kinase (GCN2), which is activated by amino-acid starvation; heme-regulated inhibitor kinase (HRI), which is stimulated by heme depletion; and PKR-like ER kinase (PERK), which is activated under circumstances of endoplasmic reticulum (ER) stress. Following stress-induced eIF2α phosphorylation, translation of normal cellular mRNAs is repressed, while the translational initiation of selected mRNAs involved in stress response is stimulated (Hinnebusch et al., 2007).

A second mechanism for nonspecifically reducing levels of protein synthesis can be done by interfering with m^7G cap recognition, thereby preventing recruitment of the translational machinery to the mRNA (Raught and Gingras, 2007). The m^7G cap is recognized by eIF4E as part of the eIF4F complex; however, there are several eIF4E-binding proteins (4E-BPs) which compete with eIF4G for a binding site on eIF4E and prevent eIF4F complex formation (Marcotrigiano et al., 1999). The strength of binding of 4E-BPs to eIF4E is controlled by phosphorylation: hypophosphorylated 4E-BPs bind strongly, while phosphorylated 4E-BPs bind weakly. Phosphorylation of the 4E-BPs is largely controlled by the mammalian target of rapamycin (mTOR) which integrates signals from several upstream signaling pathways (Brunn et al., 1997; Hay and Sonenberg, 2004). mTOR is activated by growth factors and cytokines, phosphorylating 4E-BPs, while under stress or starvation conditions the inactivation of mTOR leads to hypophosphorylated 4E-BPs, which in turn inhibits the overall protein synthesis [for review see (Hay and Sonenberg, 2004)].

I.1.4.2. Specific control of protein synthesis

The control of specific mRNAs is mediated by particular elements usually present in the 5' leader sequence or in the 3' untranslated region (UTR) of the target mRNA. Structural

features and regulatory *cis*-acting elements that determine and modulate the translational efficiency comprise: canonical end modifications of mRNA molecules – the cap structure and the poly(A) tail; upstream AUGs (uAUGs) or upstream open reading frames (uORFs); internal ribosome entry sequences (IRESs); specific binding sites for regulatory protein complexes; specific binding sites for regulatory small microRNAs (miRNAs); and secondary or tertiary RNA structures, such as hairpins and pseudoknots. Many of these features impact negatively the translation of the corresponding mRNA, limiting their translation hence resulting in lower levels of protein. Examples are uORFs, miRNAs and strong secondary RNA structures within the 5' leader sequence of the transcript. On the contrary, there are elements such as the specific binding sites of proteins, or IRES structures, that can induce mRNA translation. In the case of existence of an IRES the 43S preinitiation complex is recruited to an internal region of the mRNA possibly in close proximity to the AUG [for a review see (Mignone et al., 2002)].

Features such as uORFs or long 3'UTRs can also affect the mRNA stability of several physiological transcripts by triggering of nonsense-mediated mRNA decay (NMD) (Amrani et al., 2004; Mendell et al., 2004; Silva et al., 2008).

I.2. Upstream open reading frames (uORFs)

Translational regulation at the initiation step can be mediated *via* different *cis*-acting elements present in the RNA 5' leader sequence of specific transcripts, such as uORFs. uORFs are sequences defined by an initiation codon in-frame with a termination codon located upstream or downstream of the main AUG. uORFs correlate with significantly reduced protein expression levels because they reduce the efficiency of translation initiation of the downstream main ORF in unstressed conditions (Calvo et al., 2009; Morris and Geballe, 2000), or trigger mRNA decay (Mendell et al., 2004; Wittmann et al., 2006; Yepiskoposyan et al., 2011). However, in response to cellular stress, the presence of uORFs can promote the increased expression of certain stress-related mRNAs (Spriggs et al., 2010). Nevertheless, there are other mRNAs for which it has been shown that some or all uORFs have no effect on translation (Lammich et al., 2004; Rogers Jr et al., 2004). Indeed, from the published data, it is apparent that there are different

mechanisms, some of them uORF(s) independent, which can be used by individual uORF-containing mRNAs to control protein synthesis.

Bioinformatic studies have now shown that about 49% of the human transcriptome contains uORFs, which are mostly conserved among species, suggesting evolutionary selection of functional uORFs (Calvo et al., 2009; Iacono et al., 2005; Kochetov et al., 2008; Sathirapongsasuti et al., 2011; Suzuki et al., 2000). uORFs are conspicuously common in certain classes of mRNAs, including two-thirds of oncogenes and many other transcripts that encode proteins involved in important cellular processes, such as differentiation, cell cycle and stress response (Kozak, 1987, 1991; Morris, 1995; Morris and Geballe, 2000; Spriggs et al., 2010). As stated above, it has been suggested that uORFs are negatively correlated with protein production (Calvo et al., 2009; Matsui et al., 2007), but until now, functional activity has been demonstrated for only a limited number of uORFs. Indeed, uORF-mediated translational regulation has been validated experimentally for about 100 eukaryotic transcripts, including around thirty human transcripts (Calvo et al., 2009). In addition, recent studies have described several transcripts where changes in the 5' leader sequence that disrupt or create a uORF are associated with the development of human disease or disease susceptibility, revealing the importance of these *cis*-acting elements in gene expression regulation (Calvo et al., 2009). Bearing in mind the unequivocal examples already described, it is expected that uORF mutations may be involved in the genetic architecture of a wide variety of diseases, including malignancies, metabolic or neurologic disorders, and inherited syndromes.

I.2.1. uORFs as translational regulatory elements

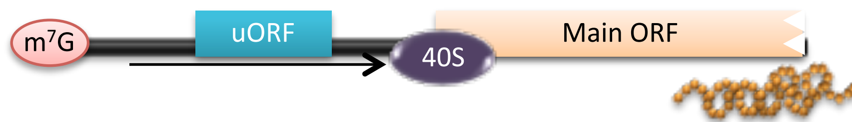
As mentioned before, translation initiation is the rate-limiting step that involves the cooperation of several eIFs in order to recruit the 43S preinitiation complex to the 5' leader region of the mRNA, which in turn, will recognize the AUG and initiate translation elongation (Livingstone et al., 2010).

Initially, it was assumed that the scanning 43S preinitiation complex would generally initiate translation at the first AUG codon encountered. However, several studies have shown that an AUG is not always recognized and there are several factors that can

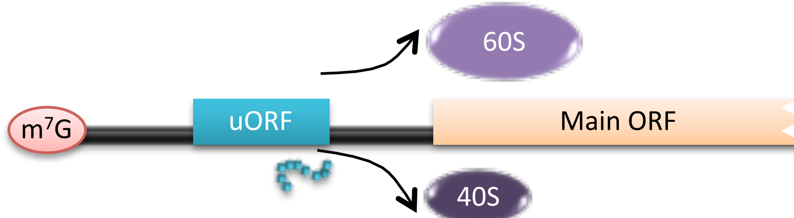
influence this recognition, such as the sequence context of the AUG codon or the presence of strong secondary structures (Sachs and Geballe, 2006). Indeed, it has been demonstrated that there are specific nucleotides surrounding the AUG codon whose presence correlates well with the strength of its recognition. The most efficient context for ribosome recognition and initiation of translation is known as the Kozak consensus sequence (GCCA/GCC**AUG**G). The nucleotides at positions -3 and +4 (underlined) are the most important ones for the definition of the context strength (Kozak, 1986). In the presence of a weaker context sequence, a mechanism called leaky scanning can occur, where the ribosome can either read the AUG codon or pass by it initiating translation at a downstream initiation codon (Kozak, 2002).

For a uORF to function as a translational regulatory element, its initiation codon must be recognized, at least at certain times, by the scanning 40S ribosomal subunit and associated initiation factors. When the uORF recognition is regulated by the so-called leaky-scanning mechanism, ribosomes either scan through the upstream AUG codon (Figure I.2.A) or recognize it, initiating translation. In the case that the uORF is recognized by a scanning ribosome, the following alternative fates are available to the ribosome: (i) translate the uORF and dissociate (Figure I.2.B); (ii) translate the uORF and stall during either the elongation or termination phase of translation, creating a blockage to additional ribosomes (Figure I.2.C) or/and inducing mRNA decay (Figure I.2.D); or (iii) translate the uORF and remain associated with the mRNA, continue scanning, and reinitiate further downstream, at either a proximal or distal AUG codon (Figure I.2.E). Translation reinitiation is thought to be an inefficient mechanism that happens only after translation of a short ORF (Meijer and Thomas, 2002). Indeed, reinitiation is dependent on (i) the time required for the uORF translation, which is determined by the relative length of the uORF and the translation elongation rate; and (ii) the translation initiation factors involved in the translation initiation event (Kozak, 2002; Poyry et al., 2004). Several initiation factors need to remain associated with the ribosome during translation and even after the termination event so that reinitiation can occur (Child et al., 1999; Roy et al., 2010). In this way, a ribosome that translates a shorter uORF (or with a higher translation rate) is more likely to reinitiate translation (Poyry et al., 2004). A key factor for translation reinitiation is the reacquisition of a new ternary complex (eIF2-GTP-Met-$tRNA_i$); this complex is essential for the recognition of a

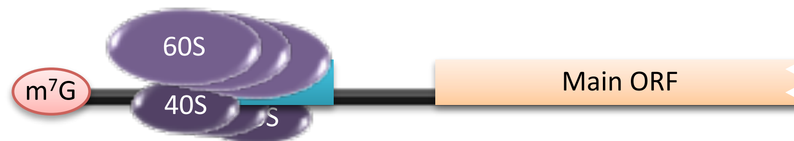
A The uORF is not always translated



B Both ribosomal subunits dissociate after uORF translation



C The uORF represses translation of the main ORF in a peptide-dependent manner



D The uORF termination codon is recognized as premature and nonsense-mediated mRNA decay is triggered

E Translation reinitiation after uORF translation

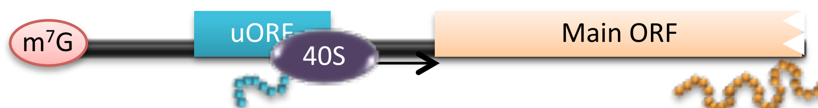


Figure I.2. Mechanisms of uORF-mediated translational control.
(A) The leaky scanning mechanism is dependent on the efficiency of uAUG recognition; sometimes the ribosome can translate the uORF, but other times the scanning machinery bypasses the uAUG, recognizing the downstream AUG and translating the main ORF. **(B)** When a scanning ribosome recognizes and translates a functional uORF, there is synthesis of a small peptide; if translation termination of the uORF is efficient, both 60S and 40S ribosomal subunits might dissociate from the transcript and the main ORF is not translated. **(C)** A uORF can repress translation of the main ORF in a peptide-dependent manner; in this case, the uORF-encoded peptide interacts with the translating machinery and promotes ribosome blockage. **(D)** The termination codon of a uORF can be recognized as premature and nonsense-mediated mRNA decay (NMD) is triggered through a mechanism involving the UPF1 protein and ribonucleases. **(E)** After translation termination of the uORF, the 40S ribosomal subunit can remain associated with the

transcript, resume scanning, and recognize the downstream main AUG – a mechanism designated as translation reinitiation.

downstream AUG by the scanning 40S subunit (Kozak, 2005). In fact, many studies have reported that longer intercistronic regions are more favorable for reinitiation, while for shorter ones the scanning time may not be sufficient for reacquisition of the ternary complex and the downstream AUG will therefore not be recognized (Child et al., 1999; Munzarová et al., 2011; Roy et al., 2010). The basis for the mechanism of translation reinitiation has not been completely elucidated. Therefore, it is essential to define more precisely which initiation factors promote reinitiation competence, as well as potential changes in the ribosomes that may be involved in this process.

As already stated, an additional feature of uORFs is their capacity to block the translational machinery in a peptide dependent manner (Lovett and Rogers, 1996); this might result in the stalling of other ribosomes that access the transcript, thereby dramatically decreasing the translation of the main ORF (Geballe and Morris, 1994). Examples of uORFs that function in a sequence-dependent manner are the receptor-like protein-tyrosine phosphatase J (*PTPRJ*) (Karagyozov et al., 2008), the β2-adrenergic receptor and the S-adenosylmethionine decarboxylase (*AdoMetDC*) (Raney et al., 2002). The few examples described in mammals make it difficult to identify the conserved peptide sequences responsible, and identification of further uORFs with this ability is only possible experimentally. One study comparing full-length cDNA sequences from different plant species aiming to identify conserved peptide uORF sequences found that uORFs rich in serine, threonine and/or tyrosine were present in nine homologous groups (Hayden and Jorgensen, 2007). These amino acids are potential targets for phosphorylation that could possibly promote or inhibit ribosome stalling or translation initiation at downstream ORFs. Nevertheless, further characterization of this type of uORF is necessary before a consensus sequence can be annotated.

Despite the obvious complexity of uORF-mediated translational regulation, results from several studies have revealed that the impact the uORFs can have on translation depends on several variables, such as (i) the distance between the 5' cap and the uORF, (ii) the context in which the uORF AUG is located, (iii) the length of the uORF, (iv) the secondary structure of the uORF, (v) conservation among species, (vi) the number of

uORFs per transcript, (vii) the position of the uORF termination codon, upstream or downstream of the main initiation codon and (viii) the length of the intercistronic sequence(s) (Figure I.3.). Although all types of uORF can reduce protein expression in unstressed cells, four uORF properties are associated with greater translational inhibition. These are: strong uAUG context, evolutionary conservation, increased distance from the cap, and multiple uORFs in the 5' leader sequence (Calvo et al., 2009). These properties reflect the impact that uORF(s) have in translational efficiency of the main ORF, when they are translated.

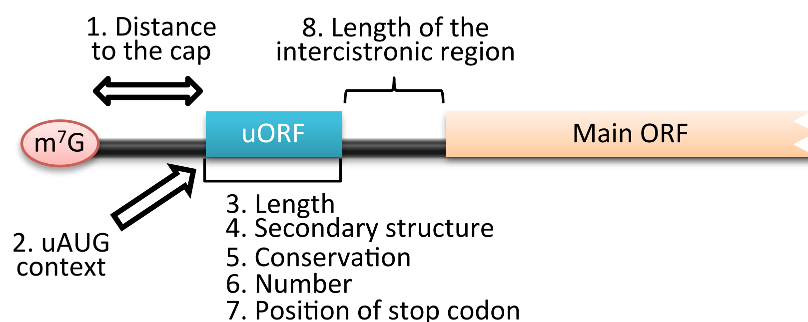


Figure I.3. Features that modulate the uORF impact.
The impact that the uORFs can have on translation depends on (1) distance between the 5' cap (m7G) and the uORF (distance to the cap), (2) context in which the uORF AUG is located (AUG context), (3) length of the uORF, (4) number of uORFs per transcript, (5) secondary structure of the uORF, (6) conservation among species, (7) length of the intercistronic sequence(s), and (8) position of the uORF termination codon, upstream or downstream of the main initiation codon (length, number, secondary structure, conservation, position of stop codon). The increase of translational repression exerted by a uORF correlates with increasing distance between the m7G and the uORF, increasing length of the uORF and intercistronic sequence, a higher number of uORFs, and a stronger uAUG Kozak context.

It is still unclear whether uORF-encoded peptides can play additional roles in the cell. Conceivably, uORF-encoded peptides could act both as translational regulators of the main ORF and as *trans*-acting factors in the cell. Further characterization of conserved uORFs might help to resolve this hypothesis.

I.2.2. uORFs and mRNA decay

I.2.2.1. Nonsense-mediated mRNA decay (NMD)

NMD is one of the better characterized quality control mechanisms which acts as an mRNA surveillance pathway by degrading transcripts harboring premature translation termination codons (PTCs) (Maquat et al., 1981). However, as previously referred, in the last decade, several studies have also implicated NMD in the regulation of steady-state levels of physiological mRNAs, and many examples of natural NMD targets are indeed transcripts containing uORFs (Mendell et al., 2004; Rehwinkel et al., 2006; Wittmann et al., 2006; Yepiskoposyan et al., 2011), in which the uORF termination codon can be recognized as premature.

The major challenge for this translation-dependent mechanism is to discriminate between a premature and a normal termination codon. This discrimination occurs when the ribosome is poised at the termination codon. According to current models, normal translation termination involves the interaction of the eukaryotic release factor 3 (eRF3) with the poly(A) binding protein cytoplasmic 1 (PABPC1) at the terminating ribosome (Figure I.4.), which stimulates a proper and efficient translation termination event (Amrani et al., 2004; Behm-Ansmant and Izaurralde, 2006; Hoshino et al., 1999). However, if the termination codon location within a certain mRNP context does not allow PABPC1 to interact with eRF3, the terminating ribosome will stall, allowing its interaction with the NMD effector UPF1 and NMD triggering (Singh et al., 2008). The "unified model" for NMD proposes that there are several features in the mRNP that can trigger the NMD response. For example, PTCs located at a greater distance from the poly(A) tail, as it is the case for mRNAs harboring long 3'UTRs, can elicit NMD due to PABPC1 failing to interact with the termination complex (Mühlemann, 2008; Shyu et al., 2008; Silva and Romão, 2009; Singh et al., 2008). Another NMD-triggering feature is the presence of at least one exon-exon junction more than 50 nucleotides downstream of the termination codon (Nagy and Maquat, 1998). During splicing, the exon junctions are marked with a dynamic multiprotein complex designated exon-junction complexes (EJC) that associates with the NMD factors UPF2 and UPF3 (Le Hir et al., 2000). The presence of an EJC downstream of a termination codon allows the interplay between UPF1, at the

terminating ribosome, and UPF2 and/or UPF3, which results in UPF1 phosphorylation, irreversibly triggering NMD (Stalder and Mühlemann, 2008). Consequently, PTCs located far, in a linear sense, from the poly(A) tail and associated PABPC1, in mRNAs containing residual downstream EJCs, are expected to elicit NMD (Mühlemann, 2008; Shyu et al., 2008; Silva and Romão, 2009; Singh et al., 2008). Nevertheless, our lab has reported that AUG-proximal nonsense-mutated mRNAs evade NMD (Inácio et al., 2004; Romão et al., 2000; Silva et al., 2006, 2008). In such cases, there is establishment of an efficient translation termination event because of the ability of PABPC1 to travel with the ribosome, due to interactions with eIF4G and eIF3. Our lab has also shown that this allows a repositioning of the PABPC1/eIF4G/eIF3 protein complex in the vicinity of the PTC at the translation termination event, blunting the NMD response and eliciting efficient termination (Peixeiro et al., 2012). Because the PABPC1/eIF4G/eIF3 complex might be still bound to the ribosome when it reaches the stop codon of a small ORF, eIF3 is in a favored position to promote reinitiation competence; as these interactions might be disrupted after some steps of translation elongation, transcripts carrying smaller ORFs are more competent for translation reinitiation than those with larger uORFs.

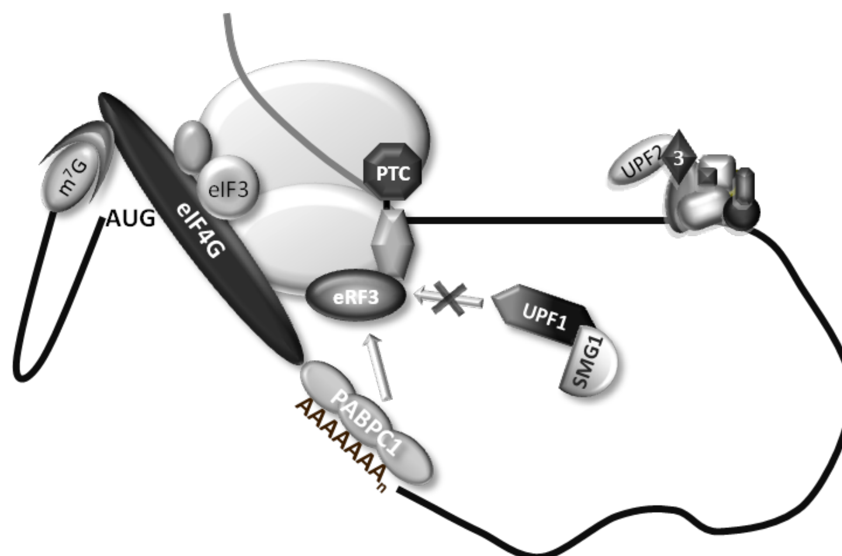


Figure I.4. A model for NMD-resistance of AUG-proximal nonsense-mutated mRNAs.
During cap-mediated translation initiation, PABPC1 interacts with the initiation factor eIF4G. This brings PABPC1 into in the vicinity of the AUG initiation codon where the PABPC1/eIF4G/eIF3 protein complex can form. During the initial phase of translation elongation this complex is repositioned to the vicinity of the PTC allowing PABPC1 to interact with eRF3 at the termination complex, resulting in efficient translation termination and inhibition of NMD (Peixeiro et al., 2012)

I.2.2.2. Example of uORFs that trigger NMD

The termination codon of a uORF can be recognized as a PTC since it is distant from the 3'UTR signals and the corresponding transcript usually presents downstream EJCs located in the coding sequence of the main ORF (McGlincy et al., 2010; Sachs and Geballe, 2006). Examples of human transcripts whose uORFs trigger NMD are the interferon-related developmental regulator 1 (*IFRD1*) (Zhao et al., 2010), the cystic fibrosis transmembrane conductance regulator (*CFTR*) (Davies et al., 2004), and the suppressor of morphological defects on genitalia 5 (*SMG5)* (Yepiskoposyan et al., 2011). However, some naturally occurring uORF containing transcripts escape NMD. Indeed, uORFs often mediate translational repression of the protein coding ORF without an associated decrease in mRNA levels (McGlincy et al., 2010; Sachs and Geballe, 2006). The length of the uORF and the time taken to translate it are characteristics that influence the triggering of NMD (unpublished data from our lab). According to the model established by our lab (Silva and Romão, 2009), only transcripts harboring at least one uORF with a critical length would trigger NMD, while those with smaller uORF(s) could be NMD-resistant because of PABPC1 proximity to the uORF termination codon due to mRNA circularization during translation (Peixeiro et al., 2012; Silva et al., 2008). In mammalian cells, the minimum size of the uORF that triggers NMD has been difficult to determine (Mendell et al., 2004); however, in plants, 35 codons is the threshold (Nyikó et al., 2009): transcripts with longer uORFs are NMD-sensitive and those with shorter uORFs are NMD-resistant. Also, in plants, increasing the reinitiation predisposition has no effect on NMD, which contradicts the notion that reinitiation would prevent the destabilization of the mRNA (Nyikó et al., 2009). Nevertheless, in mammalian cells, some transcripts with long uORFs, which are NMD-targets under normal circumstances, become resistant to NMD during stress conditions, depending on the phosphorylation of eIF2α (Gardner, 2008; Zhao et al., 2010). *IFRD1* is a documented example of a uORF with 52 codons that responds to the phosphorylation of eIF2α by increasing mRNA stability (Zhao et al., 2010). One possible explanation for NMD inhibition in response to eIF2α phosphorylation is that under these conditions, leaky scanning through the uORF increases and thus the corresponding stop codon is not recognized, which impairs NMD. This example illustrates how complex and puzzling the inhibitory effect of a uORF and

the response to stress conditions can be. In any case, these data demonstrate that cells have evolved different mechanisms that contribute to the integrated stress response, among which inhibition of NMD also contributes to increased expression of stress-response proteins.

I.2.3. uORFs and the eIF3 complex

Translation initiation is dependent on several eukaryotic initiation factors (eIFs). These proteins not only ensure the correct recognition of an AUG and the assembly of the translational machinery, but also serve as points of translational control. eIF3 is the largest complex of initiation factors, composed by 13 subunits in mammals, from eIF3a to eIF3m, with a total of 750 kDa (Hinnebusch, 2006). Although many studies have tried to reconstitute the assembly of this complex in mammalian cells, its true composition needs further clarification. In fact, many of the known functions and interactions of these subunits are the result of the studies done in yeast (Herrmannová et al., 2012; Masutani et al., 2013; Valásek, 2012).

In yeast, the eIF3 complex is composed by six subunits. Five of them – eIF3a, b, c, g, and i - are organized in a core and eIF3j is outside the core (Herrmannová et al., 2012). In mammals, *in vitro* studies suggested a functional core comprising a, b, c, e, f and h subunits (Masutani et al., 2007). However, other study based on tandem mass spectrometry and solution disruption assays identified three stable modules, one composed of subunits a, b, i, and g, resembling the yeast eIF3 core, a second one encompassing subunits c, d, e, l, and k, and a third one including subunits f, h, and m (Zhou et al., 2008b).

eIF3 interacts with eIF4G through eIF3e and eIF3f (LeFebvre et al., 2006; Masutani et al., 2013) and with 40S ribosomal subunit through eIF3a, b, c and j (Fraser et al., 2007; Hinnebusch, 2006). Also, it promotes mRNA recruitment, assembly of the preinitiation complex and AUG recognition (Chiu et al., 2010; Hinnebusch, 2006; Sokabe et al., 2011; Valásek, 2012). Recently we provided evidence of eIF3 involvement on the mechanism by which PABPC1 inhibits NMD (Peixeiro et al., 2012).

The involvement of eIF3 in translation reinitiation as also been reported. Indeed, eIF3 is one of the initiation factors required to be associated with 40S ribosomal subunit during

the elongation step of translation of a uORF and even after the termination event, so the ribosome can resume scanning and recognize an AUG further downstream (Nielsen et al., 2004; Szamecz et al., 2008). In plants, it has been unequivocally shown that the h subunit of eIF3 is necessary for reinitiation to occur after translation of a uORF (Roy et al., 2010; Zhou et al., 2010). Although this mechanism is still poorly understood, it seems that eIF3h promotes the association between eIF3 complex and 40S ribosomal subunit during translation of a short uORF, allowing ribosomes to retain competence for reinitiation (Roy et al., 2010). Translation of a longer uORF, or of a uORF with a slower translational rate, will result in the loss of the interaction of eIF3 with either eIF4G or the ribosome and reinitiation will be less efficient. This agrees with our results involving this complex with NMD evasion of AUG-proximal nonsense-mutated transcripts (Peixeiro et al., 2012).

In yeast, it was also reported the influence of eIF3a subunit in reinitiation commitment. The yeast transcription factor *GCN4* presents four uORFs in its 5' leader sequence. In this case, there are particular sequences located both 5' of several uORFs and 3' of uORF1 that act together with eIF3a subunit to promote reinitiation after translation of uORF1, but not after translation of the other three uORFs, mainly of uORF4. The underlying mechanism comprises retention of the 40S ribosomal subunit associated to mRNA after the termination event and a stabilization of this interaction, that consequently will potentiate reinitiation (Munzarová et al., 2011; Szamecz et al., 2008). However, it is not known whether eIF3 has the same importance in mammals that it has in yeast, since in the mammalian functional homologue of GCN4, ATF4, 5' and 3' sequences flanking uORF1 are not recognized (Valásek, 2012; Vattem and Wek, 2004).

Interestingly, a recent study has established a correlation between eIF3h and cancer. In this study it is shown that eIF3h is directly involved in the stimulation of protein synthesis, either in normal or cancer cells. However, in cancer cells, this subunit is overexpressed, inducing a higher translational efficiency of the oncogenic mRNA involved in cell growth, which results in a malignant phenotype (Zhang et al., 2008).

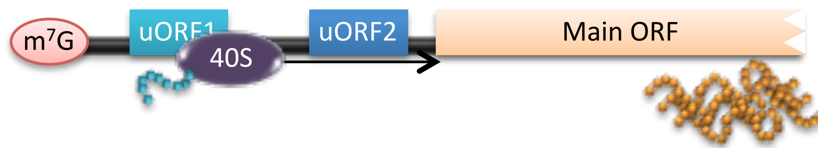
These examples illustrate not only the level of regulation and complexity underlying translation regulation, but also the involvement of eIF3 in this process. Thus, the study of this protein complex can contribute to the knowledge of the mechanisms of uORF impact in translation.

I.2.4. uORFs and the cellular response to stress conditions

As stated above, accumulating evidence has revealed that in response to abnormal stimuli, general translation is inhibited. However, alternative mechanisms of translation initiation and translational control act to maintain the synthesis of certain proteins required either for the stress response or to aid recovery from stress. These pathways are evolutionary conserved and have been shown to significantly impact translation in organisms as diverse as yeast and humans. In many cases, features in the 5' leader sequence of the corresponding mRNAs, such as IRESs and regulatory uORFs, are important for them to evade global repression of translation. For example, when eIF2 is phosphorylated and consequently global translation is inhibited, the presence of uORF(s) in a transcript can promote an increase in the corresponding protein levels (Figure I.5.). The yeast transcription factor *GCN4* is one of the best-studied examples of a transcript containing uORFs that are able to respond to cell stress. This transcript harbors four uORFs in its 5' leader sequence. The first of the four uORFs is always efficiently translated regardless of the nutritional conditions. In unperturbed cells, rapid reloading of ribosomes and initiation cofactors allows translation of uORFs 2-4 while inhibiting the translation of the main ORF. In conditions of amino acid starvation, reinitiation after translation of the uORF1 is less efficient since there is less ternary complex available. Consequently, reinitiation will take more time/distance to occur and the ternary complex will only be available by the time the 40S ribosomal subunit has already bypassed the subsequent uORFs, thereby augmenting the recognition of the main AUG (Hood et al., 2009). This mechanism allows a fast response to nutritional stress (Hinnebusch, 2005; Mueller and Hinnebusch, 1986). The stress response gene that encodes the activating transcription factor 4 (*ATF4*) is the prototypical mammalian example of this type of regulation (Lewerenz et al., 2012). ATF4 promotes transcriptional upregulation of specific target genes in response to cellular stress. *ATF4* expression at the translational level is regulated by two uORFs, with the second overlapping the AUG of the *ATF4* coding sequence, although in a different reading frame (Table I.1.). Under normal conditions, when eIF2α is not phosphorylated and ternary complex is not limiting, the scanning preinitiation complex recognizes the first uORF and translates a short peptide, and the 60S ribosome dissociates upon reaching the stop codon marking

the end of the uORF. The 40S ribosomal subunit that remains associated with the mRNA is then able to recruit the ternary complex and initiate translation of the second uORF. Because the second uORF overlaps with the main coding sequence, this prevents translation of the *ATF4* coding sequence. However, in conditions of reduced ternary complex availability, initiation of the second uORF is less likely, as there is less chance of the scanning ribosomal subunit to recruit the ternary complex required for start codon recognition (Blais et al., 2004; Lewerenz et al., 2012) (Table I.1.). By this mechanism, a reduction in active eIF2 induces increased protein expression from mRNAs carrying the correct arrangement of uORFs (Figure I.5.A) (Palii et al., 2008; Ron and Harding, 2007). This is also the case for the human ATF5 (Watatani et al., 2008); like ATF4, ATF5 is a transcription factor of the cAMP-response element binding protein (CREB)/ATF family, which is encoded by two transcripts (*ATF5α* and *ATF5β*) with alternative 5' leader sequences (Hansen et al., 2002). The 5' leader sequences of *ATF4* and *ATF5α* have similar configurations and both contain two conserved uORFs (Blais et al., 2004; Hansen et al., 2002; Palii et al., 2008; Watatani et al., 2008) (Table I.1.). Similarly to what occurs in the *ATF4* mRNA, the *ATF5α* uORFs are involved in protecting cells from amino acid limitation, as well as from arsenite-induced oxidative stress, through phosphorylation of eIF2α (Watatani et al., 2008). Interestingly, the regulatory mechanisms governing variable ATF4 and ATF5 expression in response to eIF2α phosphorylation, under different conditions of stress, is likely due to a combined effect of translational and transcriptional control of *ATF4* and *ATF5* mRNAs. In addition, global cellular adaptation to stress includes the transcriptional upregulation of ATF4 and ATF5 targets. Nevertheless, other genes activated by eIF2α phosphorylation may also function in conjunction with ATF4 and ATF5, as well as their targets.

A Stress: after uORF1 translation, high eIF2α-P increases bypass of subsequent uORFs



B Stress: high eIF2α-P increases uORF leaky scanning

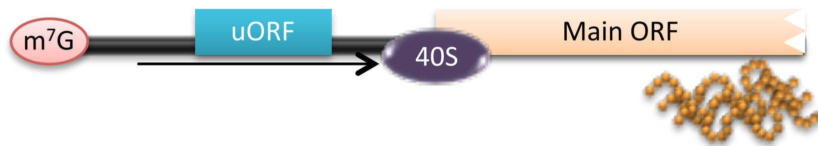


Figure I.5. uORFs response to stress conditions.
(A) In response to stress conditions, the presence of more than one uORF in a transcript can promote an increase in translation efficiency of the main ORF; the reinitiation after translation of the uORF1 is less efficient since there is less ternary complex available. Consequently, reinitiation will take more time/distance to occur and the ternary complex will only be available by the time the 40S ribosomal subunit has already bypassed the subsequent uORFs, augmenting the recognition of the main AUG. **(B)** In response to stress conditions, the presence of one uORF in a transcript can promote an increase of the corresponding protein levels; the higher levels of phosphorylated eIF2α contribute to increase leaky scanning of the uORF and translation of the main ORF is favored.

As stated, genes with uORFs in their transcripts are good candidates to be upregulated in response to eIF2α phosphorylation. An example of regulated expression *via* uORF(s) is the carnitine palmitoyltransferase 1C (*CPT1C*) gene (Table I.1.). CPT1C regulates metabolism in the brain in situations of energy surplus. The presence of a uORF in the 5' leader sequence represses the expression of the main ORF. However, this repression is relieved in response to specific stress stimuli like glucose deprivation and palmitate-BSA treatment (Lohse et al., 2011). The mRNAs that encode the CCAAT/enhancer-binding protein homologous protein (*CHOP*) (Chen et al., 2010; Palam et al., 2011), growth arrest DNA-inducible gene 34 (*GADD34*) (Lee et al., 2009) and β-site amyloid precursor protein-cleaving enzyme 1 (*BACE1*) (Mouton-Liger et al., 2012; O'Connor et al., 2008) are also examples where the phosphorylation of eIF2α is responsible for the translational derepression (Table I.1.). The majority of these transcripts bear more than one uORF resulting in an effect similar to the one seen in *GCN4*, *ATF4* or *ATF5α* (see above). Although it seems that transcripts with only one uORF can also be regulated by this

mechanism as is the case for the *CHOP* transcript, the underlying molecular basis for this remains poorly understood but it seems that the uAUG is less recognized during stress conditions (Figure I.5.B). Chen et al. have reported that in cells under anisomycin treatment, uORF-mediated *CHOP* translation is controlled by the dissociation of phosphorylated eIF4E from 4E-BP. A key finding of this study is that the phosphorylation of both eIF4E and eIF2α is crucial for *CHOP* stress-responsive translational regulation (Chen et al., 2010). These authors also shown that anisomycin activates both Mnks and mTOR signaling pathways which converge at eIF4E for *CHOP* uORF-mediated translation, in addition to phosphorylated eIF2α (Chen et al., 2010). Despite the fact that many questions still need to be answered, these two pathways have been implicated in the induction of translation of uORF-containing transcripts, such as protein kinase C (Raveh-Amit et al., 2009), *ATF4* (Palii et al., 2008) in response to amino acid starvation, Cbp/p300-interacting transactivator with Glu/Asp-rich carboxy-terminal domain 2 (*CITED2)* (van den Beucken et al., 2007) in response to hypoxia, or *CPT1C* (Lohse et al., 2011) in response to specific stress stimuli, namely glucose deprivation and palmitate-BSA treatment.

In addition, vascular endothelial growth factor A (*VEGF-A*) (Bastide et al., 2008), *p27* (Göpfert et al., 2003), endothelial cell tyrosine kinase receptor (*TIE2*) (Park et al., 2005), N-deacetylase/N-sulfotransferase (*NDST*) (Grobe and Esko, 2002), and cationic amino acid transporter 1 (*CAT1*) (Fernandez et al., 2002; Yaman et al., 2003), provide other examples of transcripts regulated by functional uORFs (Table I.1.); however, it is interesting to note that in these cases, uORFs are located within an IRES, which is translated through a cap-independent mechanism. In the case of *CAT1* mRNA, it has been demonstrated that induction of IRES activity requires the translation of the uORF located within the IRES (Yaman et al., 2003). The translation of the uORF unfolds an inhibitory structure in the mRNA 5’ leader sequence creating an active IRES through RNA-RNA interactions between the 5’ end of the leader sequence and downstream sequences, which increases CAT1 protein synthesis (Yaman et al., 2003).

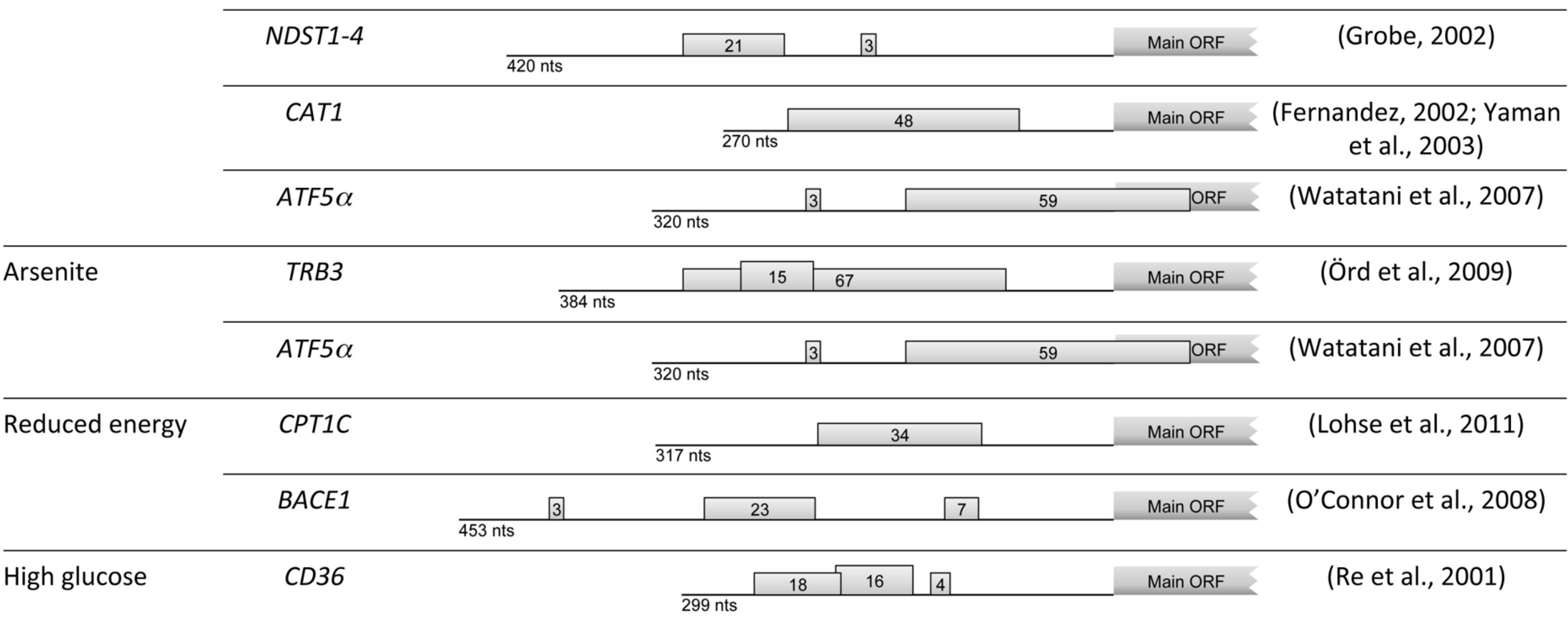
There are other interesting examples of how *cis*-acting elements and different gene expression mechanisms can act together for a specific outcome (Koschmieder et al., 2007; Örd et al., 2009; Re et al., 2001) (Table I.1.). In the case of the tribbles homolog 3 (*TRB3*) gene, in response to arsenite exposure, there is binding of ATF4 to the promoter

which leads to a switch in promoter usage; this results in the production of a transcript with no uORF, while under normal conditions two transcripts are produced: one with a uORF in the 5' leader sequence and one with no uORF (Örd et al., 2009). For the *C/EBP*α gene, 2-cyano-3,12-dioxooleana-1,9-dien-28-oic acid (CDDO) augments C/EBPα activity in acute myeloid leukemia cells by translationally enhancing the p42/p30 *C/EBP*α isoform ratio in a *C/EBP*α uORF-dependent manner (Koschmieder et al., 2007). In another case, high glucose conditions increase *CD36* mRNA translational efficiency that results in increased expression of the macrophage scavenger receptor CD36, due to ribosomal reinitiation following translation of a uORF. Increased translation of macrophage *CD36* transcript provides a mechanism for accelerated atherosclerosis in diabetics (Re et al., 2001).

A final example is the *HER2* oncogene that encodes a 185 kDa transmembrane receptor tyrosine kinase. HER2 overexpression occurs in numerous primary human tumors and contributes to 25-30% of breast and ovarian carcinomas. Synthesis of HER2 is controlled in part by a uORF that represses translation of the downstream main coding region. HER2 overexpression in cancer cells seems to be due to an interaction of 3'UTR with the uORF through an RNA-binding protein, thus overriding translational inhibition mediated by the *HER2* uORF (Mehta, 2006). Even though the precise mechanism by which this interaction occurs is still unknown, it provides further evidence of how uORFs and other gene expression pathways can act together for the modulation of the expression of regulatory genes and of the individual phenotype. In addition, the examples shown here suggest that the translational control mediated by uORFs may involve several steps of mRNA metabolism, may include unfolding of mRNA structures, specific sequences or *trans*-acting factors, may occur in a context-dependent manner and may respond differently to stress-activated translation initiation factors.

Table I.1. Examples of human genes encoding mRNAs that, under stress conditions, evade global repression of translation and are upregulated due to the presence of uORFs.
For each mRNA, the schematic representation of the 5' leader sequence is shown with the length (in nucleotides; nts) indicated below each representation; boxes with numbers represent the uORF(s), where the number indicates the corresponding length in codons.

Stress Condition	Upregulated Gene	Schematic Representation of the 5' Leader Sequence	Reference
Hypoxia	*ATF4*	281 nts; 3; 59; ORF	(Blais et al., 2004; Lewerenz et al., 2011)
	CITED2	246 nts; 3; Main ORF	(van den Beucken et al., 2007)
	VEGF-A	231 nts; 3; Main ORF	(Bastide et al., 2008)
	TIE2	467 nts; 8; 20; 17; 11; 17; Main ORF	(Park et al., 2005)
Endoplasmic reticulum	*CHOP*	250 nts; 34; Main ORF	(Chen et al., 2010; Palam et al., 2011)
	GADD34	269 nts; 27; 26; Main ORF	(Lee et al., 2008)
	BACE1	453 nts; 3; 23; 7; Main ORF	(Mouton-Liger et al., 2012)
	C/EBPα	150 nts; 5; Main ORF	(Koschmieder et al., 2007)
Nutrient deprivation/amino acid starvation	*PKC*η	659 nts; 18; 26; Main ORF	(Raveh-Amit et al., 2009)
	ATF4	281 nts; 3; 59; ORF	(Palii et al., 2008)



	NDST1-4	(Grobe, 2002)
	CAT1	(Fernandez, 2002; Yaman et al., 2003)
	ATF5α	(Watatani et al., 2007)
Arsenite	*TRB3*	(Örd et al., 2009)
	ATF5α	(Watatani et al., 2007)
Reduced energy	*CPT1C*	(Lohse et al., 2011)
	BACE1	(O'Connor et al., 2008)
High glucose	*CD36*	(Re et al., 2001)

I.2.5. uORFs and human disease

Given that uORFs reduce translational efficiency, it is clear that polymorphisms or mutations that create, disrupt, or modify uORFs are likely to affect protein expression and may impact individual phenotypes. Indeed, when Calvo and colleagues searched for uORF-altering variants within 12 million single nucleotide polymorphisms (SNPs) in the human dbSNP database (Calvo et al., 2009; Sherry et al., 2001), they identified uORFs created or deleted by a polymorphism in 509 genes; 366 of these genes encode transcripts harboring multiple uORFs, whereas the remaining 143 mRNAs have a single uORF (Calvo et al., 2009). This study also shown that these uORFs induce a 30-60% decrease in protein levels when compared to the protein levels expressed from the corresponding allele without the uORF-altering SNP variant (Calvo et al., 2009). As a concrete example, an SNP was described that alters the human clotting factor XII (*FXII*) 5’ leader sequence, and has been associated with several thromboembolic conditions due to differences in circulating FXII plasma levels (Bersano et al., 2008). This SNP consists of a common C to T polymorphism with prevalence of the T allele estimated at 20% in Caucasian and 70% in Asian populations (Bach et al., 2008; Kanaji et al., 1998). It is located at position -4 of the *FXII* 5’ leader sequence (where the A of the main AUG start codon is nucleotide +1), introduces a very short uORF (with 2 codons), and simultaneously alters the AUG Kozak sequence context of the factor *FXII* coding sequence. Kanaji and colleagues have experimentally confirmed that the T allele does not affect mRNA levels, but reduces protein levels by about 50%, increasing the predisposition to thrombosis (Kanaji et al., 1998). More recently, it was demonstrated that this protein reduction is indeed due to the presence of the 2 codon uORF, while the disruption of the Kozak consensus sequence is not responsible for the observed variation in human FXII protein levels (Calvo et al., 2009, 2009) (Table I.2.). This example shows how SNPs, found through genetic analyses in the 5’ leader sequence of transcripts, cannot be disregarded, as even if they do not affect mRNA levels they can affect protein levels and be associated with human disease. This region should, therefore, be systematically explored when investigating the molecular mechanism of a disease.

In addition to polymorphisms that can affect uORFs, rare mutations that create or disrupt uORFs may also cause disease, as has been shown for several human genes (Table I.2.). Indeed, several mutations that eliminate or create uORFs that alter protein levels have been associated with human disease. Calvo and colleagues have experimentally demonstrated, in five genes (*HBB*, *PRKAR1A*, *IRF6*, *SRY*, and *SPINK1*), that mutations that create a uORF decrease protein expression levels to 30%, or less, of those from the normal allele, and these reduced protein levels are responsible for the associated disease phenotype (Calvo et al., 2009). Notably, in the *SRY* and *SPINK1* genes, the mutation creates a second uORF within the 5' leader sequence. Thus, the strong suppression of protein expression by these mutations offers a simple mechanistic basis for their pathogenicity (Calvo et al., 2009). Another study has shown that predisposition to melanoma can be caused by mutations that introduce a uORF into the 5' leader sequence of the mRNA encoding the cyclin-dependent kinase inhibitor protein (*CDKN2A*) (Bisio et al., 2010; Liu et al., 1999). Other examples of human diseases associated with mutations that create a uORF include familial hypercholesterolemia (low-density lipoprotein receptor gene; *LDLR*) (Sözen et al., 2005), cystic fibrosis (*CFTR*) (Lukowski et al., 2011), congenital hyperinsulinism (potassium inwardly-rectifying channel, subfamily J, member 11; *KCNJ11*) (Huopio et al., 2002), rhizomelic chondrodysplasia punctata (peroxisomal biogenesis factor 7; *PEX7*) (Braverman et al., 2002), proopiomelanocortin deficiency syndrome (proopiomelanocortin; *POMC*) (Krude et al., 1998), levodopa-responsive dystonia (guanosine triphosphate cyclohydrolase I; *GCH1*) (Tassin et al., 2000) and juvenile hemochromatosis (hepcidin; HAMP) (Rideau et al., 2007) (Table I.2.). Although the majority of the polymorphisms/mutations referred here that create a uORF have been experimentally tested for their influence on translation, in the case of *LDLR*, *KCNJ11*, *PEX7*, *POMC* and *GCH1* mRNAs, further studies are needed to confirm the effect of the corresponding mutation on translational efficiency (Table I.2.).

Contrary to the effect of mutations that create a uORF, the repression exerted by a functional uORF can be modulated by mutations, or alternative processing of the transcript, that disrupt the uORF, thus influencing the translational rate of the main ORF. In either case, there is a change in organism homeostasis that affects individual phenotype. An illustration of a genetic alteration that disrupts a uORF is a mutation described in the initiation codon of an inhibitory 34 codon uORF located in the 5' leader

sequence of the mRNA that encodes the human hairless homolog (*HR*) protein. This mutation has been associated with the symptomatic condition of Marie Unna hereditary hypotrichosis, which is a rare autosomal dominant form of genetic hair loss (Baek et al., 2009; Wen et al., 2009). Functional analysis shown that this mutation results in increased translation of the main *HR* physiological ORF (Baek et al., 2009; Wen et al., 2009). Another noteworthy example is the thrombopoietin (*TPO*) gene (Cazzola and Skoda, 2000). Translation of *TPO* mRNA is physiologically strongly inhibited by the presence of seven uORFs in its 5' leader sequence. Directed mutagenesis of all uAUGs in the *TPO* mRNA restores translational efficiency, demonstrating that translational inhibition of TPO biosynthesis is mediated by uORFs (Cazzola and Skoda, 2000). The uORF defined by the seventh uAUG was shown to exert the strongest negative effect on translation. This uAUG is in a good Kozak consensus context and the uORF extends beyond the physiological start site, thus preventing reinitiation (Cazzola and Skoda, 2000). Mutations in the 5' leader sequence of the *TPO* gene, which cause hereditary thrombocytosis, inactivate the inhibitory function of uORF7 and abolish this translational control (Cazzola and Skoda, 2000; Ghilardi and Skoda, 1999; Ghilardi et al., 1999; Kikuchi et al., 1995; Kondo et al., 1998; Wiestner et al., 1998). In these cases, pathologically high TPO levels are observed, leading to an increased number of platelets in the peripheral blood and increased thrombosis risk. One particular mutation was demonstrated to introduce a translation termination codon in the 5' leader sequence in frame with uORF7. As the new in frame stop codon produces a uORF entirely located in 5' leader sequence it confers the ability to reinitiate at the main ORF. This new regulation mechanism by uORF7 produces a weaker translational repression causing an increase of the TPO protein levels (Ghilardi and Skoda, 1999; Ghilardi et al., 1999; Kikuchi et al., 1995). In another case, a point mutation (G to C transversion) in the +1 position of the splice donor site of intron 3 causes exon skipping and results in loss of exon 3 that normally encodes a large part of the 5' leader sequence. As a consequence, the mutant *TPO* mRNA lacks uORF7, which normally inhibits translation, and encodes a novel N-terminus created by fusion of uORF5 with the *TPO* coding sequence (Wiestner et al., 1998). A different mutation consists in a single G nucleotide deletion in the 5' leader sequence of the *TPO* gene that causes a frameshift in the 5' leader sequence of *TPO* mRNA, which places uORF7 in frame with the *TPO* coding sequence, neutralizing the

strong inhibitory effect of uORF7 and creating a novel N-terminus for the TPO protein (Kondo et al., 1998). These data clearly illustrate how TPO expression is tightly regulated at the translational level.

As mentioned above, uORFs may differ in their efficiency and in the mechanisms by which they exert translational repression of the main ORF. In some cases uORFs repress translation because the corresponding encoded peptide is able to promote a blockage in the translating ribosome (Lovett and Rogers, 1996). Consequently, specific nucleotide substitutions that alter the uORF coding sequence and originate an amino acid substitution, might affect the efficiency of ribosomal blockage and thus protein expression from the main ORF. For example, amino acid substitutions that decrease efficiency of ribosomal blockage might decrease the translational repression exerted by the uORF, and therefore they might increase protein levels, which might lead to clinical manifestations. This is the case for the human dopamine D3 receptor (*DRD3*) gene (Sivagnanasundaram et al., 2000). Sivagnanasundaram and colleagues have screened for polymorphisms to assess their contribution to the association of DRD3 with schizophrenia. Their data have shown that one of the SNPs found in the 5' leader sequence encodes a change of one amino acid residue from lysine to glutamic acid within a 36 codon uORF, which correlates to an increased schizophrenia predisposition (Sivagnanasundaram et al., 2000) (Table I.2.). Another example is the G to A transition described in the *WDR46* gene that originates an amino acid change from glycine to arginine at codon 18 of a uORF in the *WDR46* transcript; this variant is associated with higher risk of aspirin-exacerbated respiratory disease (Pasaje et al., 2012) (Table I.2.). In a different study, authors identified the transforming growth factor-β3 (*TGFβ3*) gene as being involved in arrhythmogenic right ventricular cardiomyopathy, a progressive and genetically determined myocardial disease, due to a G to A transition in the *TGFβ3* 5' leader sequence, which leads to an arginine to histidine substitution at codon 36 of a uORF with 88 codons; it has been experimentally proven that this change causes an increase in the *TGFβ3* protein levels (Beffagna et al., 2005) (Table I.2.). Moreover, the human *HT3A* mRNA, which encodes the subunit A of the type 3 receptor for 5-hydroxytryptamine (serotonin) contains two uORFs, in frame with the main ORF. A -42C to T mutation in the second uORF of *HT3A* is associated with bipolar affective disorder and major depression; it has been experimentally shown that this mutation increases

translation efficiency of the 5-HT3A subunit (Niesler et al., 2001) (Table I.2.). For these pathologies, elucidating the mechanisms through which uORFs can affect downstream translational efficiency depending on the amino acid sequence of the uORF-encoded peptide, may constitute a tool for the development of new and more effective drug treatments.

Another intriguing regulatory function of uORFs is observed in transcripts harboring alternative downstream initiation codons within their main ORF. This is exemplified by CCAAT/enhancer binding protein β and α (*C/EBPβ* and *C/EBPα*, respectively), in which uORFs control the expression ratio of functionally distinct protein isoforms by sensing the translational status of the cell (Wethmar et al., 2010a). Recently, an interesting work using *C/EBP* uORF mice has corroborated the role of uORFs in pathophysiology (Table I.2.). This genetic mouse model has provided the proof-of-principle for the physiological relevance of uORF-mediated translational control in mammals (Wethmar et al., 2010a, 2010b), as targeted disruption of the uORF initiation codon within the *C/EBPβ* mRNA resulted in deregulated C/EBPβ protein isoform expression, associated with defective liver regeneration and impaired osteoclast differentiation (Wethmar et al., 2010a, 2010b).

Another fascinating regulatory function of uORFs occurs in transcripts encoded by genes with cryptic promoters – e.g. the oncoprotein MDM2, which is overexpressed in a number of human tumors, particularly in osteosarcomas (Oliner et al., 1992). This overexpression can result from a change in mRNA structure due to a switch in promoter usage. There are two transcripts from the *MDM2* gene that differ only in their 5' leader sequence: a long form (*L-MDM2*) that carries two uORFs and a short form (*S-MDM2*) without uORFs. In these tumors, the switch in promoter usage yields enhanced cellular levels of the *S-MDM2* mRNA isoform, which is efficiently translated. On the contrary, the *L-MDM2* mRNA is less efficiently transcribed and its translation is repressed by two functional uORFs (Brown et al., 1999). Overall, MDM2 becomes overexpressed in tumors due to the preferential transcription of the *S-MDM2* isoform that is not under translational regulation (Table I.2.) (Brown et al., 1999). This set of data illustrates how disrupted uORF-mediated translational regulation can affect expression levels of oncogenes or tumor suppressor genes, and thus contribute to the pathophysiology of many forms of cancer.

As previously discussed, uORF mediated translational regulation has the ability to respond to stress conditions, which is a feature that can also be associated to human disease. This may be the case for *BACE1* gene, which encodes an enzyme involved in the production of beta-amyloid plaques in the brain of patients with Alzheimer's disease (AD). The enhanced production of this enzyme occurs without corresponding changes in *BACE1* mRNA levels and seems to occur at the translational level. The complex *BACE1* 5' leader sequence contains three uORFs preceding the *BACE1* initiation codon that might be involved in the enhanced production of this enzyme characteristic of humans with AD. It has been hypothesized that aging and other factors such as cardiovascular disease or traumatic brain injury might impair brain energy metabolism that leads to a higher phosphorylation of eIF2α. Indeed, it has been shown that energy deprivation induces phosphorylation of the eIF2α, which increases the translation of *BACE1* mRNA (O'Connor et al., 2008). Under these conditions, the BACE1 protein levels might increase due to a uORF(s) mediated translational derepression leading to beta-amyloid overproduction, which could be an early, initiating molecular mechanism in sporadic AD (Table I.1.) (Lammich et al., 2004; Mihailovich et al., 2007; Mouton-Liger et al., 2012; O'Connor et al., 2008; Rogers Jr et al., 2004; Zhou and Song, 2006). However, some other data is consistent with the hypothesis that the translation efficiency of the *BACE1* initiation codon may be increased in patients with Alzheimer's disease by molecular mechanisms that enhance shunting or increase the relative accessibility of the *BACE1* initiation codon, without the involvement of uORF(s) (Rogers Jr et al., 2004).

Although phosphorylation of eIF2α in response to cellular stress has been unequivocally shown to increase *BACE1* translation (Mouton-Liger et al., 2012; O'Connor et al., 2008), the involvement of uORF(s) in the stress-dependent mechanism of translation initiation is more controversial (Lammich et al., 2004; Mihailovich et al., 2007; Rogers Jr et al., 2004; Zhou and Song, 2006). Indeed, it has been shown that the *BACE1* uORF(s) have little or no effect on BACE1 expression in unstressed cells (Lammich et al., 2004; Rogers Jr et al., 2004). Instead, it may be the GC-rich region of the *BACE1* 5'UTR that forms a constitutive translation barrier, which could prevent the ribosomes from efficiently translating the *BACE1* mRNA (Lammich et al., 2004). The exact role of the three *BACE1* uORFs in its translational regulation needs further evaluation.

In the examples discussed here, all the uORF-altering polymorphisms/mutations have been reported in the literature as demonstrating segregation with the disease. However, some of them, although present within a gene known to underlie the disease when disrupted, were not followed up experimentally (by using reporter assays) to confirm their impact on translational efficiency (Table I.1.). In any case, these examples highlight the importance of searching for uORF changes, in addition to coding alterations, underlying disease and draw attention to the need for recognition of these structures as potential therapeutic targets.

The recent advances in next-generation sequencing technologies certainly represent a quantum leap towards (i) the identification of a large number of novel disease-associated uORF alterations, (ii) the subsequent uncovering of predictive genotype-phenotype correlations in many areas of human pathology, and (iii) the recognition of uORFs as possible therapeutic targets.

Table I.2. Examples of human diseases associated with polymorphisms or mutations that introduce/eliminate uORFs or modify the encoded uORF peptide

Disease	Gene	Mode of Pathogenesis	Reference
		Polymorphisms/mutations that create uORFs	
1. Thrombotic predisposition	*FXII*	The -4C to T polymorphism creates a uORF that reduces mRNA translation efficiency from the main ORF[(a)]	(Bach et al., 2008; Bersano et al., 2008; Calvo et al., 2009; Kanaji et al., 1998)
2. β-Thalassemia	*HBB*	The -29G to A mutation creates a new translation initiation codon in a favorable Kozak consensus sequence, which leads to the introduction of a new uORF that overlaps with the main ORF, but out of frame, and decreases translation efficiency from the main ORF[(a)]	(Calvo et al., 2009; Oner et al., 1991)
3. Carney complex type 1	*PRKAR1A*	The -97G to A mutation creates a uORF that overlaps with the main ORF, but out of frame, and decreases translation efficiency from the main ORF[(a)]	(Calvo et al., 2009)
4. Van der Woude syndrome	*IRF6*	The -48A to T mutation creates a uORF that overlaps with the main ORF, but out of frame, and decreases translation efficiency from the main ORF[(a)]	(Calvo et al., 2009; Kondo et al., 2002)
5. Gonadal dysgenesis	*SRY*	The -75G to A mutation creates a second uORF and reduces mRNA translation efficiency from the main ORF[(a)]	(Calvo et al., 2009; Poulat et al., 1998)
6. Hereditary pancreatitis	*SPINK1*	The -53C to T mutation creates a uORF and reduces mRNA translation efficiency from the main ORF[(a)]	(Calvo et al., 2009; Witt et al., 2000)
7. Melanoma predisposition	*CDKN2A*	Both described -21C to T and -34G to T mutations create a uORF that reduces mRNA translation efficiency from the main ORF[(a)]	(Bisio et al., 2010; Liu et al., 1999)
8. Familial hypercholesterolemia	*LDLR*	A single C nucleotide deletion (at position -22) creates a uORF and reduces mRNA translation efficiency from the main ORF[(b)]	(Sözen et al., 2005)

9. Disseminated bronchiectasis	*CFTR*	The -34C to T mutation creates a uORF overlapping, but out of frame, with the CFTR protein coding sequence, which decreases gene expression by reducing mRNA stability and translation efficiency from the main ORF[(a)]	(Lukowski et al., 2011)
10. Congenital hyperinsulinism	*KCNJ11*	The -54C to T mutation creates a new translation initiation codon in a favorable Kozak consensus sequence, which leads to the introduction of a new uORF that overlaps with the main ORF, but out of frame, and decreases translation efficiency from the main ORF[(b)]	(Huopio, 2002)
11. Rhizomelic chondrodysplasia punctata	*PEX7*	The -45C to T mutation creates a new translation initiation codon in a favorable Kozak consensus sequence, which leads to the introduction of a new uORF that overlaps with the main ORF, but out of frame, and decreases translation efficiency from the main ORF[(b)]	(Braverman et al., 2002)
12. Proopiomelanocortin deficiency	*POMC*	The -11C to A mutation creates a new translation initiation codon in a favorable Kozak consensus sequence, which leads to the introduction of a new uORF that overlaps with the main ORF, but out of frame, and decreases translation efficiency from the main ORF[(b)]	(Krude et al., 1998)
13. Levodopa responsive dystonia	GCH1	The -22C to T mutation creates a new translation initiation codon that leads to the introduction of a new uORF overlapping with the main ORF, but out of frame, and decreases translation efficiency from the main ORF[(b)]	(Tassin et al., 2000)
14. Juvenile hemochromatosis	*HAMP*	The -25G to A mutation creates a new translation initiation codon, which leads to the introduction of a new uORF overlapping with the physiological ORF, but out of frame, and decreases translation efficiency from the main ORF[(a)]	(Rideau et al., 2007)
		Polymorphisms/mutations that disrupt uORFs	
15. Marie Unna hereditary hypotrichosis	*HR*	The -321A to G mutation disrupts one of the existing uORFs and results in an increased translational efficiency of the main *HR*	(Baek et al., 2009; Wen et al., 2009)

		physiological ORF[a]	
16. Thrombocythemia	*TPO*	-31G to T mutation generates a new stop codon in uORF 7 and thereby shortens uORF 7 by 42 nucleotides. The truncated uORF 7 no longer extends past the physiological initiation codon, and thus it improves translational efficiency by allowing translation reinitiation[a]	(Cazzola and Skoda, 2000; Ghilardi and Skoda, 1999; Ghilardi et al., 1999; Kikuchi et al., 1995)
		The G to C transversion in the splice donor site of intron 3 of the *TPO* gene leads to mRNAs with shortened 5' leader sequence that are more efficiently translated than the normal *TPO* transcripts because they lack uORF 7, which normally inhibits translation; a novel N-terminus is created by fusion of uORF 5 with the *TPO* coding sequence[a]	(Wiestner et al., 1998)
		A single G nucleotide deletion (at position -50) in the 5' leader sequence of the *TPO* gene causes a frameshift in the 5' leader sequence of *TPO* mRNA that places uORF 7 in frame with the *TPO* coding sequence, neutralizing the strong inhibitory effect of uORF 7 and creating a novel N-terminus for the TPO protein[a]	(Kondo et al., 1998)
Polymorphisms/mutations that modify the encoded uORF peptide			
17. Schizophrenia predisposition	*DRD3*	The -204A to G polymorphism within a 36-codon uORF originates a Lys9Glu amino acid substitution in the uORF-encoded peptide that might decrease efficiency of ribosomal blockage; this change causes an increase in the DRD3 protein levels[b]	(Sivagnanasundaram et al., 2000)
18. Aspirin-exacerbated respiratory disease	*WDR46*	The -36G to A polymorphism originates a Gly18Arg amino acid substitution in the uORF-encoded peptide[b]	(Pasaje et al., 2012)
19. Arrhythmogenic right ventricular cardiomyopathy	*TGFβ3*	The -30G to A mutation within an 88-codon uORF originates a Arg36His amino acid substitution in a putative 88–amino acid inhibitory peptide encoded by the uORF; this change causes an	(Beffagna et al., 2005)

		increase in the TGF-β3 protein levels[a]	
20. Bipolar affective disorder and major depression	*HT3A*	-42C to T mutation originates a Pro16Ser amino acid substitution in the uORF-encoded peptide and is postulated to decrease the efficiency of the uORF repression causing an increase in the HT3A protein levels[a]	(Niesler et al., 2001)
		Other alterations	
21. Acute myeloid leukemia	*C/EBP*α	The *C/EBP*α uORF modulates the expression ratio of three N-terminally distinct protein isoforms that are translated from subsequent in frame initiation codons within the *C/EBP*α transcript; an increase in expression of the shorter isoform is associated with acute myeloid leukemia[a]	(Wethmar et al., 2010a)
22. Breast cancer	*C/EBP*β	The *C/EBP*β uORF modulates the expression ratio of three N-terminally distinct protein isoforms that are translated from subsequent in frame initiation codons within the *C/EBP*β transcript; an increase in expression of the shorter isoform due to the inactivation of the uORF is associated with breast cancer[a]	(Wethmar et al., 2010a, 2010b)
23. Several tumors	*MDM2*	A switch in promoter usage favors transcription of an isoform without uORFs which overexpresses MDM2 protein in comparison with what occurs in normal cells, where one isoform with two uORFs is mainly expressed[a]	(Brown et al., 1999)
24. Alzheimer's disease	*BACE1*	Elevated levels of phosphorylated eIF2α induce a bypass of the inhibitory mechanism exerted by *BACE1* uORFs, which leads to enhanced BACE1 expression[a]	(Mihailovich et al., 2007; Zhou and Song, 2006)

Position of the mutation is relative to the main AUG start codon, where the A is nucleotide +1. [a]It has been experimentally tested to affect translational efficiency. [b]It is not experimentally tested.

I.3. Human Erythropoietin (EPO)

Erythropoiesis is a process that has been described almost a century ago. Soon emerged the general understanding that erythropoiesis had to be regulated and later on, in the middle of the twentieth century, such role was assigned to erythropoietin (EPO). Shortly after, a recombinant human EPO protein (rhEPO) form was produced for research and clinical purposes (Egrie et al., 1985). Indeed, administration of EPO in patients with anemia proved to be greatly efficient and thereafter it has been broadly used to treat such disorders (Hino et al., 1989; Stein et al., 1991), and hence EPO is one of the most well-studied proteins. Many efforts across the world have been done towards a deep knowledge of EPO structure, regulation, and mode of action.

The kidney was described as the primary source of EPO production in the adult. This organ is able to sense the differences in the blood flow during hypoxia and increase the production and secretion of EPO. As a result, the erythroid precursor cells, induced by EPO signaling, start to proliferate, thus restoring the levels of red blood cell mass, which decreases EPO levels creating a feedback mechanism (Bunn, 1990; Hambley and Mufti, 1990; Krantz, 1991).

Many studies have tried to identify which renal cells are able to produce EPO. The latest reports show that the peritubular fibroblasts in the cortex are the cells expressing EPO mRNA in the kidney (Haase, 2013; Paliege et al., 2010). Apart from the kidney, in the adult, there is also a minor contribution from hepatocytes to the production of circulating EPO. Indeed, during fetal life, the major site of EPO production is the liver (Dame et al., 1998; Hambley and Mufti, 1990). After birth, there is a switch on the production site of EPO from the liver to the kidney. This switch has been the focus of many studies but the precise mechanism capable to regulating this change of production sites still needs further clarification.

The human *EPO* gene is located in chromosome 7 (q11-12). This gene encodes for a 1340 nucleotide transcript (NM_000799), which in turn produces a 193 amino acid protein, given rise to a mature protein of 166 amino acids after cleavage. Then, EPO protein is glycosylated resulting in an increase of the EPO molecular weight from around 18kDa to approximately 30,4-34,4kDa (Bunn, 1990; Krantz, 1991). Its three-dimensional structure

consists in a globular form of four α-helices associated by a hydrophobic interaction (Brines et al., 2008; Bunn, 2013).

I.3.1. EPO signaling pathways

The circulating EPO is able to regulate erythropoiesis by binding to the EPO receptor (EPOR) present in the surface of hematopoietic cells in the bone marrow. When EPO interacts with EPOR, it triggers the homodimerization of EPOR that, in turn, will activate various intercellular signaling pathways that will end up in the control of proliferation, differentiation and death of the erythroid cells (Chateauvieux et al., 2011; Watowich et al., 1994).

The Janus Kinase (JAK)-2 is associated with the cytoplasmic tails of the EPOR. When the EPO protein is associated with the dimeric EPOR, there is a *trans*-phosphorylation of the associated JAK-2, which in turn will phosphorylate eight tyrosine residues on the cytoplasmic region of EPOR, consequently activating: (i) the JAK2/signal transducer and activator of transcription (STAT) pathway, that through the activation of STAT5 is able to increase transcription of specific antiapoptocic genes, such as Bcl-X_L (Bittorf et al., 2000); (ii) the phosphatidylinositol-3 kinase (PI3K)/AKT pathway, that culminates in the enhanced activity of GATA-1, a key transcription factor for the regulation of erythro-specific genes, that also activates Bcl-X_L (Uddin et al., 2000); and (iii) the mitogen-activated protein kinases (MAPKs) family member cascade, that is initiated by the recruitment of the Src homology-2 (SH2) domain-containing adapter proteins, such as Src homology-2 domain-containing transforming protein (SHC), growth factor receptor bound protein-2 (GRB2) and son of sevenlees (SOS) protein, which will activate RAS, inducing a cascade that comprises RAF1, MAPK/extracellular signal-regulated kinase (ERK) kinases (MEK) and ERK itself, a protein associated with the activation of genes related to with the cell proliferation, survival and differentiation (Figure I.6.) (Chen and Sytkowski, 2004).

The hematopoietic properties of the circulating EPO, produced and released by the kidney, depend on the aforementioned signaling pathways. However, many reports shown that the EPO mRNA is also detected in other organs such as the brain (neurons and glial cells), the lung, the heart, the bone marrow, the spleen, the hair follicles, and

the reproductive tract (Dame et al., 1998; Fandrey and Bunn, 1993; Ghezzi and Brines, 2004; Hoch et al., 2011; Weidemann and Johnson, 2009; Yasuda et al., 1998). The EPO protein synthesized in these organs appears to act locally, modulating, for instance, regional angiogenesis and cellular viability. It does not seem to contribute to erythropoiesis (Gassmann and Soliz, 2009; Maiese et al., 2008).

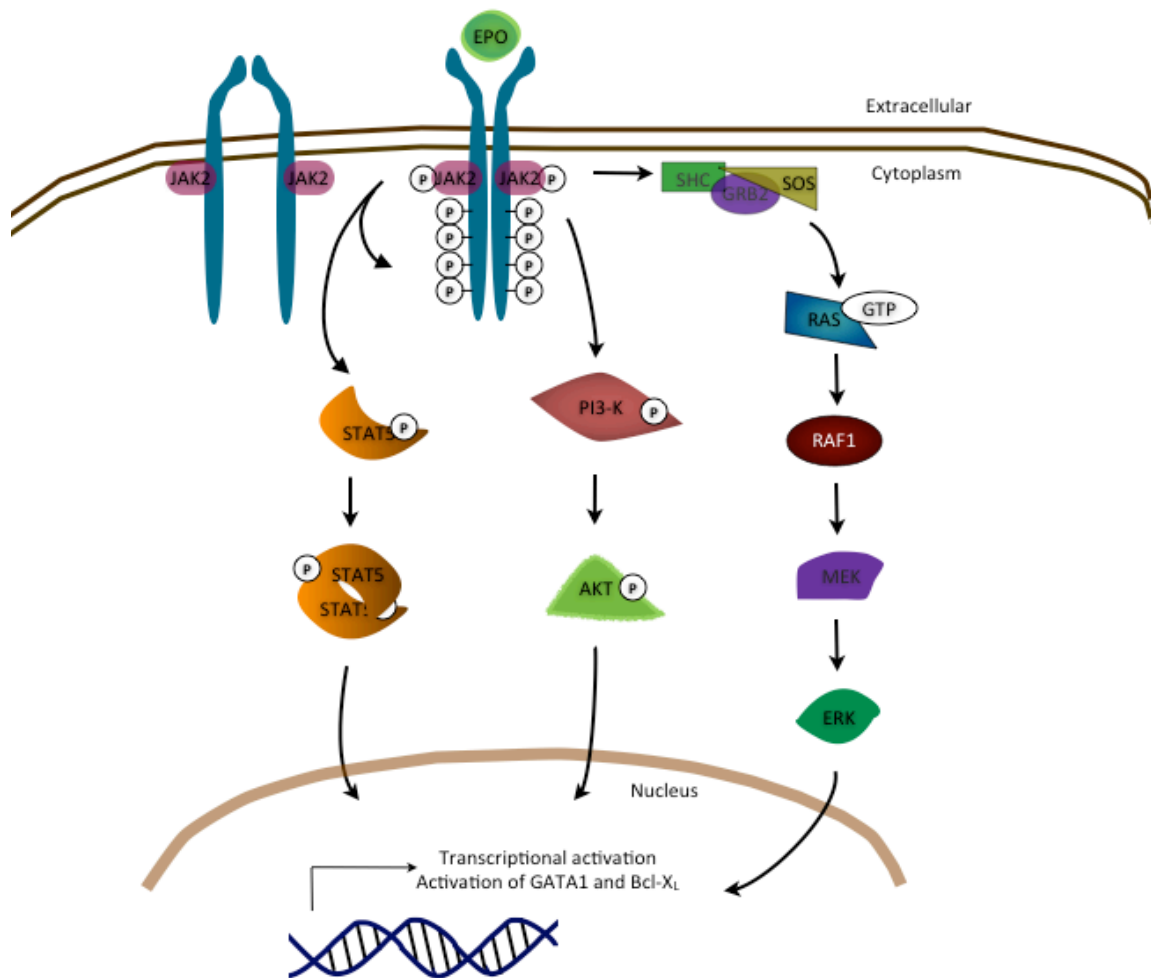


Figure I.6. EPO signalling pathways.
The homodimerization of the EPOR (in blue) occurs only when EPO is bound. This interaction triggers a *trans*-phosphorylation of the JAK2 associated with the cytoplasmic rails of the EPOR. In turn JAK2 will phosphorylate eight tyrosine residues on the cytoplasmic region and several pathways will be activated. One pathway is the phosphorylation and activation of STAT5 by JAK2, represented at the left side of the figure. Also, it can activate PI3-K, which in turn will activate AKT, shown at the middle of the panel. The third pathway is the activation of the MAPK family member cascade. In this pathway SHC, GRB2 and SOS are recruited and activate RAS inducing the RAF1/MEK/ERK cascade, exemplified at the right side of the panel. All of these pathways culminate to the activation of protein that in turn will induce transcriptional

activation of several genes and activation of several factors, such as GATA1 and Bcl-X_L. This will induce cell proliferation, survival and differentiation.

The EPO/EPOR axis was observed in the referred organs and also in a range of tumor and cancer cell lines (Ghezzi and Brines, 2004; Knabe et al., 2005; Moriconi et al., 2013; Yasuda et al., 2010). However, the EPOR structure in non-hematopoietic cells, mainly in neurons, glial and cardiomyocytes, seems to differ from the one characterized above. It has been hypothesized the formation of a heterodimer with the EPOR and CD131, the β common cytokine receptor in these cells (Brines et al., 2004, 2008; Bunn, 2013; Jubinsky et al., 1997). The pathways activated by this receptor are potentially the same previously described to cause inhibition of cell apoptosis, proliferation and migration. However, a slight difference has been reported: a cross-talk between the JAK2 and the nuclear factor-kappa B (NF-kB). The NF-kB is translocated to the nucleus and activates the transcription of antiapoptotic and neuroprotective genes (Digicaylioglu and Lipton, 2001; Ghezzi and Brines, 2004; Kumral et al., 2011). This differential signaling pathway on neuronal cells highlights and contributes to the notion that EPO is putatively important for neuronal protection and, consequently, it might be used has a therapeutic target for the treatment of many neuronal disorders (Dame et al., 2001; Ryou et al., 2012). Indeed, since the cascade downstream the EPO/EPOR binding is mainly the same, the difference between the EPO hematopoietic and non-hematopoietic functions has to be on the EPO protein itself. Actually, EPO binds with different affinities to both receptors and contains specific domains for that interaction. This led to the creation of a modified rhEPO protein that retains its non-hematopoietic functions without stimulation of erythropoiesis (Arcasoy, 2008; Brines et al., 2008; Hoch et al., 2011; King et al., 2007). However, the specific expression, regulation, and action of EPO in the brain are still unclear.

The therapeutic potential of EPO is undoubetly massive and many are the efforts for creating safer and broader therapies using EPO. Thus, the knowledge of the EPO regulation and expression is of great importance and has been the focus of intensive studies. This is a highly complex network that comprises transcriptional and post-transcriptional regulatory mechanisms, and disturbing of one of these mechanisms can lead to clinical disorders. For example, if EPO is underproduced patients develop a

severe anemia. However, if it is overexpressed the result is either polycythemia vera or erythrocytosis, which increases the incidence of thrombotic and hemorrhagic complications (Bunn, 2013; Bushuev et al., 2006).

I.3.2. Transcriptional regulation of the *EPO* gene

Transcriptional regulation has been the major focus of several studies concerning gene expression regulation and *EPO* gene was no exception. This level of regulation is certainly well-studied for *EPO*, being the hypoxia inducible factor (HIF), a transcription factor that is able to increase the amount of *EPO* transcripts during hypoxia, one of the main effectors. HIF is a heterodimer composed of two basic helix-loop-helix proteins, HIFα and HIFβ. HIFα comprises three hypoxia-inducible subunits (HIF-1α, HIF-2α and HIF-3α), while HIFβ is the previously cloned and characterized aryl hydrocarbon receptor nuclear translocator (ARNT), a constitutive nuclear protein (Mole and Ratcliffe, 2007; Wang et al., 1995). HIF-1α is the most abundant of the three subunits and is present in most organs and tissues. On the contrary, HIF-2α has a more limited expression being detected only in endothelial cells under normal physiological conditions (Tian et al., 1997). However, after hypoxia exposure, its levels are increased in several other tissues (Wiesener et al., 2003). HIF-2α together with HIF-1α facilitates oxygen delivery and cellular adaptation to hypoxia by stimulating multiple biological processes, such as erythropoiesis, angiogenesis, and anaerobic glucose metabolism (Semenza, 2001).

In normoxia, the three HIFα subunits are hydroxylated by prolyl-4-hydroxylase domain (PHD) proteins. Then HIFα is ubiquilated and targeted for rapid proteasomal degradation by the von Hippel-Lindau tumor suppressor (VHL)-E3 ubiquitin ligase complex. On the other hand, in low oxygen conditions, HIFα is not hydroxylated, escaping degradation, and is able to exert its function as a transcription factor (Maxwell et al., 1999; Maynard et al., 2003; Semenza, 2001).

HIF recognizes a specific sequence called hypoxia responsive element (HRE), which, in the *EPO* gene, is located in different positions either for kidney or liver EPO induction. Upstream the *EPO* gene is located a kidney inducible element (KIE) that comprises the HRE responsible for the increase of *EPO* mRNA levels in the kidney. On the contrary, for regulation of the *EPO* mRNA at the liver level, the region recognized by HIF is located 3'

of the gene, a region described as dispensable for the renal EPO synthesis (Suzuki et al., 2011; Wang and Semenza, 1993). Despite this study, others shown that in Hep3B, a model cell line for EPO expression, the HRE in the 3' enhancer interacts with the HRE in the promoter to increase *EPO* mRNA expression during hypoxia in about 50- to 100-fold, the levels observed *in vivo* during hypoxia (Blanchard et al., 1992).

This 3' enhancer contains the HRE but also the binding site of another protein, the nuclear receptor HNF-4. This two proteins interact with p300, a transcriptional activator forming a macromolecular complex that is able to activate transcription (Ebert and Bunn, 1998). Interestingly, for the *EPO* upregulation during hypoxia, the essential transcription factor is HIF-2α, which is contrary to the majority of other hypoxia inducible genes that require HIF-1α for their transcriptional regulation (Frede et al., 2011; Warnecke et al., 2004).

As mentioned before, the expression of EPO after birth shifts from the liver to the kidney. This is also due to a transcriptional regulation of the *EPO* gene. The *EPO* promoter is weak and presents some negative regulatory elements, such as the conserved GATA-2 sequence and a NF-kB binding site (Blanchard et al., 1992; Lee-Huang et al., 1993; Imagawa et al., 1994, 1997). However, the difference between fetal and adult expression of *EPO* gene in the liver was credited to the action of GATA-4. GATA-4 production is restricted to the fetal hepatocytes, in which it binds to the EPO promoter, opening the chromatin and leading to the expression of the gene. For that reason, the lack of GATA-4 expression in the adult liver, probably in combination with the interaction of GATA-2 or -3, may lead to the formation of a repressive chromatin structure and inactivation of *EPO* gene transcription (Dame et al., 2004).

Another mechanism that is able to regulate the transcriptional levels of *EPO* gene is the methylation of a CpG island present in the 5' promoter and in the 5'UTR of this gene. It seems that this methylation is able to decrease the levels of transcription by two different mechanisms of control: (i) the methylation of the CpG sites in the promoter block the binding of proteins that would enhance transcription, such as HIFβ; (ii) the methylation of the 5'UTR allows the binding of proteins that repress transcription or recruits corepressors, histone deacetylases, or both (Yin and Blanchard, 2000). Also, the report of an unmethylated promoter in Hep3B, along with the hypermethylation of the

same region in different cancer cell lines and primary tumors, may reflect how this mechanism is involved in the tissue-specific expression of EPO (Steinmann et al., 2011).
The regulation of the *EPO* gene is a complex and multilayered network of mechanisms. Further studies are still required for the integrated knowledge of how these mechanisms interact and work together to the correct expression of EPO.

I.3.3. Post-transcriptional regulation of the *EPO* transcript

As mentioned before, emerging examples illustrate the importance and the diversity of post-transcriptional regulatory mechanisms and how they are responsible for a more quick and reversible response of the organisms to their environment.
Studies of the increased expression of EPO during hypoxia in Hep3B reported the cooperation of transcriptional and post-transcriptional regulatory mechanisms. The hypothesis that arose for its post-transcriptional regulation was the stabilization of the *EPO* mRNA (Goldberg et al., 1991). Also, the homology studies with the human and murine *EPO* gene granted for the recognition of several evolutionary conserved regions: the position and number of introns, the sequence of the first intron, the 5' promoter region, the 5' leader sequence that presents a conserved upstream open reading frame of 14 codons, the 3'UTR, and the amino acid sequence of the EPO protein (Shoemaker and Mitsock, 1986). Altogether, these observations led to the conclusion that, besides the influence of several transcriptional regulatory mechanisms, there is also a contribution of post-transcriptional regulation for the *EPO* transcript.
Rondon et al. (1991) described, for the first time, a *cis*-element present in the proximal region of the *EPO* 3'UTR that is able to increase *EPO* mRNA stability under hypoxia. Under hypoxic conditions there is a specific binding of a complex of proteins called *EPO* RNA binding protein (ERBP) that was initially proposed to modulate the *EPO* mRNA turnover rather then translation (Rondon et al., 1991). Latter, it was suggested that the ERBP would protect the *EPO* mRNA from endonucleolytic cleavage in a region downstream and adjacent to the ERBP binding site, increasing the steady-state mRNA levels during hypoxia (McGary et al., 1997). Another question has to do with the lack of increased ERBP levels during hypoxia, which does not correlate with the induction of the *EPO* mRNA stability in these conditions. One possibility is that the binding of ERBP to the

EPO mRNA is modulated by the heat-shock protein 70 (hsp70). Thus, during normoxia the interaction between ERBP and hsp70 would prevent the binding of this complex to *EPO* mRNA. On the contrary, during hypoxia, the hsp70 is set apart, leaving the ERBP free to interact with *EPO* mRNA hence increasing its stability (Scandurro et al., 1997). Latter on, two isoforms of a poly(C)-binding protein (PCBP), $PCBP_1$ and $PCBP_2$, were identified as part of a ribonucleoprotein complex associated with the 3'UTR of *EPO* mRNA that also seems to modulate the *EPO* mRNA stability (Czyzyk-Krzeska and Bendixen, 1999; Zhu et al., 2002). However, it is still unknown whether these proteins are actually elements of the ERBP.

Although the complete basis of this regulatory mechanism is still unclear, the *EPO* 3'UTR and 3' enhancer were proposed as a potential system for hypoxia inducible gene therapy, mainly to enhance the production of VEGF (Lee et al., 2006). Then, cloning the gene of interest along with both 3'UTR and 3' enhancer of *EPO* transcript in a plasmid will result in a more stable and controlled expression of the corresponding gene for gene therapy purposes (Choi et al., 2007).

I.3.4. EPO as a therapeutic target

As already stated, a form of rhEPO is largely used for clinical purposes. Several disorders such as chronic renal failure, cancer, acquired immunodeficiency syndrome (AIDS), and also surgical patients that develop severe anemia, are treated by administration of rhEPO. However, the cost and possible secondary effects of this treatment are of major concern and have motivated the production of safer forms of EPO.

The negative outcomes of rhEPO administration are consequence of its pro-coagulant, pro-thrombotic and vasoactive activities [for review see (Bunn, 2013; Maiese et al., 2008)]. This may contribute to increased thrombosis, mortality, progression of cancer, and cerebral ischemia (Bennett et al., 2008; Frietsch et al., 2007; Leyland-Jones, 2003; Phrommintikul et al., 2007). Despite these negative aspects of rhEPO, its application is effective in most of the patients with anemia that have received treatment.

The non-hematopoietic functions of EPO are responsible for the negative secondary effects of rhEPO administration for treatment of anemic patients, but also contribute for EPO recognition as a possible therapy for several other disorders. EPO presents cardiac

and neuronal tissue protection activities and its expression in these tissues is increased during injury. However, this endogenous increase is not sufficient to induce a complete protective effect, being necessary the administration of rhEPO or its derivatives. This has urged the development of new EPO-associated therapies for the treatment of several other disorders (Chateauvieux et al., 2011). In fact, more than 300 clinical trials involving EPO are presently registered in the National Institute of Health website (www.clinicaltrials.gov). The disorders for which the use of EPO protein is suggested include AD, cerebral malaria, retinopathy, cerebral and cardiac ischemia, and several muscle disorders (Arabpoor et al., 2012; Caprara and Grimm, 2012; Casals-Pascual et al., 2009; Lipsic et al., 2006; Scoppetta and Grassi, 2004; Undén et al., 2013).

In the search for new approaches for EPO administration, modified forms of EPO have emerged. One example, was a modified version where the peptide retains the non-hematopoietic functions of EPO but lacks the ability to induce blood cell proliferation (Arcasoy, 2008; Brines et al., 2008; Hoch et al., 2011; King et al., 2007). Also, EPO derivatives that are able to cross the blood-brain barrier more easily, and that present an increased half life, so that the administration frequency can be decreased, are being developed for the treatment of neuronal disorders (Nett et al., 2012; Zhou et al., 2011).

In summary, although EPO is a protein discovered several years ago and it has been intensively studied, there are still several questions regarding its synthesis, regulation and signaling under active investigation. These questions can contribute not only for the knowledge of EPO physiological role but also for the so expected elaboration of new and safer therapeutic strategies either by modulating EPO or other effector involved in this pathway.

I.4. Aims

uORFs are negative regulatory elements present in almost half of the human transcriptome. These elements allow a fine-tuned regulation of protein levels: under normal conditions they repress translation of the main ORF thus guaranteeing low levels of protein production, whereas under specific stress conditions, in which the protein is needed, they lead to an increased production of the corresponding protein. Thus, they constitute a level of translation regulation. Since uORFs are also related to several diseases, their study has been recently directed towards therapeutics.

The human *EPO* transcript presents a 14-codon uORF totally located at the 5' leader sequence. EPO is a multifaceted protein that induces the proliferation and differentiation of the erythroid precursor cells, regulating erythropoiesis. Moreover, it is also a neuro- and cardio-protective protein since it enhances the proliferation, differentiation and survival of cardiac and neuronal cells. EPO is mainly produced in the embryonic liver and in the adult kidney, but it is also produced in several other tissues as neurons, glial cells, lung, heart, bone marrow, among others. Furthermore, it has several layers of gene expression regulation, both at the transcriptional and post-transcriptional levels.

In the present work, our main goal was to identify and characterize the *EPO* uORF regulatory mechanism and also its biological relevance. First, we investigated whether *EPO* uORF is functional in the major EPO production tissues. Thus, we expected to dissect the molecular basis of *EPO* uORF regulation on those tissues under normal conditions. For that, we intent to test: (i) the leaky scanning and translation reinitiation efficiency; (ii) if translational inhibition by the uORF is peptide sequence-dependent; (iii) if nonsense-mediated mRNA decay (NMD) downregulates EPO expression. Additionally, we aimed to identify any interaction between the uORF and 3'UTR of the *EPO* transcript, and whether *EPO* uORF repression is overridden under stress conditions, particularly under hypoxia, and what is the underlying mechanism.

Then, considering that *EPO* uORF is functional, we expected to broaden our study and to test whether *EPO* uORF is functional in other tissues of EPO production, such as neuronal tissue.

Another goal of this work was to investigate the mechanisms through which the reininitation occurs, mainly, the role of the uORF length in this process and which eIF3 subunits are involved.

Understanding oxygen and tissue-specific regulation of EPO expression and production is of high relevance for physiology. Moreover, this knowledge is useful to design improved EPO-based therapies for several human diseases, including acute nervous system syndromes.

CHAPTER II – Translation of the human erythropoietin transcript is regulated by an upstream open reading frame in response to hypoxia

Author's note

The results contained in this chapter are submitted to publication.

Barbosa C and Romão L. Translation of the human erythropoietin transcript is regulated by an upstream open reading frame in response to hypoxia. (under review)

II.1. Abstract

Erythropoietin (EPO) is a key mediator hormone for hypoxic induction of erythropoiesis that also plays important non-hematopoietic functions. Regulation of EPO occurs at different levels, including transcription and mRNA stabilization. In this report, we show that the expression of *EPO* is also regulated at the translational level by a small upstream open reading frame (uORF) of 14 codons. As judged by comparisons of protein and mRNA levels, the uORF acts as a *cis*-regulatory element that represses translation of the EPO main ORF, in unstressed HEK293, HepG2 and REPC cells. Furthermore, we present an analysis of the translational mechanism by which the human EPO uORF affects downstream translation. Despite its conservation among mammalian species, the uORF-encoded peptide is not required for this inhibitory effect. Rather, a minority of ribosomes gain access to the *EPO* initiation codon by leaky scanning past the upstream AUG codon and the majority of ribosomes that load on the *EPO* mRNA most likely translate the uORF and some are then able to reinitiate at the downstream AUG codon. These results show that the *EPO* uORF controls synthesis of this hormone by limiting ribosomal access to the downstream main initiation codon. However, in response to hypoxia, this repression is significantly released, specifically in renal REPC cells, through a mechanism that involves processive scanning of ribosomes from the 5' end of the *EPO* transcript and enhanced ribosome bypass of the uORF. In addition, we demonstrate that hypoxia induces the phosphorylation of eukaryotic initiation factor 2α (eIF2α) and that eIF2α phosphorylation significantly increases translation of the *EPO* mRNA translation, in response to hypoxia, in renal cells. These findings provide a framework for understanding that production of high levels of EPO induced by hypoxia also involves regulation at the translational level.

II.2. Introduction

Regulation of mRNA translation is a key mechanism by which cells and organisms can rapidly change their gene expression patterns in response to extra- and intracellular stimuli. Translational control can occur on a global basis by modifications of the basic translation machinery, or selectively target defined subsets of messenger RNAs (mRNAs)

to maintain the synthesis of certain proteins required either for the stress response or to aid recovery from the stress. These pathways are evolutionary conserved and have been shown to significantly impact translation in organisms as diverse as yeast and humans. In many cases, features in the 5' leader sequences of the corresponding mRNAs, such as internal ribosome entry sites (IRESs) and/or regulatory upstream open reading frames (uORFs), are important for them to evade global repression of translation (Sonenberg and Hinnebusch, 2009).

uORFs are regulatory *cis*-acting elements present in the 5' leader sequence of a transcript that are spread among different species and throughout the genome, but their prevalence has been difficult to calculate (Mignone et al., 2002). The most recent studies estimate that about 49% of the human transcripts contains at least one uORF (Calvo et al., 2009), being conspicuously common in certain classes of genes, including oncogenes and genes involved in the control of cellular growth and differentiation (Morris and Geballe, 2000; Wethmar et al., 2010a). The fact that mutations that introduce or disrupt a uORF can cause human diseases illustrates their role in translational regulation (Cazzola and Skoda, 2000; Chatterjee and Pal, 2009).

For a uORF to function as a translational regulatory element, its initiation codon (upstream AUG; uAUG) must be recognized, at least at certain times, by the scanning 40S ribosomal subunit and associated initiation factors (Hernández et al., 2010). When the uORF recognition is regulated by a so-called leaky scanning mechanism, ribosomes either scan through the uAUG codon or recognize it, initiating translation. Indeed, the recognition of an AUG can be affected by its context, the AUG proximity to the cap site and the presence of nearby secondary structures. The optimal context is GCC(A/G)CCAUGG, being the -3 and +4 functionally the most important positions (Kozak, 2002). In the case the uORF is recognized and translated by a scanning ribosome, multiple alternative fates are available to the ribosome: the ribosome may (i) terminate and leave the mRNA, resulting in downregulation of translation of the downstream main ORF, (ii) translate the uORF and stall during either the elongation or termination phase of uORF translation, creating a blockade to additional ribosome scanning, (iii) terminate and reinitiate (Meijer and Thomas, 2002; Poyry et al., 2004). When the option is for the ribosome to remain associated with the mRNA, it continues scanning, and reinitiates further downstream, at either a proximal or distal AUG codon. The potential of a

ribosome to reinitiate further downstream depends on a number of factors, including the length of the uORF and the time it took to translate the uORF (Kozak, 2001; Rajkowitsch et al., 2004). It has been shown that it is not so much the length *per se* that is the critical parameter, but rather the time taken for the ribosome to translate the uORF (Poyry et al., 2004). Indeed, a uORF that is short enough to be permissive for translation reinitiation becomes nonpermissive if it has a pseudoknot structure that causes ribosome pausing (Kontos et al., 2001; Kozak, 2001; Poyry et al., 2004). The reason why the uORF translation needs to be completed rapidly is related to the initiation factors necessary to remain associated with the 40S subunit to promote 40S scanning from the uORF termination codon to the downstream initiation site (Poyry et al., 2004). However, other initiation factors have to be acquired *de novo*. One initiation factor that has to be reacquired is eIF2 in the form of an eIF2/GTP/Met-$tRNA_i$ ternary complex (Hinnebusch, 1997), because the Met-$tRNA_i$ in the ternary complex associated with the 40S subunit as it scans to the uORF initiation codon has been used to initiate the uORF translation (Kozak, 2005; Poyry et al., 2004; Sachs and Geballe, 2006).

When mammalian cells encounter stress conditions such as pathogenic infection, chemical exposure, nutrient deprivation and hypoxia, and even during differentiation and development, a family of protein kinases is activated to phosphorylate eIF2α. Phosphorylation of the α subunit of eIF2 on Ser 51 prevents the exchange of GDP for GTP by sequestering eIF2B, lowering the available pool of eIF2/GTP that is required for binding of initiator tRNA to the small ribosomal subunit, and thus repressing protein synthesis (Sonenberg and Hinnebusch, 2009). Concomitant with the general inhibition of translation, phosphorylation of eIF2α selectively promotes translational upregulation of a subset of mRNAs. Such mRNAs include those containing uORFs, such as the activating transcription factor (*ATF4*) mRNA. Indeed, the 5’ leader sequence of *ATF4* mRNA contains two uORFs that are conserved among species and regulate translation. Under basal conditions, the levels of eIF2/GTP are high and thus, the ribosomes translate the first 5’ proximal uORF. Following translation termination of the first uORF, the small ribosomal subunit resumes scanning recharges the eIF2/GTP/Met-$tRNA_i$ ternary complex and reinitiates translation at the uORF2, which precludes translation of the ATF4 ORF, as the uORF2 overlaps the ATF4 ORF. In contrast, under conditions where eIF2α is phosphorylated, initiation of the second uORF is less likely, as there is less chance of the

scanning ribosomal subunit recruiting the ternary complex required for start codon recognition, and thereby initiation of translation occurs at the main ORF (Lu et al., 2004; Vattem and Wek, 2004).

uORFs have also the potential to affect gene expression by altering mRNA stability. The similarity in the cistronic organization of a uORF-containing mRNA to that of a mRNA containing a nonsense mutation has suggested the potential of a uORF-bearing mRNA to trigger nonsense-mediated mRNA decay (NMD). Indeed, it has been shown that NMD functions to control the physiologic levels of transcripts bearing uORFs (Mendell et al., 2004). However, not all mRNAs that contain uORFs are targets for NMD (Lee et al., 2009; Yaman et al., 2003; Zhou et al., 2008a), and the critical determinants of sensitivity remain to be fully appreciated (Barbosa et al., 2013).

The human erythropoietin (EPO) is a circulating 34,4-kDa glycoprotein hormone that controls erythropoiesis by stimulating the proliferation of erythroid precursors (Ebert and Bunn, 1999; Jelkmann, 1992; Mole and Ratcliffe, 2007). Indeed, its major action is the prevention of apoptosis in EPO-dependent colony-forming unit-erythroid cells and erythroblasts that have not begun hemoglobin synthesis. Expression of the *EPO* gene is tightly controlled and in the adult organism, kidneys produce around 90% of circulating EPO, being its production markedly up-regulated by hypoxia (Jelkmann, 1992). In the adult, liver *EPO* mRNA levels, which are very difficult to detect at baseline, rise substantially following stimulation with moderate to severe hypoxia (Bunn, 2013). Aside from the kidney and liver as the two major sources of synthesis, *EPO* mRNA expression has also been detected in the brain (neurons and glial cells), lung, heart, bone marrow, spleen, hair follicles, and the reproductive tract (Dame et al., 1998; Fandrey and Bunn, 1993; Ghezzi and Brines, 2004; Hoch et al., 2011; Weidemann and Johnson, 2009; Yasuda et al., 1998). EPO synthesized in these organs appears to act locally, modulating, for example, regional angiogenesis and cellular viability and does not seem to contribute to erythropoiesis (Gassmann and Soliz, 2009; Maiese et al., 2008).

As above referred, EPO expression under normoxic conditions is low, but increases during exposure to hypoxia in the cells of the kidney cortex and outer medulla (Besarab et al., 2009; Chin et al., 2000), and also in the two Hep3B and HepG2 cell lines derived from liver tumors (Cohen et al., 2004), as well as in the REPC (renal EPO-producing) cells (Frede et al., 2011), an human kidney cell line recently established from an explanted

human kidney that exhibits correct *EPO* gene expression and release of the EPO protein in an oxygen-dependent manner (Bushuev et al., 2006). Undeniably, hypoxia is the primary physiological stimulus for EPO production, which, depending on the hypoxic condition, increases serum EPO levels up to several hundred-fold (Ebert and Bunn, 1999). Hypoxia inducible factor 1 (HIF1) is the transcriptional activator responsible for the hypoxic induction of EPO that binds to the hypoxia-responsive element located in the 3' untranslated region (UTR) of the *EPO* gene augmenting the *EPO* transcriptional rate (Noguchi et al., 2008; Wang et al., 1995). Together with HIF2α, HIF1α facilitates oxygen delivery and cellular adaptation to hypoxia by stimulating multiple biological processes, such as erythropoiesis, angiogenesis, and anaerobic glucose metabolism (Semenza, 2001). Under normoxia, all three known HIF α-subunits, HIF1α, HIF2α, and HIF3α, are targeted for rapid proteasomal degradation by the von Hippel-Lindau tumor suppressor (VHL), which acts as the substrate recognition component of an E3 ubiquitin ligase complex (Maxwell et al., 1999; Maynard et al., 2003).

Although the notorious effect of the many transcriptional mechanisms involved in *EPO* gene expression regulation, there is also an important contribution of post-transcriptional mechanisms that are less characterized (Goldberg et al., 1991). For example, different studies have reported the binding of proteins to the 3'UTR of *EPO* transcript that increase the mRNA stability, such as protein kinase C-α (PKC) (Ohigashi et al., 1999), poly(C) binding protein (PCBP) (Zhu et al., 2002), and *EPO* mRNA binding protein (ERBP) (McGary et al., 1997). Alignment of the human and mouse *EPO* 5' leader sequences revealed high identity and the presence of a 14 codons uORF located upstream of the main AUG. These observations led us to investigate the potential role of this naturally occurring uORF in the translational control of the human *EPO* expression. In this study, we report that the single uORF located in the human *EPO* mRNA inhibits translation in unstressed cells. However, this repression is significantly released by hypoxia in renal cells, via eIF2α phosphorylation. These findings provide a framework for understanding that production of high levels of EPO induced by hypoxia also involves regulation at the translational level.

II.3. Materials and Methods

II.3.1. Plasmid constructs

The plasmid pGL2-enhancer (Promega) that encodes for the firefly luciferase (FLuc) was digested with BglII/HindIII, where it was inserted the BglII/HindIII digested CMV promoter sequence from pcDNA3.1/hygro+ (Invitrogen), originating the so-called pGL2-Luc construct. Then, the 181 base pairs fragment corresponding to the 5'UTR of the human *EPO* transcript was inserted into the HindIII/XbaI sites of the pGL2-Luc plasmid, originating the pGL2-WT construct. The fragment corresponding to the 5'UTR of the human *EPO* transcript was previously obtained by overlap-extension PCR. For that, two PCR products were obtained: one corresponding to the human *EPO* 5'UTR amplified from human genomic DNA with the primer #1 (with the HindIII linker; Table II.1.) and the overlapping reverse primer #3 (Table II.1.), and another corresponding to the amplification of the pGL2-Luc plasmid with the overlapping primer #4 (Table II.1.) and the reverse flanking primer #2 that produce a 160 base pairs fragment beginning at the FLuc AUG codon through downstream (Table II.1.). Then, the flanking primers #1 and #2 and the two PCR amplified products were used to amplify the final overlap-extension fragment encompassing the human *EPO* 5'UTR and the 5' part of the FLuc cistron, which was further digested with HindIII and XbaI enzymes (the fragment sequence includes a XbaI restriction site) before ligation to the pGL2-Luc vector to create the pGL2-WT construct.

The pGL2-no_uAUG variant, carrying a mutation at the uORF AUG codon (ATG→TTG), was created by replacing the native HindIII/XbaI fragment by the corresponding fragment carrying the ATG→TTG mutation. For that, the same overlap-extension approach was used, in which the two first PCR reactions were performed to amplify fragments from the pGL2-WT plasmid with primers #1 and #5, (reverse primer #5 carries the mutation; Table II.1.) and with primers #6 and #2 (forward primer #6 carries the mutation; Table II.1.), respectively. Then, primers #1 and #2 where used to amplify the overlapped fragment, using the previous fragments as template. The resulting DNA fragment was digested with HindIII/XbaI, and then ligated to the pGL2-WT plasmid previously digested with the same enzymes.

The pGL2-Luc_fusion1 carrying a mutation of the uORF stop codon in *cis* with a deletion of one nucleotide four nucleotides downstream the stop codon (TGAgggac→AGAggg-c), so that both AUG codons (the uAUG and the main AUG) are in frame, the pGL2-Luc_fusion2 carrying the same mutations as pGL2-Luc_fusion1 and a mutation (ATG→TTG) of the main AUG, the pGL2-no_uSTOP carrying a mutation (TGA→AGA) at the uORF STOP codon, and the pGL2-optimal_uAUG carrying a mutation of the uAUG sequence context (gggAUGa→gccAUGg), were all created by site-directed mutagenesis, using the pGL2-WT plasmid as template and the mutagenic primers #7 to #14, respectively (Table II.1.).

The pGL2-frameshift construct was obtained, by site-directed mutagenesis, by insertion of one *A* nucleotide at the beginning of the uORF (5'-ATGA*A*GG…-3') and deletion of one *T* nucleotide at the 3' end (5'-…GG*T*CGCTGA-3'), using the primers #15 to #18 (Table II.1.), and the pGL2-WT as template.

The sequence corresponding to the 3'UTR of the human *EPO* transcript was inserted into the EcoRI/PflMI sites of the pGL2-Luc, pGL2-WT and the pGL2-no_uAUG constructs to obtain the pGL2-Luc-3'UTR, pGL2-WT-3'UTR and pGL2-no_AUG-3'UTR constructs, respectively. For this purpose, overlap-extension PCR was performed as above, with the flanking primers #19 and #20 (primer #20 with PflMI linker; Table II.1.), using as templates, the PCR amplification of the 3' part of the FLuc cistron (450 base pairs fragment located upstream the FLuc stop codon, which encompasses a EcoRI site) from the pGL2-Luc construct with primers #19 and #22 (Table II.1.), and the PCR amplification of the 3'UTR fragment of the human *EPO* gene with primers #20 and #21 (Table II.1.).

The dicistronic constructs carrying both *Renilla* and firefly luciferase (RLuc and FLuc, respectively) cistrons were derived from the psiRF vector previously described (Tahiri-Alaoui et al., 2009). In order to prevent ribosome read-through, the sequence encompassing a stable hairpin structure was PCR amplified from the plasmid p53 "construct A" (Candeias et al., 2006), with the primers #23 and #24. This PCR product was digested with XhoI and cloned into the XhoI site of the psiRF vector. The resulting construct was named "RLuc-empty". To obtain the "RLuc-β-globin 5'UTR", the same overlap-extension approach described above was used. In this case, the two first PCR reactions were performed to amplify the following fragments: the human *β-globin* 5'UTR with the primers #25 (with the XmaI linker) and #26 (Table II.1.) and the 663 base

pairs fragment located downstream of the FLuc AUG codon from the “RLuc-empty” plasmid encompassing one AccI site, with the primers #27 and #28 (Table II.1.). Then, primers #25 and #28 where used to amplify the two overlapped fragments. The resulting DNA fragment was digested with XmaI/AccI, and then ligated to the “RLuc-empty” plasmid previously digested with the same enzymes. Using the same approach, the “RLuc-c-myc_IRES” construct was generated where the two first PCR reactions were performed to amplify the *c-myc* IRES (Stoneley et al., 1998) with the primers #29 (with the EcoRI linker) and #30 (Table II.1.), and to amplify a 663 base pairs fragment from the AUG codon through downstream into the FLuc coding sequence of the “RLuc-empty” construct, with the primers #31 and #28 (Table II.1.). Then, primers #29 and #28 where used to amplify the two overlapped fragments. The resulting DNA fragment was digested with EcoRI/AccI, and then ligated to the “RLuc-empty” plasmid previously digested with the same enzymes.

To obtain the “RLuc-WT” and the “RLuc-no_uAUG” dicistronic constructs, the “RLuc-empty” plasmid was digested with BglII/PmeI and the *Renilla* luciferase cistron sequence along with the SV40 promoter sequence and the hairpin structure was inserted into the BglII/PmeI sites of pGL2-WT and pGL2-no_uAUG plasmids, creating, respectively, the RLuc-WT and the RLuc-no_uAUG constructs.

II.3.2. Cell culture and plasmid transfection

HEK293 cells were grown in Dulbecco’s modified Eagle’s medium (DMEM) supplemented with 10% fetal bovine serum. HepG2 and REPC cells were grown in Roswell Park Memorial Institute (RPMI) 1640 medium supplemented with 10% fetal bovine serum (Gibco). Cells were grown at 37°C in humidified incubator containing 5% CO_2. Transient transfections were performed using Lipofectamine 2000 Transfection Reagent (Invitrogen), following the manufacturer’s instructions, in 35-mm plates. Cells were co-transfected with 750 ng of the test DNA construct corresponding to the pGL2-Luc, pGL2-WT, or its derivative plasmids, and 500 ng of the pRL-TK plasmid (Promega), which encodes *Renilla* luciferase as an internal control, and, then, harvested after 24h. Dicistronic plasmids were single transfected at the same conditions, but using 1 µg of test DNA construct. To mimic hypoxia, 6h post-transfection, the cultures were changed

to fresh medium supplemented with 200 µM $CoCl_2$ (Sigma-Aldrich). To induce nutrient starvation, 6h post-transfection, the cultures were changed to fresh medium supplemented with 10% of dialyzed fetal bovine serum (Gibco) and then, cells were harvested 24h post-treatment. Cells were treated with 1 µM thapsigargin (Sigma) to activate eIF2α kinases and induce eIF2α phosphorylation.

II.3.3. siRNA transfection

The siRNA oligonucleotides used for transfections [Luciferase (59-CGUACGCGGAAUACUUCGA-39) and UPF1 (59-UUACCGCGUUCUGUGUGAA-39)] were purchased as annealed, ready-to-use duplexes from MWG. HepG2 cells cultured in 35-mm plates were transfected using 200 pmol of each oligonucleotide and 10 µl of Lipofectamine TM RNAiMAX transfection reagent (Invitrogen), following the reverse-transfection protocol indicated by the manufacturer. Seventy-two hours later, cells were collected for RNA and protein extracts.

II.3.4. SDS-PAGE and Western blotting

Protein lysates were resolved, according to standard protocols, in 10% SDS-PAGE and transferred to PVDF membranes (Bio-Rad). Membranes were probed using mouse monoclonal anti-α-tubulin (Sigma) at 1:10000 dilution, goat polyclonal anti-hUPF1 (Bethyl Labs) at 1:500 dilution, rabbit polyclonal anti-eIF2α (Cell Signaling) at 1:500 dilution, rabbit polyclonal anti-phospho Ser 51 eIF2α (Life Technologies) at 1:500 dilution or rabbit polyclonal anti-HIF1α (BD Biosciences) at 1:500. Detection was carried out using secondary peroxidase-conjugated anti-mouse IgG (Bio-Rad), anti-rabbit IgG (Bio-Rad) or anti-goat IgG (Sigma) antibodies followed by chemioluminescence.

II.3.5. Luminometry assay

Lysis was performed in all cell lines with Passive Lysis Buffer (Promega). The cell lysates were used to determine luciferase activity with the Dual-Luciferase Reporter Assay System (Promega) and a Lucy 2 luminometer (Anthos Labtec), according to the manufacturer's standard protocol. One µg of extract was assayed for firefly and *Renilla*

luciferase activities. Ratio is the unit of firefly luciferase after normalized with *Renilla* luciferase, and each value was derived from three independent experiments.

II.3.6. RNA isolation

Total RNA from transfected cells was isolated using the Nucleospin RNA extraction II kit (Marcherey-Nagel), following the manufacturer's instructions. Then, all RNA samples were treated with RNase-free DNase I (Ambion) and purified by phenol:chloroform extraction.

II.3.7. Reverse transcription-quantitative PCR (RT-qPCR)

Synthesis of cDNA was carried out using 1μg of total RNA and Superscript II Reverse Transcriptase (Invitrogen), according to the manufacturer's instructions. Real-time PCR was performed in ABI Prism 7000 Sequence Detection System, using SybrGreen Master Mix (Applied Biosystems). Primers specific for the firefly luciferase cDNA (primers #32 and #33; Table II.1.) and *Renilla* luciferase cDNA (primers #34 and #35; Table II.1.), were designed using the ABI Primer Express software. Primers specific for human *EPO* cDNA (primers #36 and #37; Table II.1.) were as previously described (Frede et al., 2011). Primers specific for human *HFE* hemochromatosis (primers #38 and #39; Table II.1.) and G protein pathway suppressor 1 (*GPS1*) (primers #40 and #41; Table II.1.) cDNAs were previously described (Martins et al., 2012). Quantification was performed using the relative standard curve method (ΔΔCt, Applied Biosystems). The following cycling parameters were used: 10 min at 95°C and 40 cycles of 15 sec at 95°C and 1 min at 61°C. Technical triplicates from three to four independent experiments were assessed in all cases.

II.3.8. Statistical analysis

Results are expressed as mean ± standard deviation. Student's *t* test was used for estimation of statistical significance. Significance for statistical analysis was defined as a $p < 0.05$.

Table II.1. DNA oligonucleotides used in the current work.

Primer	Sequence (5' → 3')
#1	CCCAAGCTTCCCGGAGCCGGACCGG
#2	CGTACGTGATGTTCACCTC
#3	GCCAGGCGCGGAGATGGAAGACGCCAAAAACATAAAG
#4	CTTTATGTTTTTGGCGTCTTCCATCTCCGCGCCTGGC
#5	CTTCCCGGGTTGAGGGC
#6	GCCCTGAACCCGGGAAG
#7	GTCGCAGAGGGCCCCGG
#8	CCGGGGCCCTCTGCGAC
#9	GCGGAGTTGGAAGACGCC
#10	GTCTTCCAACTCCGCGCC
#11	CAGGTCGCAGAGGGACCCCGGCCAC
#12	CTGGCCGGGGTCCCTCTGCGACCTG
#13	CCGCCGAGCTTCCCGGCCATGGGGCCCCCGG
#14	CCGGGGGCCCCCATGGCGGGAAGCTCGGCGG
#15	GAGCTTCCCGGGATGAAGGGCCCCCGG
#16	CCGGGGGCCCTTCATCCCGGGAAGCTC
#17	GCGCCCCAGGCGCTGAG
#18	CTCAGCGCCTGGGGGCGC
#19	GTAAACAATCCGGAAGCGACC
#20	GGGCCATAGGTTGGTTGGTGGTTTCAGTTCTTGTC
#21	GTCCAAATTGTAACCAGGTGTGTCCACCTGG
#22	GTGGACACACCTGGTTACAATTTGGACTTTCCGCC
#23	CCGCTCGAGCGGGGTACCAATGACGCGCGC
#24	GAATTCTGCAGTCGACGGTACC
#25	TCCCCCCGGGGGGAACATTTGCTTCTGACACAAC
#26	CATCGGCCTAGGTGTCTGTTTGAGGT
#27	ACAGACACCATGGCCGATGCTAAGAACA
#28	GTGAGAGAAGCGCACACAG
#29	GGAATTCCAATTCCAGCGAGAGGCAGAG
#30	TAGCATCGGCCATCGTCTAAGCAGCTGCAAGGAGA
#31	GCAGCTGCTTAGACGATGGCCGATGCTAAGAACA
#32	CAACTGCATAAGGCTATGAAGAGA
#33	ATTTGTATTCAGCCCATATCGTTT
#34	AACGCGGCCTCTTCTTATTT
#35	ACCAGATTTGCCTGATTTGC
#36	TGGGAGCCCAGAAGGAAGCCA
#37	TGGTCATCTGTCCCCTGTCCTGC
#38	AAGCATTCTGTCTTGAAGGGCA
#39	CTGAGCTGTATATGGTATCCTGAAGC
#40	CGAGTCCAAGTACGCCTCATG
#41	GGTTGTCCTTCATCTCGTCCA

II.4. Results

II.4.1. The human *EPO* 5' leader sequence comprises a conserved uORF

It has been shown that the 5' leader sequence of the human *EPO* transcript has a significant homology with the murine *EPO* 5'UTR, with both sequences having a high percentage of GC content (Shoemaker and Mitsock, 1986). These authors also reported the presence of a 14-codons-ORF located upstream of the main AUG in both human and murine *EPO* transcripts. Here, Figure II.1.A shows a broader alignment of the *EPO* 5' leader sequences from human, chimpanzee, gorilla, orangutan, common marmoset, mouse, and rat, which exhibits the high degree of similarity between all sequences and the conservation of the 14-codons uORF. These uORFs have a Kozak match of G/A at -3 of the A(+1)UG. In addition, the position of the *EPO* uORF relatively to the main AUG shows a significant similarity among species, being the intercistronic region of 22 or 25 nucleotides length (Figure II.1.A). The alignment of the amino acid sequences of the *EPO*

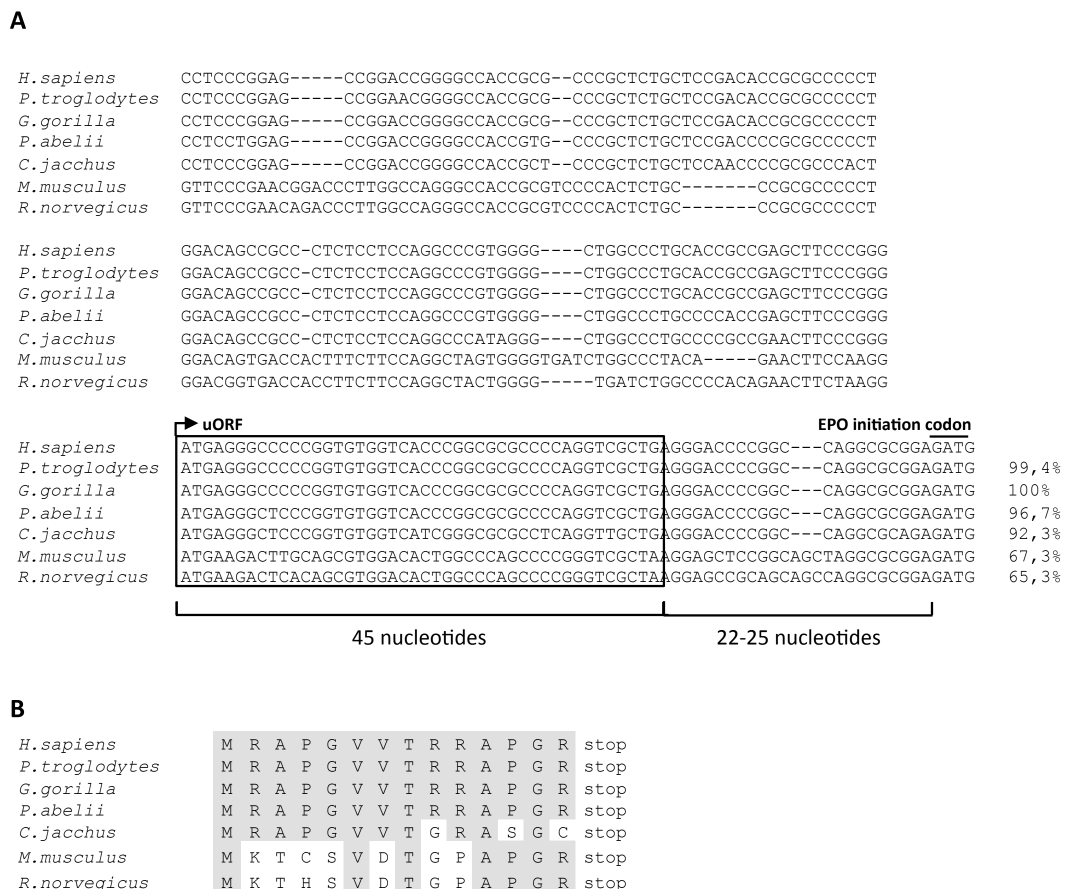


Figure II.1. The 5' leader sequence of the *EPO* transcript includes a highly conserved uORF.

(A) Nucleotide sequence alignment of the human (*Homo sapiens*), chimpanzee (*Pan troglodytes*), gorilla (*Gorilla gorilla*), orangutan (*Pongo abelii*), common marmoset (*Callithrix jacchus*), mouse (*Mus musculus*) and rat (*Rattus norvegicus*) *EPO* mRNA 5' leader regions. The uORF sequences of the different species are framed, where the arrow indicates position of the upstream initiation AUG codon (in grey). The uORF termination codon (UGA or UAA), as well as the main *EPO* AUG codon (EPO initiation codon), is also shown in grey. The uORF and the intercistronic region lengths (in nucleotides) are indicated below. On the right, it is indicated the percentage of homology relatively to the human 5' leader sequence. **(B)** Amino acid sequence alignment of the uORF in the human, chimpanzee, gorilla, orangutan, common marmoset, mouse and rat *EPO* transcript. The conserved amino acids are indicated in grey.

uORF in these species also shows a high degree of similarity (Figure II.1.B). The conservation of the *EPO* uORF among species may reflect an important evolutionary selection pressure, and may suggest a potential regulatory function in EPO expression. These observations directed us to investigate the role of the human *EPO* uORF in the translational control of the downstream main ORF.

II.4.2. The *EPO* uORF represses translation of a downstream main ORF

To determine the importance of the human *EPO* uORF in modulating translation efficiency of the downstream main ORF, the 181 base pairs sequence corresponding to the intact human *EPO* 5' leader sequence was cloned into the pGL2 expression vector, flanking the FLuc reporter gene to create the pGL2-WT construct (Figure II.2.A). In addition, the *EPO* uORF was disrupted by site directed mutagenesis of the uAUG (ATG→TTG), using the previous pGL2-WT construct as template, originating the pGL2-no_uAUG construct (Figure II.2.A). Expression of each of these reporter gene constructs was studied after transient transfection into a panel of cell lines – human embryonic kidney 293 (HEK293), human hepatoma (HepG2), and human kidney REPC (renal EPO-producing cells) cells. For that, cellular extracts were prepared and assayed for luciferase activity and total RNA was isolated to quantify the relative luciferase mRNA levels by RT-qPCR (Figure II.2.B). FLuc activity of each construct was normalized to the activity units from RLuc expressed from the co-transfected pRL-TK plasmid. The relative luciferase activity was compared to that of the empty pGL2-Luc vector (Figure II.2.A), arbitrary defined as 1 (Figure II.2.B). Results show that in all cell lines studied, the human *EPO* 5' leader sequence with the intact uORF induces a 3-fold repression of translation of the reporter transcript, when compared with the relative luciferase activity from the pGL2-no_uAUG construct without uORF, whereas the relative luciferase mRNA levels are not

affected (Figure II.2.B). Thus, the intact *EPO* uORF induces a repression of gene expression at the translational level.

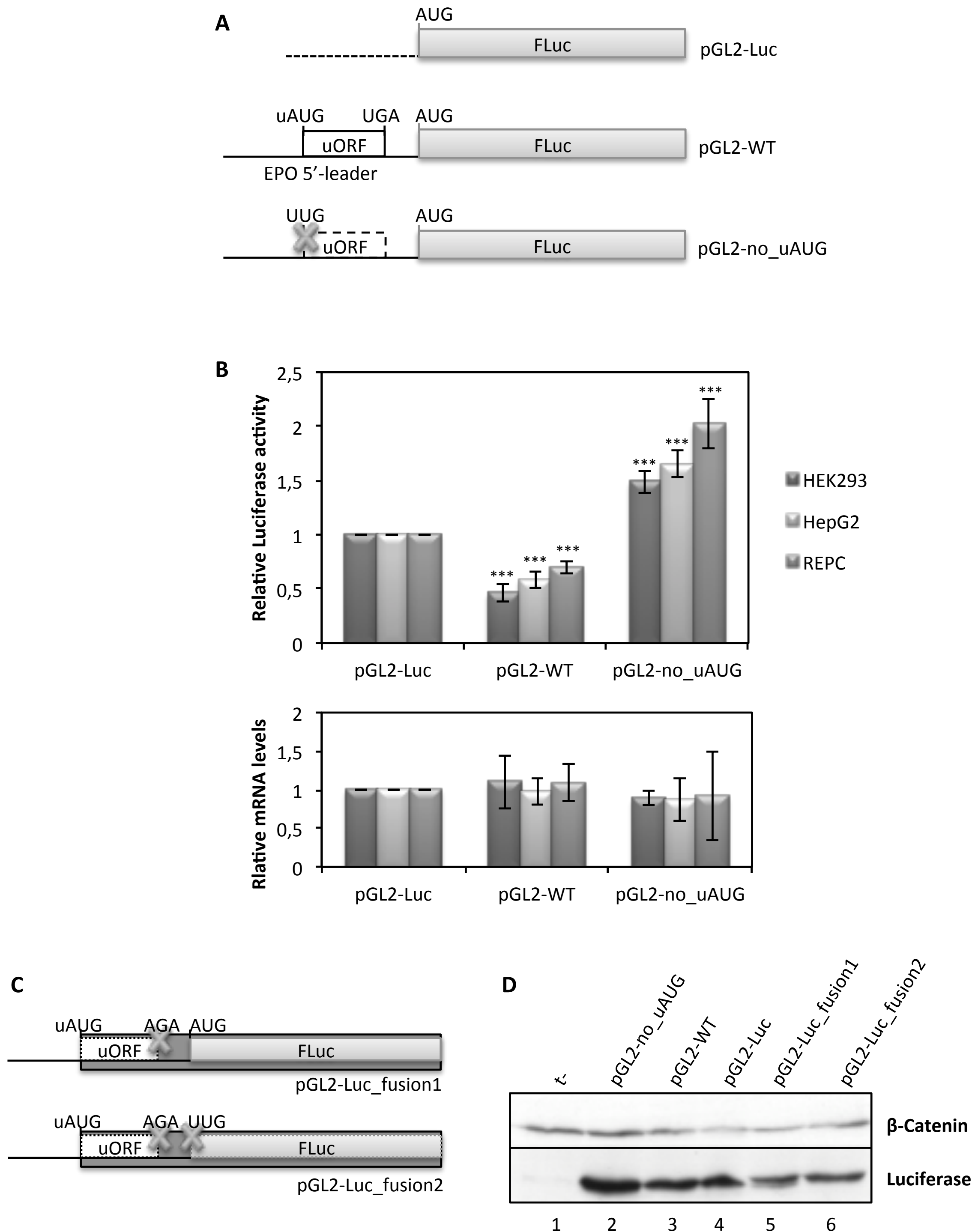


Figure II.2. The *EPO* uORF represses translation of the downstream main ORF.
(A) Schematic representation of reporter constructs. The human *EPO* 5' leader sequence encompassing its uORF (open box) with the intact initiation (uAUG) and termination (UGA) codons, was cloned into the empty vector (pGL2-Luc), upstream of the firefly luciferase coding region (FLuc; grey boxes) to create the pGL2-WT construct. In the pGL2-no_uAUG construct, the uORF initiation codon is mutated (AUG→UUG)

(the cross represent the point mutation and the dashed lined box represent the non-functional uORF). **(B)** The *EPO* 5' leader sequence represses protein expression of the downstream reporter. HEK293, HepG2 and REPC cells were transiently co-transfected with each one of the constructs described in (A) and with the pRL-TK plasmid encoding the *Renilla* luciferase (RLuc). Cells were lysed twenty-four hours later and the luciferase activity was measured by luminometry assays. FLuc activity values were normalized to RLuc activity to control for transfection efficiency. Relative luciferase activity of the pGL2-Luc was defined as one. In parallel, the luciferase mRNA levels were quantified by RT-qPCR. The FLuc mRNA levels were normalized to those of the RLuc mRNA and analyzed by the ΔΔCt method. The relative pGL2-Luc mRNA levels were also defined as one. Average values and standard deviation (SD) of three independent experiments are shown. Statistical analysis was performed using Student's *t* test (unpaired, two tailed); (∗) p<0.05; (∗∗) p<0.01; (∗∗∗) p<0.001. **(C)** Schematic representation of additional reporter constructs. The uORF stop codon was mutated in *cis* with a deletion of one nucleotide four nucleotides downstream the stop codon (TGAgggac→AGAggg-c), so that both initiation codons (uAUG and AUG) are in frame in the pGL2-Luc_fusion1 construct. This construct encodes a fusion protein represented by a darker box and the native luciferase protein if leaky scanning occurs. The AUG of the firefly luciferase coding region was mutated (AUG→UUG) in the pGL2-Luc_fusion1 construct to produce the pGL2-Luc_fusion2 construct; this construct exclusively encodes the fusion protein represented by the darker box. Crosses represent the point mutations. **(D)** Translation initiation can occur at the *EPO* uAUG. The constructs specified above each lane were transiently transfected in HEK293 cells. Twenty-four hours later, lysates were prepared and analyzed by Western blot. Immunoblotting was performed by using a firefly luciferase specific antibody and a human β-catenin specific antibody as a control for variations in protein loading.

To confirm that the uAUG of the *EPO* uORF is recognized by the ribosome, we cloned a construct in which the *EPO* uORF was fused in-frame to the luciferase ORF. This was achieved by site-directed mutagenesis of the pGL2-WT construct to introduce a mutation at the uORF stop codon in *cis* with a base pair deletion four nucleotides downstream the stop codon (TGAgggac→AGAggg-c) (see Materials and Methods; pGL2-Luc_fusion1 construct; Figure II.2.C). As a control for the presence of the extended form of the luciferase protein, another construct was cloned derived from the pGL2-Luc_fusion1 construct by mutating the FLuc main AUG (ATG→TTG) (pGL2-Luc_fusion2 construct; Figure II.2.C). These constructs, as well as pGL2-Luc, pGL2-WT and pGL2-no_uAUG constructs were transiently transfected into HEK293 cells. Twenty-four hours later, cell extracts were purified and analyzed by Western blotting using a specific antibody that recognizes firefly luciferase, using the detection of β-catenin as a loading control (Figure II.2.D). As shown in Figure II.2.D (lane 5), pGL2-Luc_fusion1 construct expresses two different proteins: one corresponding to the FLuc protein, as it presents the same molecular weight as that one expressed from the pGL2-Luc (lane 5 *versus* lane 4), and another protein with higher molecular weight that corresponds to the uORF-Luc fusion protein as it shows the same molecular weight as that one expressed from the

pGL2-Luc_fusion2 construct (lane 5 *versus* lane 6). In conclusion, the slightly larger band detected in lanes 5 and 6 (Figure II.2.D) corresponds to the fusion protein, while the other lanes (lanes 2 to 4; Figure II.2.D) only show a smaller band that corresponds to the native form of the FLuc protein. These data unequivocally demonstrate that indeed the human *EPO* uORF is recognized by the ribosome and translated and thus, it is functional.

II.4.3. Both translation reinitiation and uAUG leaky scanning are involved in the translational initiation at the main AUG codon

Since the AUG codon of the *EPO* uORF is in a good, but not optimal, context for initiation (gggAUGa), we expected that some ribosomes that load onto the *EPO* mRNA would initiate at the uAUG codon, but others could leak past the uAUG codon and initiate at the main AUG. Nevertheless, a few ribosomes that translate the uORF may reinitiate at the main AUG codon. To evaluate these possibilities we first mutated the stop codon of the uORF (TGA→AGA; pGL2-no_uSTOP construct), creating an extended uORF that terminates at the next in-frame stop codon, 83 nucleotides downstream from the FLuc initiation codon (pGL2-no_uSTOP construct; Figure II.3.A). This mutation completely abrogates the possibility that FLuc can be made by reinitiation after translation of the uORF, giving the possibility to evaluate the efficiency of ribosome leaky scanning. In addition, we mutated in the pGL2-WT vector, the context of the uAUG codon (gggAUGa→gccAUGg), to obtain the pGL2-optimal_uAUG construct (Figure II.3.A) with a uAUG sequence context shown by Kozak to yield maximum initiation frequency in higher eukaryotes (Kozak, 1997; Loughran et al., 2012; Wang and Rothnagel, 2004). In this case, the majority of the ribosomes that load on the *EPO* mRNA 5' leader sequence are unable to leak past the uAUG codon and most likely they translate the uORF and may reinitiate at the downstream AUG codon. To estimate the relative contribution of leaky scanning and translation reinitiation, HEK293, HepG2 and REPC cells were transiently transfected with the pGL2-no_uSTOP or pGL2-optimal_uAUG constructs, or with the pGL2-WT construct, and translational efficiencies were monitored by dual luciferase assays, as before. Results were compared to those obtained from the pGL2-WT construct (Figure II.3.B). As shown in Figure II.3.B, mutation of the uORF stop codon (pGL2-no_uSTOP construct) reduces relative luciferase activity to approximately 25% of that of the pGL2-

WT construct, suggesting that the percentage of ribosomes that leak past the uORF is low and thus translation of the main ORF mostly occur by reinitiation of the ribosomes after translation termination of the uORF. In fact, the analysis of the pGL2-optimal_uAUG expression allowed us to observe that translation reinitiation at the main ORF can account for about 60% of relative luciferase activity, in comparison to the relative luciferase activity of the pGL2-WT construct (Figure II.3.B). Our data demonstrate that the minority of ribosomes gain access to the main initiation codon by leaky scanning past the uAUG codon, while the majority of ribosomes by reinitiating translation after having translated the uORF.

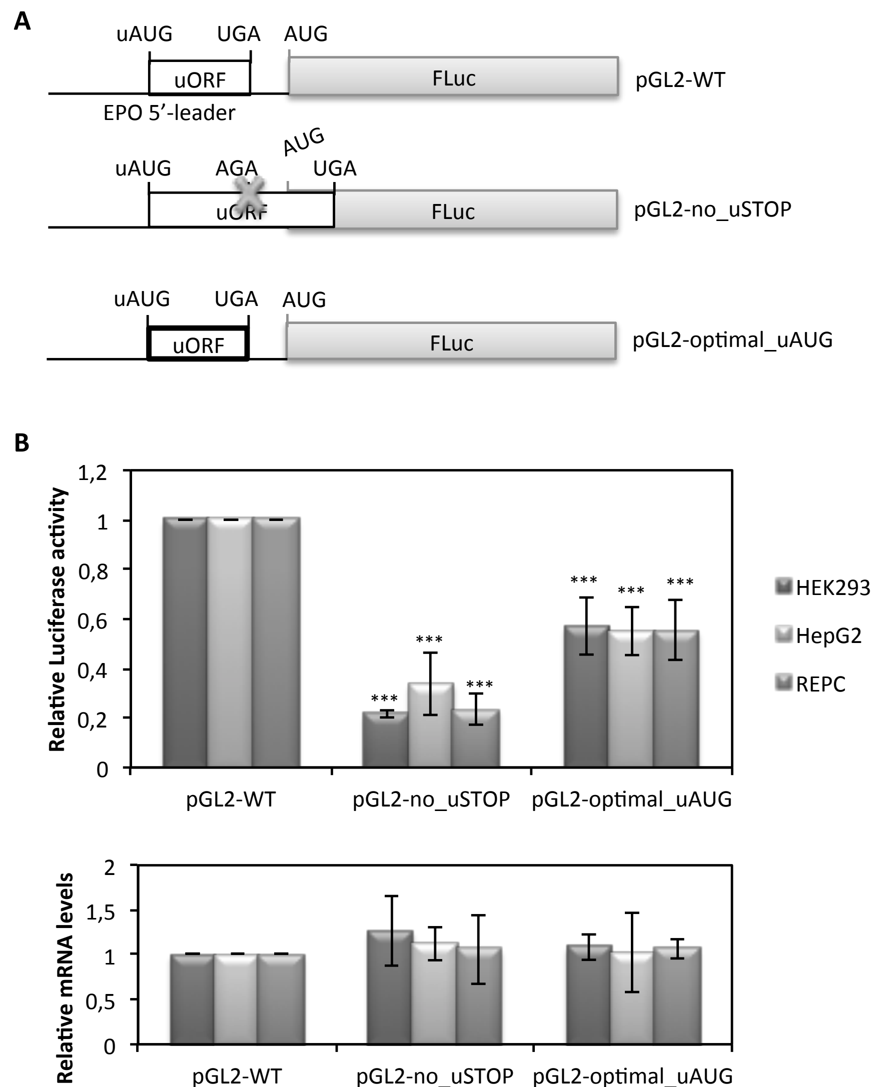


Figure II.3. Both translation reinitiation and uAUG leaky scanning are involved in the translational initiation at the main AUG codon.

(A) Schematic representation of reporter constructs. The pGL2-WT plasmid contains the wild-type human

EPO 5' leader sequence, the pGL2-no_uSTOP construct presents the *EPO* 5' leader sequence with a mutation (UGA→AGA) at the uORF translation termination codon, which makes the uORF to overlap with the luciferase ORF (the cross represent the point mutation), and the pGL2-optimal_uAUG contains the *EPO* 5' leader sequence with a optimal uAUG sequence context (gggAUGa→gccAUGg; represented by a bold lined box). **(B)** HEK293, HepG2 and REPC cells were transiently co-transfected with each one of the constructs described in (A) and with a plasmid encoding *Renilla* luciferase (pRL-TK) and analyzed as described in the legend to Figure II.2.B.

II.4.5. Translational repression exerted by the *EPO* uORF is peptide sequence-independent

Translational inhibition by uORFs in some eukaryotic transcripts is dependent upon the peptide-coding sequence of the uORF (Karagyozov et al., 2008; Wei et al., 2012). Based on these data, and knowing that the peptide encoded by the *EPO* uORF is conserved among mammalian species (Figure II.1.B), which may indicate its functional role, we next examined whether the peptide encoded by the *EPO* uORF is required for inhibition of downstream translation. For that, we cloned the pGL2-frameshift construct (Figure II.4.A) in which the uORF was modified by shifting the reading frame to generate a different amino acid sequence while preserving the uAUG context and most of the nucleotide sequence (see Materials and Methods). The pGL2-frameshift construct was used to transiently transfect the same panel of cells as before. The corresponding relative luciferase activity and mRNA accumulation levels were analyzed as previously and the results were compared to those of the pGL2-WT construct (Figure II.4.B). Our data show that in HepG2 and REPC cells, the mutant uORF of pGL2-frameshift construct allows for a relative luciferase activity slightly lower than that of the pGL2-WT construct with the normal uORF. This difference is significant in HEK293 cells but not in HepG" and REPC cells (Figure II.4.B). The lower level of relative luciferase activity expressed by the pGL2-frameshift construct may reflect the effect of a rare codon encoded by the uORF of pGL2-frameshift construct that may decrease the efficiency of translational reinitiation. Since both constructs show similar mRNA levels in all cell lines (Figure II.4.B) and frameshifting the *EPO* uORF does not derepresse translation, we can conclude that the native *EPO* uORF functions in a peptide sequence-independent manner to inhibit translation.

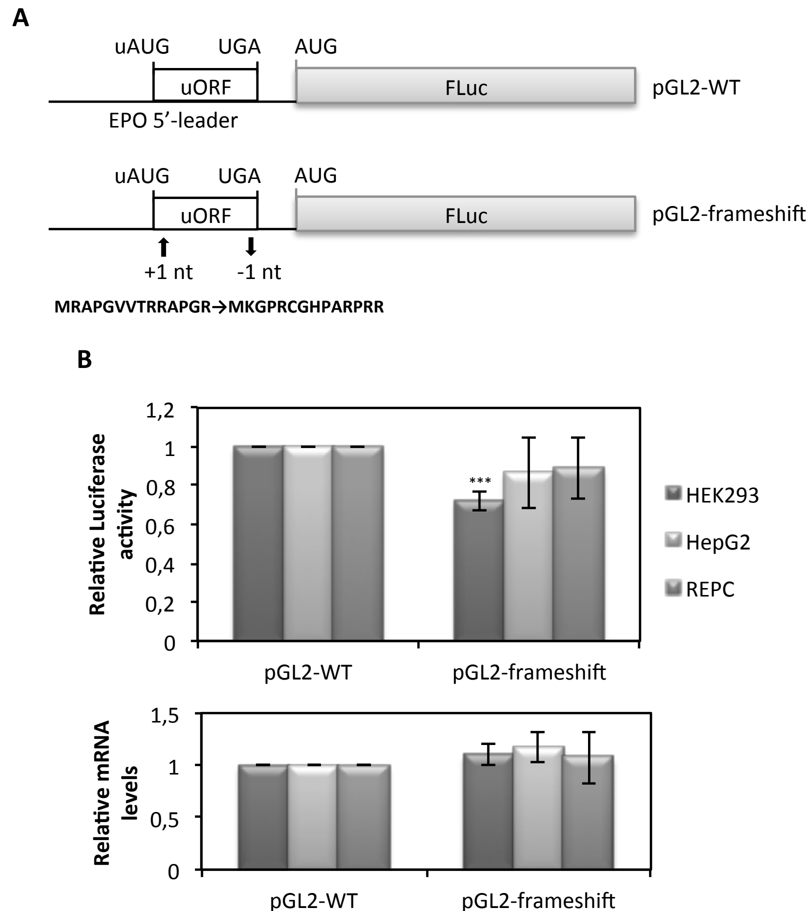


Figure II.4. Translational repression exerted by the *EPO* uORF is peptide sequence-independent.
(A) Schematic representation of the expression constructs. The pGL2-WT plasmid contains the human normal *EPO* 5' leader transcript sequence, the pGL2-frameshift vector carries a *EPO* uORF sequence modified by frameshift mutations, which consist in the insertion of one nucleotide in the second codon (+1 nt) and the deletion of one nucleotide in 13th codon (-1 nt). The resulting uORF-encoded peptide sequence is shown below. **(B)** HEK293, HepG2 and REPC cells were transiently co-transfected with each one of the constructs described in (A) and with a plasmid encoding *Renilla* luciferase (pRL-TK) and analyzed as described in the legend to Figure II.2.B.

II.4.5. The 3'UTR of the *EPO* mRNA has no impact on the inhibitory effect of the uORF

It has been shown that translational repression exerted by a uORF present in a transcript can be modulated by the corresponding 3'UTR, through protein interactions between both UTRs of the mRNA (Mehta, 2006). On the other hand, it has been shown that a pyrimidine-rich region within the human *EPO* 3'UTR is implicated in regulation of *EPO* mRNA stability and shown to bind two isoforms of a 40 kD poly(C) binding protein (PCBP), $PCBP_1$, and $PCBP_2$ (Czyzyk-Krzeska and Bendixen, 1999; McGary et al., 1997;

Ohigashi et al., 1999; Wang et al., 1995). Since these data show that the 3'UTR of the *EPO* transcript influences its expression, we hypothesized that the *EPO* 3'UTR could affect the translational inhibition exerted by the *EPO* uORF. To address this question, we first tested the effect of the *EPO* 3'UTR on the firefly luciferase activity. For that, the *EPO* 3'UTR was cloned into the pGL2-Luc vector downstream from the firefly luciferase cistron (pGL2-Luc-3'UTR construct; Figure II.5.A). Expression of this reporter gene construct was studied by transfection into HEK293, HepG2 and REPC cell lines. Firefly luciferase activity was normalized to the activity units from co-transfected *Renilla* luciferase reporter construct, as before, and the relative luciferase activity of the pGL2-Luc-3'UTR that carries the *EPO* 3'UTR was compared to that of the empty pGL2-Luc construct. The results show that the *EPO* 3'UTR alone induces about a 5-fold increase in relative luciferase activity in all cell lines studied, when compared to the relative luciferase activity of the pGL2-Luc control (Figure II.5.B). In addition, differences in the relative mRNA levels of the pGL2-Luc 3'UTR construct, quantified by RT-qPCR, were specifically observed in REPC cells (Figure II.5.B), which may reflect an higher mRNA stability specifically induced in these cells, by the *EPO* 3'UTR. Nevertheless, the level of relative luciferase activity is similar in all cell lines studied (Figure II.5.B). Therefore, it seems that the *EPO* 3'UTR affects expression of the reporter gene through different mechanisms in the three cell lines.

Then, we monitored the relative luciferase activity of the pGL2-WT reporter that harbors both *EPO* 5' and 3'UTRs (pGL2-WT-3'UTR construct; Figure II.5.C) and we compared it to the relative luciferase activity of the corresponding construct with the disrupted uORF (pGL2-no_uAUG-3'UTR; Figure II.5.C). For that, each one of these constructs was co--transfected with pRL-TK in the same cell lines as above and the luciferase activities and mRNA levels were obtained as previously (Figure II.5.D). The results were striking since insertion of the *EPO* 3'UTR into the construct pGL2-WT do not abrogate the ability of the *EPO* uORF to inhibit reporter translation in the three cell lines studied (Figure II.5.D). Indeed, the intact *EPO* 5' leader sequence in the pGL2-WT-3'UTR construct allows a significant 3-fold decrease in relative luciferase activity when compared to that observed from the pGL2-no_uAUG-3'UTR construct with the disrupted uORF. The fact that the relative mRNA levels of the pGL2-WT-3'UTR construct are higher in REPC cells than in HEK293 and HepG2 cells (Figure II.5.D) indicates that translational repression exerted by

the *EPO* uORF is stronger in REPC cells than in HEK293 and HepG2 cells, to lastly achieve the same levels of protein expression (Figure II.5.D). This led us to conjecture that the repressive effect of the *EPO* uORF independent of the effect that the *EPO* 3'UTR has in increasing mRNA levels. Thus, the *EPO* 3'UTR fails to overcome translational repression induced by the *EPO* uORF, if all it enhances the *EPO* uORF repression in REPC cells.

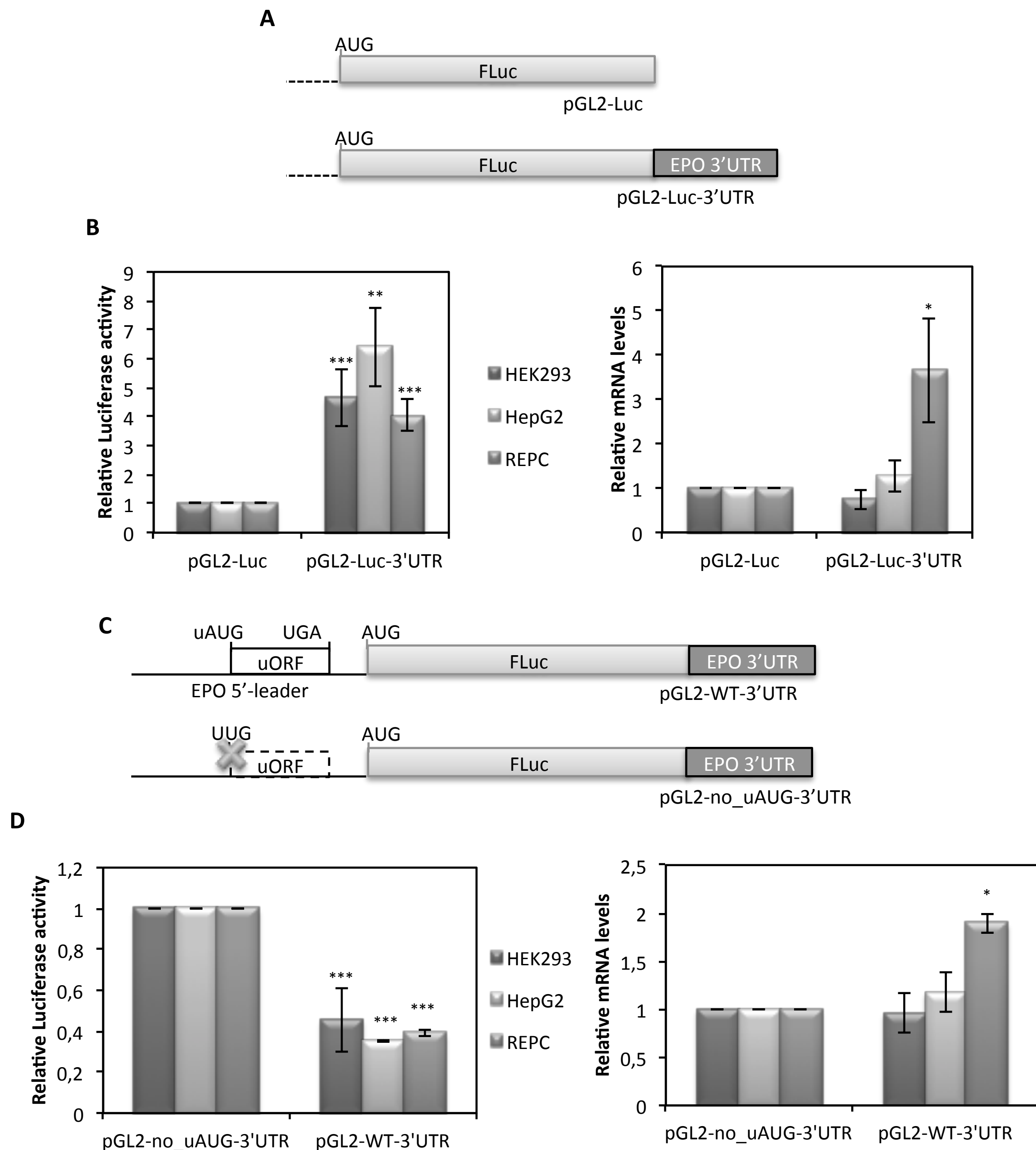


Figure II.5. The 3'UTR of the *EPO* mRNA enhances the inhibitory effect of the uORF in REPC cells.
(A) Schematic of the firefly luciferase (FLuc) reporter constructs containing the native luciferase 3'UTR (pGL2-Luc) or the 3'UTR sequence (dark grey box) of the human *EPO* transcript (pGL2-Luc-3'UTR). **(B)** HEK293, HepG2 and REPC cells were transiently co-transfected with each one of the constructs described

in (A) and with a plasmid encoding *Renilla* luciferase (pRL-TK) and analyzed as described in the legend to Figure 2B. **(C)** Schematic of the firefly luciferase (FLuc) reporter constructs containing the human *EPO* 5' leader sequence with the intact uORF and the 3'UTR sequence (dark grey box) of the human *EPO* transcript (pGL2-WT-3'UTR), or the *EPO* 5' leader sequence with a disrupted uORF due to the uAUG→UUG mutation (represented by a cross) and the EPO 3'UTR sequence (dark grey box; pGL2-no_uAUG-3'UTR). **(D)** HEK293, HepG2 and REPC cells were transiently co-transfected with each one of the constructs described in (C) and with a plasmid encoding *Renilla* luciferase (pRL-TK) and analyzed as described in the legend to Figure II.2.B.

II.4.6. The *EPO* uORF does not trigger nonsense-mediated mRNA decay

Nonsense-mediated decay (NMD) is an mRNA surveillance mechanism that rapidly degrades mRNAs carrying premature termination codons (PTCs) (Silva and Romão, 2009). In addition to its important role in mRNA quality control, it is now clear that the NMD mechanism also plays a role in regulating the steady-state level of a set of wild-type transcripts (Mendell et al., 2004; Wittmann et al., 2006; Yepiskoposyan et al., 2011). These physiological NMD substrates structurally mimic nonsense-mutated transcripts as they possess a translation termination codon that is recognized as premature. In face of this knowledge, we hypothesized that the termination codon of the human *EPO* uORF could be defined as a PTC, which would target the transcript to rapid degradation. To examine whether *EPO* transcripts could be physiological substrates for the UPF1-dependent NMD pathway, we quantified the endogenous *EPO* mRNA levels after short interfering RNA (siRNA)-mediated depletion of UPF1 in HepG2 cells. All results were compared to those obtained in NMD-competent cells transfected with nonspecific control (Luciferase) siRNAs (Figure II.6.A). At seventy-two hours after siRNAs transfection, the Western blot analysis demonstrated a decrease in UPF1 protein levels induced by siRNA of about 60%, when compared with results obtained after treatment with Luciferase siRNAs (Figure II.6.A). Under these conditions, the *EPO* mRNA levels were quantified by reverse transcription-coupled quantitative PCR assays, relative to the *EPO* mRNA levels obtained in cells treated with the control siRNA (Luc siRNA). We have previously shown that the human *HFE* transcript is a natural target of NMD (Martins et al., 2012); here, we also quantified the human *HFE* mRNA levels, as a positive control for a natural NMD target. Our data have shown that depletion of UPF1 results in a 2.5-fold increase of the abundance of the *HFE* mRNA, as expected (Figure II.6.B) (Martins et al., 2012). However, quantification of *EPO* mRNA levels did not reveal any increase at conditions of UPF1 depletion (Figure II.6.B), which indicates that, contrary to

the *HFE* transcripts, the physiological *EPO* transcripts are not natural substrates for NMD. In fact, this data is in accordance with our previous results, which are consistent with a model in which short uORFs fail to induce NMD due to the AUG-proximity effect (Peixeiro et al., 2012; Silva et al., 2008).

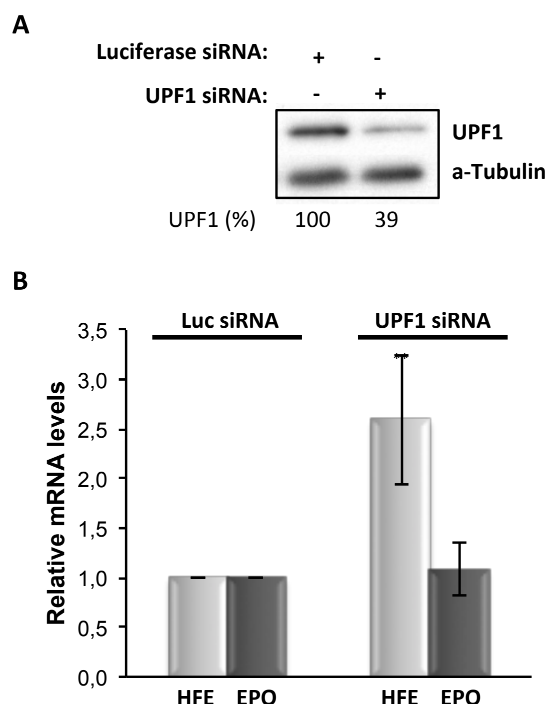


Figure II.6. The human *EPO* transcript is resistant to nonsense-mediated mRNA decay.
(A) Representative Western blot analysis of HepG2 cell extracts transfected with human UPF1 siRNA or a control siRNA target (Luciferase siRNA). Forty-eight hours after siRNA treatment, cells were harvested for protein and RNA. Immunoblotting was performed using a human UPF1 specific antibody and an α-tubulin specific antibody to control for variations in protein loading. The percentage (%) of UPF1 protein expressed in the cells after siRNA treatment is indicated below each lane. **(B)** Relative changes in *HFE* and *EPO* mRNA levels were analyzed by RT-qPCR, normalized to the levels of endogenous G protein pathway suppressor 1 (*GPS1*) mRNA. Levels of *HFE* and *EPO* mRNA obtained after cellular UPF1 siRNA treatment were compared to those obtained after luciferase siRNA treatment at the same conditions, defined as 1. The histogram shows the mean and standard deviations from three independent experiments, corresponding to three independent transfections. Statistical analysis was performed using Student's *t* test (unpaired, two tailed); (*) $p<0.05$; (**) $p<0.01$; (***) $p<0.001$.

II.4.7. EPO is regulated at the translational level in response to hypoxia, but not to nutrient deprivation, specifically in renal cells

A number of stresses, including temperature shock, DNA damage, nutrient stress, and hypoxia can lead to changes in gene expression patterns caused by a general shutdown

and reprogramming of protein synthesis. This process of translational control decreases global protein synthesis rates and increases synthesis of selective subsets of mRNAs, such as those carrying IRESs and/or uORFs, which encode stress-response proteins (Barbosa et al., 2013; Blais et al., 2004). Given that EPO is widely known to respond to hypoxia (Ebert and Bunn, 1999; Haase, 2013), here, we decided to investigate whether the *EPO* uORF directs translational control in response to this cellular stress. For that, HEK293, HepG2 and REPC cells were transiently transfected with the pGL2-WT and pGL2-no_uAUG constructs that carry the intact or disrupted *EPO* uORF, respectively, and then, cells were untreated or treated with 200μM of cobalt chloride ($CoCl_2$) to mimic hypoxia. Twenty-four hours after exposure, cells were lysed and protein and RNA were extracted and analyzed. As the transcription factor HIF complex is the key regulator of hypoxia-inducible *EPO* gene expression in HepG2 and REPC cells (Fandrey and Bunn, 1993; Frede et al., 2011), the cellular hypoxic stimulus was monitored by Western blot against HIF1α. This analysis demonstrated an increase in HIF1α protein levels induced by hypoxia, when compared with results obtained in untreated cells. Detection of α-tubulin was used to control the amount of loaded protein (Figure II.7.A). Under normoxic and hypoxic conditions, the relative luciferase activity of pGL2-WT construct was evaluated by dual luciferase assays and compared to that obtained from the pGL2-no_uAUG construct with the disrupted *EPO* uORF. The relative mRNA levels were quantified by RT-qPCR, as before (Figure II.7.B). We observed that the relative luciferase activity of pGL2-WT construct, in comparison with the relative luciferase activity of the pGL2-no_uAUG construct, does not significantly change in response to hypoxia in HEK293 and HepG2 cells (Figure II.7.B). In contrast, the pGL2-WT protein expression is increased by 1.7-fold in the REPC cells in response to hypoxia. In addition, there is no significant changes in pGL2-WT relative mRNA levels in the tested conditions (Figure II.7.B), consistent with the intact 5' leader sequence of the *EPO* mRNA directing translational derepression in response to hypoxia.

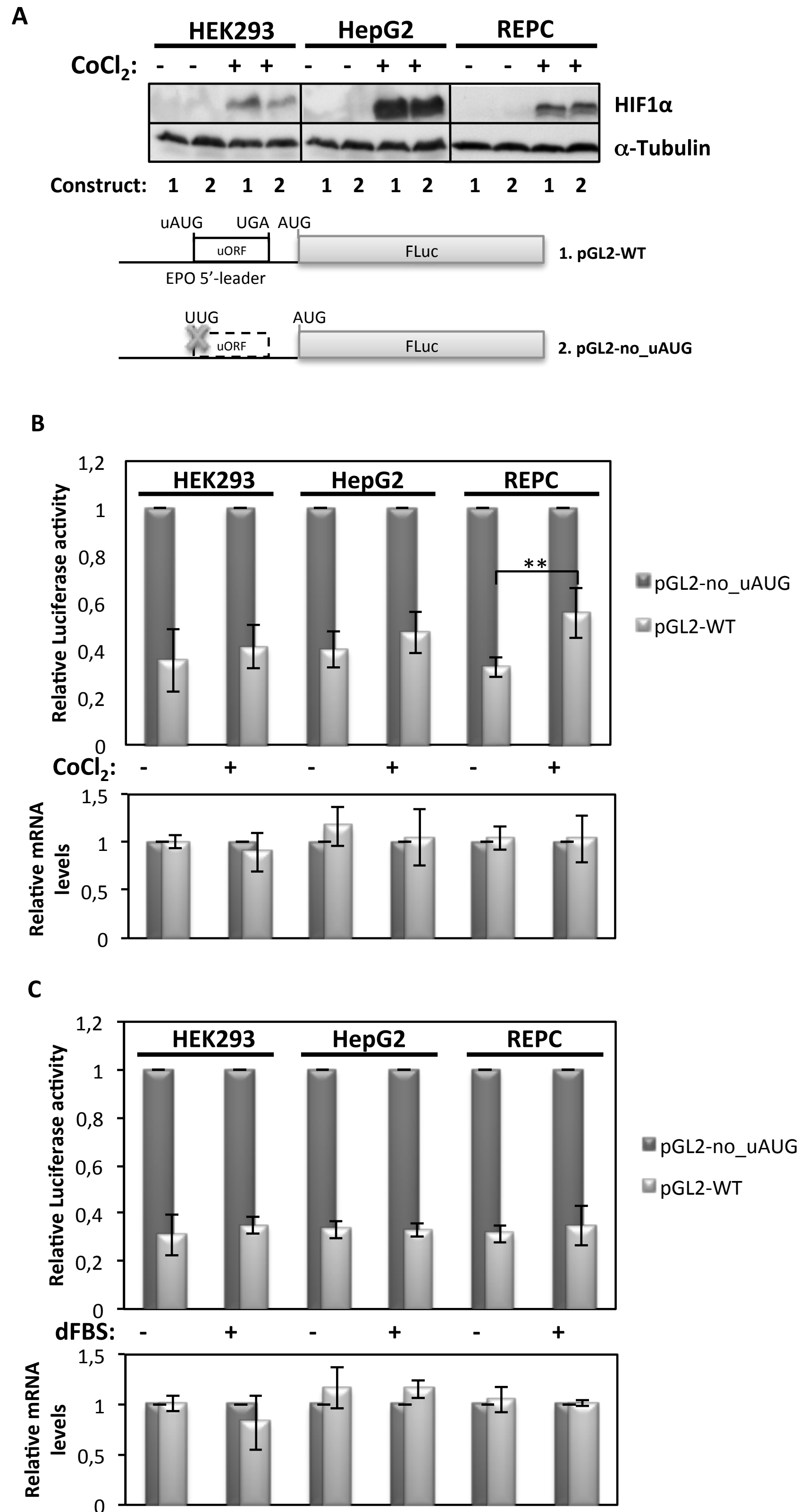


Figure II.7. The *EPO* uORF responds to hypoxia but not to nutrient starvation, specifically in REPC cells. The pGL-WT (construct 1) and pGL2-no_uAUG (construct 2) vectors represented as in Figure 2, were separately co-transfected with a plasmid encoding *Renilla* luciferase (pRL-TK) in HEK293, HepG2 and REPC cells. Six hours after transfection, cells were untreated (-) or treated (+) for twenty-four hours with 200 µM $CoCl_2$ to mimic hypoxic conditions, or with medium supplemented with 10% (v/v) dialyzed fetal bovine

serum (dFBS) to induce nutrient starvation. **(A)** Representative Western blot analysis of HEK293, HepG2 and REPC cell extracts untreated or treated with $CoCl_2$ as described. Immunoblotting was performed using a human HIF1α specific antibody to control the stress conditions, and a human α-tubulin specific antibody to control for variations in protein loading. **(B)** Normoxic ($CoCl_2$: -) and hypoxic ($CoCl_2$: +) transfected cells were lysed and analyzed as described in the legend to Figure 2. **(C)** Cells cultured in nutrient deprivation (dFBS: +) or in control conditions (dFBS: -) were lysed and analyzed as described in the legend to Figure II.2.B.

Thereafter, we investigated if this effect is observed in stimuli other than hypoxia. To test this hypothesis, HEK293, HepG2 and REPC cells transiently transfected with the pGL2-WT or pGL2-no_uAUG constructs, were cultured in nutrient deprivation using dialyzed fetal bovine serum (dFBS) in the media, as described in material and methods. Protein expression of the pGL2-WT construct was compared to that of the pGL2-no_uAUG construct as before. Our data have shown that the relative luciferase activity of pGL2-WT, in comparison with the relative luciferase activity of the pGL2-no_uAUG construct, does not significantly change in response to nutrient deprivation in all cell lines studied. Also, there were no significant changes in mRNA levels in the tested conditions (Figure II.7.C). Taken together, this set of data show that the *EPO* uORF controls translation specifically in response to hypoxia in REPC cells.

II.4.8. *EPO* translational derepression in response to hypoxia in REPC cells is not mediated by an internal ribosome entry site

As already stated, translational control can be regulated by elements within the 5’ and 3’UTRs of mRNAs, including uORFs and internal ribosome entry sites (IRESs), among others. These elements can act singly or in combination. The 5’ leader sequence of the *EPO* mRNA is relatively long, with an high CG content (81%) forming a Y secondary structure with strong and structured hairpins, as predicted by Mfold (ΔG=-93.18kcal/mol) (Figure II.8.A), which are common features of the 5’UTRs of cellular mRNAs reported to have IRES activity (Bastide et al., 2008; Komar and Hatzoglou, 2011; Tahiri-Alaoui et al., 2009). Based on these data, we first aimed to test whether the 5’ leader sequence of the human *EPO* mRNA exhibits IRES activity. For that, we cloned a dicistronic reporter construct in which the *EPO* 5’ leader sequence was inserted between the *Renilla* and firefly luciferase cistrons, but downstream of a stem-loop structure (ΔG value of -55.50 kcal/mol) to impede ribosomal reinitiation (Figure II.8.B).

In parallel, the 5'UTR from the *β-globin* and the IRES from *c-myc* mRNAs were also inserted in the dicistronic reporter system, to serve as negative and positive controls for IRES activity, respectively (Cobbold et al., 2008; Stoneley et al., 1998) (Figure II.8.B). The resulting plasmids, called "RLuc-β-globin_5'UTR", "RLuc-c-myc_IRES" and "RLuc-WT", were each transiently transfected into REPC cells. Twenty-four hours post transfection, cells were harvested and lysates subjected to a dual luciferase reporter assay and the subsequent ratio of firefly luciferase to *Renilla* luciferase was compared to that from the control "RLuc-empty" plasmid, which contains a short linker sequence between the two luciferase cistrons (Figure II.8.B). Results show that the *c-myc* IRES significantly increased the relative luciferase activity by 2.5-fold (Figure II.8.C), which indicates that the *c-myc* IRES is able to drive cap-independent firefly luciferase expression, in accordance to what it is expected for an mRNA containing IRES activity (Cobbold et al., 2008; Stoneley et al., 1998). However, the *EPO* 5'UTR behaved as the *β-globin* 5'UTR and both mantained the relative luciferase activity at levels similar to those of the "RLuc-empty" construct, showing that none of them allow detectable internal translation initiation (Figure II.8.C). These results illustrate that the *EPO* 5' leader sequence does not exhibit IRES activity in normoxic REPC cells.

Knowing that in response to hypoxia, some transcripts may increase efficiency of translation by facilitating internal translation initiation through a process such as IRES-mediated initiation (Schepens et al., 2005), we subsequently tested if EPO translational derepression in response to hypoxia in REPC cells is mediated by IRES. Thus, we analyzed the expression of the RLuc-WT construct with the intact *EPO* uORF, relatively to that of the RLuc-no_uAUG construct carrying the disrupted *EPO* uORF, both cloned in the dicistronic reporter plasmid (Figure II.8.D), as before, in normoxic and hypoxic REPC cells. To mimic hypoxia, cells were treated with $CoCl_2$ and induction of hypoxia was monitored, as previously, by Western blotting using an antibody against HIF1α (Figure II.8.E). As expected, results show high accumulation of HIF1α protein during hypoxic incubation. In addition, data show that relative luciferase activity from the RLuc-WT construct in normoxic cells is at about 45% of the relative luciferase activity of RLuc-no_uAUG construct. Under hypoxia, the same relative luciferase activity was observed (Figure II.8.F). This result supports the idea that *EPO* 5' leader sequence does not allow for internal translation initiation irrespectively of stress conditions. Thus, in normoxic or

hypoxic renal cells, *EPO* translation involves the processive scanning of ribosomes from the 5'-end of the *EPO* transcript.

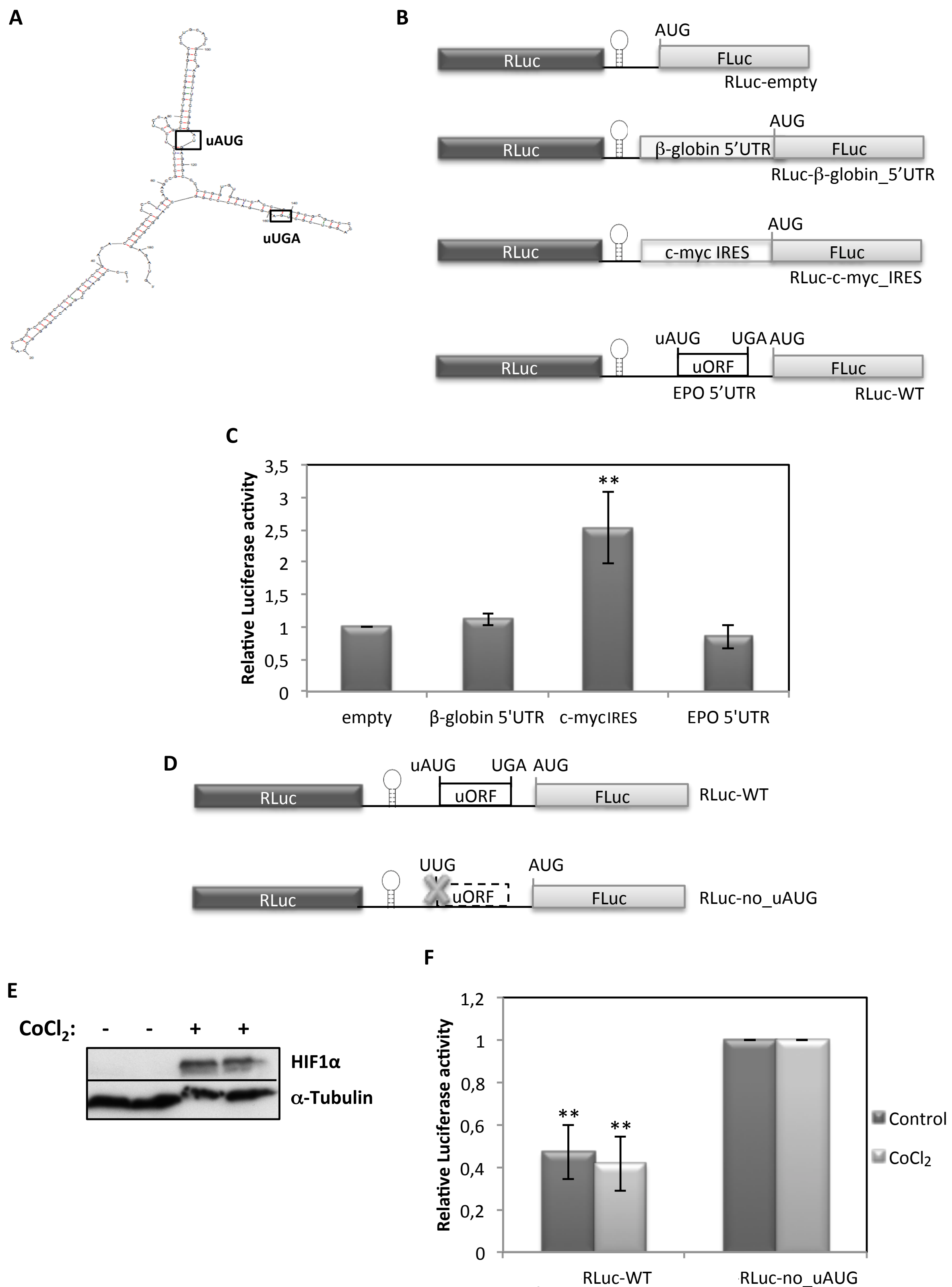


Figure II.8. EPO translational derepression in response to hypoxia in REPC cells is not mediated by an internal ribosome entry site (IRES).

(A) Representation of the secondary structure of the 5' leader sequence of the human *EPO* transcript as predicted by Mfold webserver. **(B)** Schematic of the dicistronic luciferase vectors. The 5'UTR sequence of the human *EPO* transcript (EPO 5'UTR) to be tested for IRES activity, as well as the 5'UTR sequence of the human *β-globin* transcript (β-globin 5'UTR), or the *c-myc* IRES sequence previously described by Stoneley et al. (1998), were inserted between the Renilla (RLuc) and firefly (FLuc) luciferase cistrons, downstream of a hairpin structure (represented by a stem loop) in the multiple cloning site spacer of the RLuc-empty vector, to create the RLuc-WT, RLuc-β-globin and the RLuc-c-myc constructs, respectively. **(C)** REPC cells were transiently transfected with each one of the constructs described in (B) and with a plasmid encoding *Renilla* luciferase (pRL-TK) and analyzed as described in the legend to Figure 2B. **(D)** Schematic of the dicistronic reporter constructs used to test if the *EPO* 5' leader sequence contains an IRES activated during hypoxia. RLuc-WT contains the human *EPO* 5' leader sequence with the intact uORF, as defined in (A) and the RLuc-no_uAUG contains the *EPO* 5' leader sequence with a disrupted uORF due to the uAUG→UUG mutation (represented by a cross). REPC cells were transfected with these constructs. Six hours later, cells were untreated or treated with 200μM $CoCl_2$ for twenty-four hours. **(E)** Representative Western blot analysis of REPC cell extracts untreated (-) or treated (+) with $CoCl_2$. Immunoblotting was performed using a human HIF1α specific antibody to control the stress conditions, and a human α-tubulin specific antibody to control for variations in protein loading. **(F)** Relative luciferase activity was quantified as described in the legend to Figure II.2.B.

II.4.9. *EPO* translational derepression in response to hypoxia is mediated by leaky scanning of ribosomes through the inhibitory uORF

Based on our data showing that both translation reinitiation and uAUG leaky scanning are involved in the translational initiation at the main AUG codon of the pGL2-WT construct, we next addressed which of these two mechanisms of initiation occur to overcome the inhibitory function of the *EPO* uORF in response to hypoxia in REPC cells. Thus, as previously in Figure II.3., protein expression of pGL2-WT (construct carrying the wild type *EPO* uORF), pGL2-no_uSTOP [construct carrying a mutation at the stop codon (TGA→AGA) of the uORF], and pGL2-optimal_uORF (construct carrying the uAUG in an optimal context) constructs was analyzed in REPC cells under normoxic and hypoxic conditions. To mimic hypoxia, cells were treated with $CoCl_2$ and induction of hypoxia was monitored, as before, by Western blotting using an antibody against HIF1α (Figure II.9.A). As expected, hypoxia conditions led to the accumulation of HIF1α (Figure II.9.A). Furthermore, our results show that relative luciferase activity from the pGL2-no_uSTOP construct with the extended uORF increases 1.6-fold in hypoxic conditions, when compared to its activity in normoxia, meaning that, under hypoxia, the uAUG is less efficiently recognized (Figure II.9.B). In contrast, increasing the translation initiation sequence context to an optimal start codon context at the pGL2-optimal_uAUG construct does not affect the corresponding relative luciferase activity under hypoxic

versus normoxic conditions (Figure II.9.B). In addition, no significant changes in mRNA levels were observed in the tested conditions (Figure II.9.B). These results are consistent with a model in which derepression of translation in hypoxic REPC cells occurs by an increase in leaky scanning of ribosomes through the inhibitory *EPO* uORF.

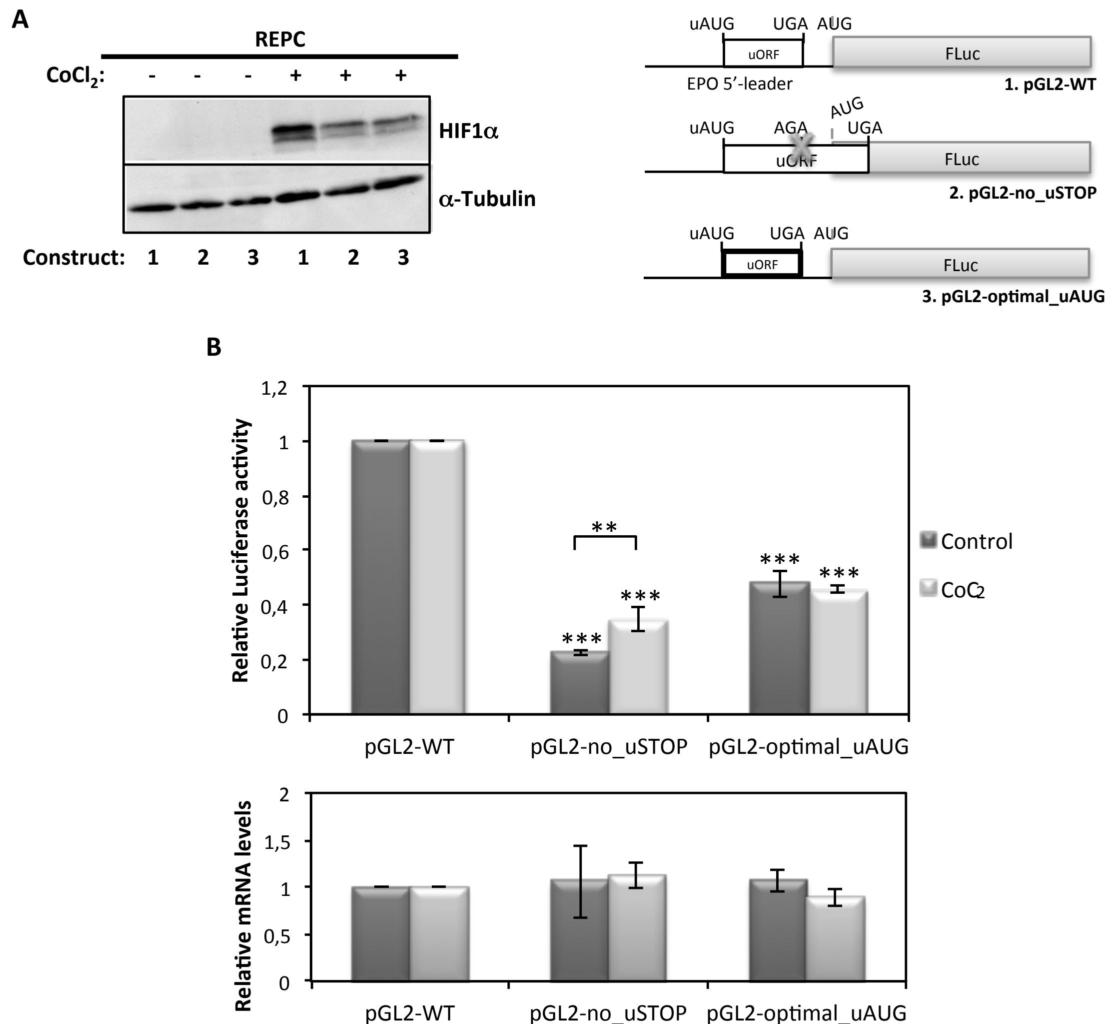


Figure II.9. *EPO* translational derepression in response to hypoxia of REPC cells is mediated by leaky scanning of ribosomes through the inhibitory uORF.

The pGL-WT (construct 1), pGL2-no_STOP (construct 2) and pGL2-optimal_uAUG (construct 3) vectors represented as in Figure 3, were separately co-transfected with a plasmid encoding *Renilla* luciferase (pRL-TK) in REPC cells. Six hours after transfection, cells were untreated (-) or treated (+) for twenty-four hours with 200 µM $CoCl_2$ to mimic hypoxic conditions. **(A)** Representative Western blot analysis of transfected REPC cell extracts untreated (-) or treated (+) with $CoCl_2$ as shown. Immunoblotting was performed using a human HIF1α specific antibody to control for hypoxia, and a human α-tubulin specific antibody to control for variations in protein loading. **(B)** Relative luciferase activity was quantified as described in the legend to Figure II.2.B.

II.4.10. Hypoxia-induced phosphorylation of eIF2α is required for *EPO* translational regulation

Phosphorylation of eIF2α is a rapid consequence of hypoxic stress, reducing the availability of competent initiation complexes. Indeed, when eIF2α is phosphorylated, the ternary complex becomes scarce and global translation compromised (Sonenberg and Hinnebusch, 2009). Despite eIF2α phosphorylation, the presence of uORFs can promote the increased expression of certain stress-related mRNAs (Barbosa et al., 2013; Dang Do et al., 2009; Koritzinsky et al., 2007; Vattem and Wek, 2004). This occurs through a mechanism that can involve two or more uORFs and reduced ternary complex, which makes reinitiation to take longer allowing the bypass of a second uORF improving the recognition of the main AUG located further downstream. In some other mRNAs, the mechanism to promote translational derepression appears to involve bypass of a single inhibitory uORF (Lee et al., 2009; Lewerenz et al., 2012). Based on these data, we next addressed whether the mechanism by which *EPO* uORF significantly derepresses translation in hypoxic REPC cells occurs through eIF2α phosphorylation. For that, the state of eIF2α phosphorylation was examined by immunoblotting using anti-phospho-eIF2α antiboby, in REPC cells transiently transfected with pGL2-WT or pGL2-no_uAUG constructs, and untreated or treated with $CoCl_2$ 200μM during 24 hours. Results show that the extent of eIF2α phosphorylation in this cell line is increased by induction of hypoxia (Figure II.10.A). Taking advantage of these data, we next tested whether the treatment of these cells with thapsigargin, a potent ER stress agent that directly activates eIF2α kinases without activating any other signaling pathway (Harding et al., 2001; Koumenis et al., 2002), would induce translational derepression of the luciferase reporter. REPC cells transiently transfected with pGL2-WT or pGL2-no_uAUG constructs were untreated or treated with thapsigargin 1μM. Twenty-four hours later, cells were lysed and protein levels were measured by luminometry assays and the mRNA levels quantified by RT-qPCR, as previously. The extent of eIF2α phosphorylation in thapsigargin-treated cells was examined by immunoblotting, as before. Figure II.10.A shows that the extent of eIF2α phosphorylation was increased by thapsigargin treatment. Figure II.10.B shows that phosphorylation of eIF2α effectively induces a significant increase of translation of the pGL2-WT mRNA relatively to that of pGL2-no_uAUG (2-fold increase), specifically in treated when compared to untreated REPC

cells. In these experiments, relative mRNA levels were comparable in all conditions tested. From these findings, we conclude that eIF2α phosphorylation regulates the translation of the pGL2-WT reporter mRNA *via* the *EPO* uORF, in REPC cells in response to hypoxia. Taking together these results and those from Figure II.9., we suggest that in REPC cells exposed to hypoxia, eIF2α is phosphorylated, which up-regulates the translation of *EPO* mRNA by increasing the rate of ribosomal bypass of the inhibitory uORF.

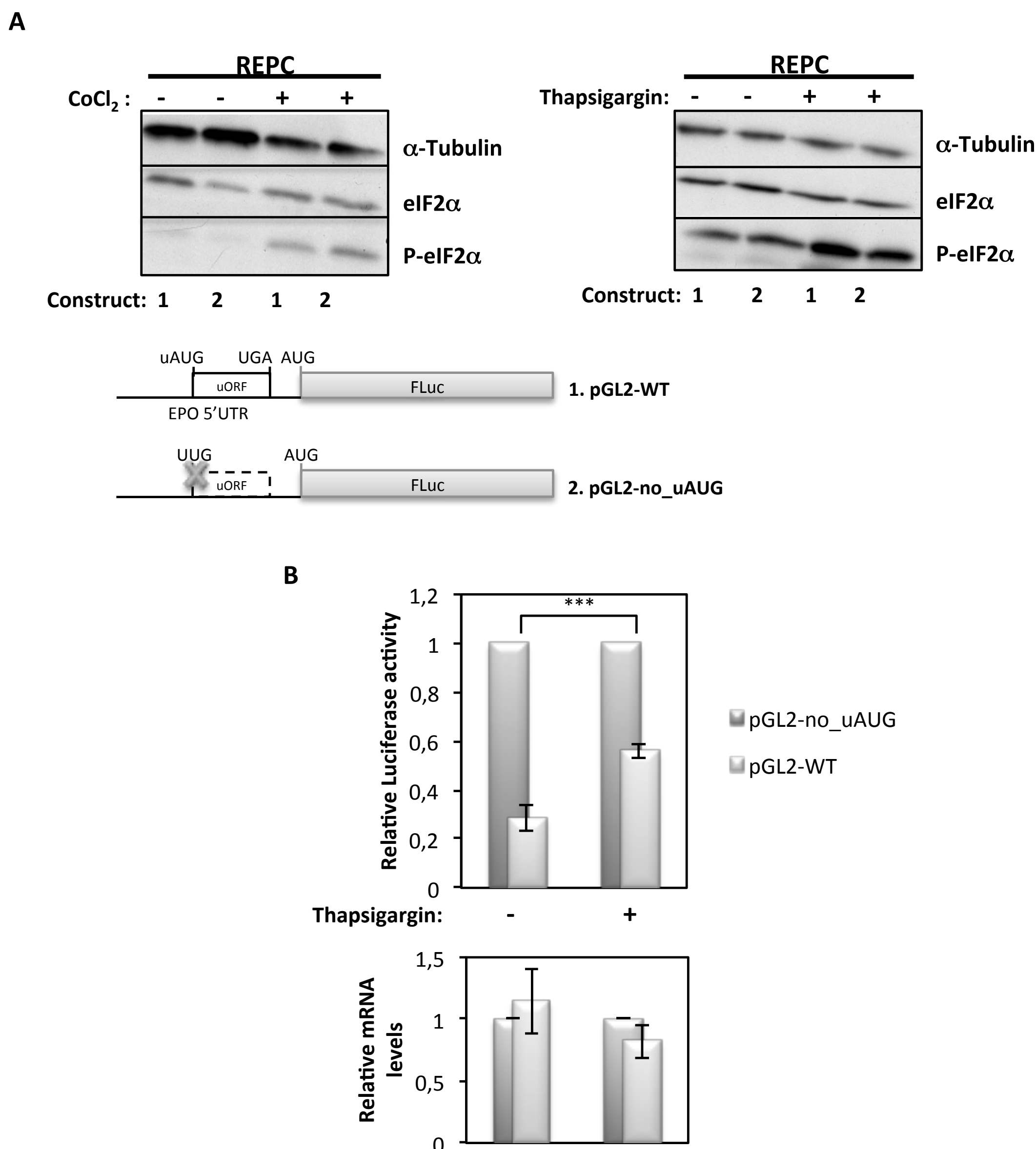


Figure II.10. Hypoxia induces phosphorylation of eIF2α, which is required for *EPO* translational regulation in REPC cells.

The pGL-WT (construct 1) and pGL2-no_uAUG (construct 2) vectors represented as in Figure 2, were separately co-transfected with a plasmid encoding *Renilla* luciferase (pRL-TK) in REPC cells. Six hours after transfection, cells were untreated (-) or treated (+) for twenty-four hours with 200 µM $CoCl_2$ to mimic

hypoxic conditions or with 1µM thapsigargin. **(A)** Representative Western blot analyses of REPC cell extracts untreated or treated as described. Immunoblotting was performed using human eIF2α and human phosphorylated eIF2α specific antibodies to control for stress conditions, and human α-tubulin specific antibody to control for variations in protein loading. **(B)** Relative luciferase activity was quantified as described in the legend to Figure II.2.B.

II.5. Discussion

The 5' leader sequences of about 49% of eukaryotic mRNAs are known to harbor one or more uORF(s) (Calvo et al., 2009). Interestingly, a high percentage of RNAs containing uORFs encode oncogenes, hormones and growth factors (Kozak, 1991), and expression of these genes is highly regulated, as their protein products are important in cell growth and proliferation. Studies of a subset of RNAs harboring uORFs have shown that uORFs can function by reducing the efficiency of translation initiation of the main downstream ORF in unstressed conditions (Calvo et al., 2009; Morris and Geballe, 2000). Modulation of translation efficiencies of the downstream ORF can occur via a number of distinct mechanisms including translation termination and reinitiation, as well as uORF-encoded peptide dependent ribosome stalling and mRNA decay induction (Barbosa et al., 2013; Morris and Geballe, 2000).

EPO is an essential protein for stimulating the differentiation and proliferation of erythroid progenitors in the bone marrow (Fandrey, 2004). During fetal development, EPO is produced mainly in the liver. Following birth, expression of EPO in the liver is reduced to low levels and the kidney accounts for about 90% of EPO production (Bunn, 2013). The notion that EPO production is markedly up-regulated by hypoxia and that it stimulates erythropoiesis in a dose-dependent manner led to the now well accepted paradigm of a negative feedback loop where hypoxia induces an increase in EPO hormone production in the kidney, which then circulates in the plasma and binds to receptors abundantly expressed on erythroid progenitor cells, thereby promoting the viability, proliferation, and terminal differentiation of erythroid precursors, and causing an increase in red blood cell mass. The oxygen-carrying capacity of the blood is thereby enhanced, increasing tissue oxygen tension, thus completing the feedback loop and suppressing further expression of EPO (Bunn, 2013).

EPO gene is one in many examples of genes that presents several layers of expression regulation. The most well characterized mechanisms are at the transcriptional level, and are correlated to the increase of the *EPO* mRNA levels during hypoxic conditions (Jelkmann, 2011). Human *EPO* mRNA (NM_000799), which encodes a 166 amino acid hormone, presents a 5' leader sequence with 181 nucleotides that encompasses a uORF with 14 codons, located 22 nucleotides upstream of the *EPO* AUG codon (Figure II.1.). The 5' leader sequences of the *EPO* mRNAs of human, chimpanzee, gorilla, orangutan, common marmoset, mouse, and rat show high sequence similarity, being the uORF highly conserved among these species (Figure II.1.). In addition, a high percentage of transcripts encoding hormones hold uORFs involved in their translational control, responding to cell type and external stimuli (Hood et al., 2009; Medenbach et al., 2011; Morris, 1997; Sachs and Geballe, 2006; Wethmar et al., 2010a). These findings prompted us to investigate the function of the human *EPO* uORF in its translational control. We found that the *EPO* uORF is translatable and the presence of the intact uORF significantly inhibits the translation of the downstream ORF in different cell lines (Figure II.2.). The preservation of the uORF repressive effect on downstream translation in different cell types, suggests that the uORF is a major determinant of EPO protein expression.

Aiming to know how the ribosomes ever gain access to the *EPO* main AUG codon, the results shown in Figure II.3. suggest that a small percentage of ribosomes bypass the uAUG codon and the corresponding uORF and that additional ribosomes are able to reinitiate at the EPO start site after translation of the uORF. Knowing that the presence of a purine at the -3 position relative to the AUG codon, is usually thought to be sufficient for efficient initiation (Kozak, 2001), it is really not too surprising that some ribosomes leak past the *EPO* uORF AUG codon, but the majority of them recognize the uAUG, translate the uORF and may reinitiate at the downstream main ORF, even though the uORF AUG codon has the non-optimal gggAUGa sequence context. As an A at position -3 can be superior to G (Kozak, 2001), which occurs in the mouse and rat sequences (Figure II.1.), we hypothesize that in these species, *EPO* uORF may be even more efficiently recognized than in humans.

Given the *EPO* uORF is highly conserved in sequence among different mammalian species (Figure II.1.), we hypothesized that the *EPO* uORF-encoded peptide could induce ribosome stalling in a sequence-dependent manner. However, our results show that this

is not the case as uORF sequence frameshifting still retains the inhibitory effect on downstream translation, meaning that the uORF-dependent repression mechanism does not need a specific peptide (Figure II.4.). Thus, we currently do not understand the significance of the conservation of the peptide sequence. The fact that during evolution the A at position -3 was changed to G, which allows for weaker uAUG recognition, is in accordance with the fact that the uORF-mediated repression effect is uORF-encoded peptide independent. Nonetheless, it might be required for an unidentified function of the uORF other than inhibition of downstream translation.

Knowing that tissue-specific expression of the *EPO* gene and its induction by hypoxia are dependent on far upstream *cis* elements and an enhancer element downstream from the polyadenylation signal (Madan et al., 1995; Semenza, 2001), we also aimed to investigate the potential role for translation control of the *EPO* 3'UTR. Utilizing reporter constructs where the *EPO* 5' leader sequence and/or the *EPO* 3'UTR flank the firefly luciferase cistron, we demonstrated that the presence of the *EPO* 3'UTR induces an increase in the reporter mRNA levels, specifically in REPC cells. In addition, we also observed that the reporter protein expression is increased in the presence of the *EPO* 3'UTR in all cells tested (Figure II.5.B). However, when the intact 5' leader sequence of the *EPO* transcript is also present in the reporter construct, the translational repression exerted by the uORF is not released in the presence of the *EPO* 3'UTR sequence, in fact it seems that in REPC the uORF repression in enhanced in the presence of the *EPO* 3'UTR (Figure II.5.D). These results show that the *cis*-acting elements present in the *EPO* 3'UTR involved in increasing EPO gene expression do not seem to affect the mechanism by which the uORF represses translation. Thus, these two regions may act at different layers of *EPO* gene expression regulation.

It is well known that transcripts carrying uORFs are natural targets for NMD (Mendell et al., 2004). However, some naturally occurring uORF containing transcripts escape degradation (Lee et al., 2009; Yaman et al., 2003; Zhou et al., 2008a). In the present work, we show that the human endogenous *EPO* transcript, with the 14-codons uORF, is not an NMD target (Figure II.6.). This result is in accordance with our previous data showing that transcripts carrying a PTC in close proximity to the AUG (for example, a PTC at position 15) escape NMD (Romão et al., 2000; Silva et al., 2006). According to our model (Barbosa et al., 2013; Silva and Romão, 2009), only transcripts harboring at least

one uORF with a critical length would trigger NMD, while those with smaller uORF(s), such as *EPO* transcript, could be NMD-resistant because of the poly(A) binding protein cytoplasmic 1 (PABPC1) proximity to the uORF termination codon, due to mRNA circularization during translation, would induce an efficient uORF translation termination and inhibit NMD (Peixeiro et al., 2012; Silva et al., 2008).

Studies in both eukaryotes and prokaryotes have demonstrated that uORF-encoded peptides can cause ribosomal stalling by a range of mechanisms, including interference with the peptidyl transferase center activity (Gu et al., 1994; Lovett and Rogers, 1996), thereby inhibiting translation termination by preventing peptidyl-tRNA hydrolysis (Janzen et al., 2002) or by blocking elongating or terminating ribosomes in response to a cellular signal (Hood et al., 2009; Luo et al., 1995; Wang and Sachs, 1997). Also, we and others have demonstrated an association between defects in translation termination and NMD (Amrani et al., 2004, 2006; Peixeiro et al., 2012; Singh et al., 2008). The fact that the inhibitory function of the *EPO* uORF is peptide independent (Figure II.4.) corroborates with the data showing that that the human *EPO* transcript is NMD resistant (Figure II.6.).

EPO is the primary regulator of mammalian erythropoiesis and is produced by the kidney and the liver in an oxygen-dependent manner. However, it is now clear that EPO is a multifunctional molecule produced and utilized by many tissues that rapidly responds to different cell stress stimuli and tissue injuries (Arcasoy, 2008; Brines et al., 2008; Ruifrok et al., 2008; Ryou et al., 2012). Based on these data, we have investigated the role of the *EPO* uORF in three different cell lines derived from embryonic kidney, liver and kidney, in the response to chemical hypoxia or to nutrient deprivation. We found that the protein expression from the construct with the intact *EPO* 5' leader sequence is significantly increased, specifically in REPC cells in response to hypoxia, but not during nutrient limitation (Figure II.7.). A small increase in translational efficiency was also observed in HepG2 cells, but it is not significant; by another hand, no effect was observed in HEK293 cells (Figure II.7.). Thus, our data reveal that reporter translation is controlled by the *EPO* uORF in renal cells to ensure maximal expression during hypoxia stress. Indeed, these results mimic the *in vivo* EPO expression in response to hypoxia: low increase in liver cells, and a robust increase in renal cells (Fandrey, 2004; Jelkmann, 1992). These results show that the translational control mediated by the *EPO* uORF is

another layer in the already complex control of *EPO* gene expression in response to hypoxia, but it seems to parallel its transcriptional control in what concerns cell type specificity and external stimuli response. Therefore, our results suggest that there is a coordinated transcriptional and translational control of EPO expression, which is necessary for optimal expression in hypoxic renal cells.

Trying to understand the mechanism by which *EPO* translational derepression occurs in response to hypoxia in REPC cells, we have observed that *EPO* translation does not implicate the induction of IRES activity, despite the high CG content of the *EPO* 5' leader sequence forming a Y secondary structure with strong and structured hairpins, characteristics of IRES sequences. Instead, it involves the processive scanning of ribosomes from the 5'-end of the *EPO* transcript whether in normoxic or hypoxic renal cells (Figure II.8.), suggesting that *EPO* translation is not controlled *via* different elements located in its 5' leader region, condition that occurs in some transcripts such as VEGF-A isoform, in which an uORF is located within an IRES (Bastide et al., 2008).

To comprehend how scanning ribosomes better reach the main ORF to increase translation when REPC cells are hypoxic, we have observed that other than the reinitiation mechanism, more ribosomes bypass the *EPO* uORF in response to hypoxia, and thus, the uORF decreases its barrier function to scanning ribosomes and translational rate of the main ORF is significantly increased (Figure II.9.). What is the biochemical mechanism by which scanning ribosomes bypass *EPO* uORF and reach the main AUG in hypoxic renal cells? It is known that hypoxia activates eIF2α phosphorylation (Sonenberg and Hinnebusch, 2009), which is also in accordance with our data shown in Figure II.10.A. On another hand, stress-induced eIF2α phosphorylation significantly increases translation of the reporter main ORF in REPC cells (Figure II.10.B.), data may reflect the *in vivo EPO* gene expression profile observed in renal cells in response to hypoxia. Taking together, our results show that in basal conditions where eIF2α phosphorylation is low, translation of the *EPO* uORF serves as a barrier that inhibits translation of the downstream *EPO* main ORF in different cell types. During hypoxia, enhanced eIF2α phosphorylation significantly increases ribosome bypass of the uORF in renal cells, probably due to the non-optimal uAUG sequence context, and translation of the downstream main ORF occurs with higher efficiency. Indeed, the non-optimal sequence context of the uAUG is a feature conserved among

each of the species illustrated in Figure II.1., which may reflect its functional role in translational control. The finding that both transcriptional and translational mechanisms control EPO expression in renal cells suggests that EPO is tightly regulated in response to hypoxia.

The best studied mechanism of translational control is the one governing yeast *GCN4* and mammalian *ATF4* and *ATF5* transcripts (Lewerenz et al., 2012; Vattem and Wek, 2004; Watatani et al., 2008; Zhou et al., 2008a, 2008a). In the case of yeast *GCN4*, *ATF4* and *ATF5* mRNAs, the major principle of this mechanism is that the translation of the upstream uORF stimulates translation or reinitiation at a downstream AUG, whereas translation of the downstream uORF leads to translation termination and dissociation of ribosomes. Following translation of the upstream uORF, if eIF2 levels become limiting, the recruitment of the ternary complex by the ribosome is markedly reduced and as a result, ribosomes have a higher probability of reinitiating translation after the downstream uORF and thereby reinitiate translation at the main ORF. Like *ATF4/5* transcripts, the CCAAT/enhancer-binding protein homologous protein (*CHOP*) mRNA is also translationally regulated in a uORF-dependent manner under stress. In this case, a single uORF element is a significant barrier to CHOP translation in non-stressed conditions. However, in response to stress, induced eIF2α phosphorylation facilitates bypass of the repressing uORF, allowing scanning ribosomes to instead initiate translation at the CHOP coding sequence (Palam et al., 2011). Our results suggest the hypothesis that the *EPO* uORF may serve to control the access of ribosomes to the downstream main AUG codon by a mechanism different from that described for *ATF4* and *ATF5*, but related to that described for the uORF in the *CHOP* mRNA. Although the EPO regulatory model shares with CHOP and ATF4/5 translational control the idea that eIF2α phosphorylation can bypass an inhibitory uORF, EPO and CHOP accomplish this without the aid of a positive-acting uORF that facilitates translation reinitiation. Instead, *CHOP* as well as *EPO* transcripts have a similar 5' leader sequence configuration. Contrary to what occurs in *ATF4/5* transcripts, *CHOP* and *EPO* mRNA have single uORFs. However, CHOP uORF differs from *EPO* uORF, because it has two uAUG codons with a poor translation initiation context. In contrast, *EPO* uORF has a single uAUG that shares with the second uAUG of the *CHOP* uORF a comparable sequence context for initiation that can be bypassed in response to eIF2α phosphorylation (Chen et al., 2010; Jousse et

al., 2001; Palam et al., 2011). Another difference among these two systems, seems to be the tissue specificity observed in the translational control mediated by the *EPO* uORF in response to hypoxia, which may suggest the involvement of potential tissue specific regulator(s) that would facilitate the bypass of the *EPO* uORF, specifically in renal cells during hypoxia. In the future it will be interesting to determine whether EPO responds to other stress stimuli in combination with eIF2α phosphorylation through the uORF bypass mechanism in different cell types.

A different potential mechanism by which there is an increase in ribosome bypass of the *EPO* uORF in hypoxic renal cells is the possibility that the sequence length (117 nucleotides) preceding the *EPO* uORF is not enough for scanning ribosomes to acquire the ternary complex in conditions where eIF2 levels become limiting (i.e. eIF2α is phosphorylated). Although this hypothesis has been tested, for the stress-induced CHOP translation, the insertion of a 120 nucleotides sequence in the 130 nucleotides sequence present upstream of the uORF did not change the translational rate of the reporter mRNA (Palam et al., 2011). The fact that both transcripts show the sequence preceding the uORF with similar lengths (130 nucleotides in *CHOP* mRNA *versus* 117 nucleotides in *EPO* mRNA) is indicative of no influence of the sequence length preceding the uORF in the *EPO* transcript.

We do not yet completely understand the biochemical basis for the ribosomal bypass of the uORF in our model of *EPO* translational control in response to hypoxia specifically in renal cells. Lowered eIF2-GTP levels may contribute to the reduced recognition of the *EPO* uORF. Additional contributors to this bypass may be the eIF2α phosphorylation mediated expression regulation of other critical translation factors, or tissue specific regulators that would then facilitate the bypass of the *EPO* uORF during hypoxia. Also, this mechanism may involve specific sequences or conditions that have not yet been identified but will be challenged to investigate.

Overall, the current results report a new mechanism involved in the human *EPO* gene expression regulation. The translational control by the *EPO* uORF and its response to hypoxia might present a new target for therapeutic interventions in diseases related to the hematopoietic functions of EPO.

II.6. Acknowledgements

We are grateful to Margarida Gama Carvalho, Marco Candeias and Abdessamad Tahiri-Alaoui for supplying the pRL-TK, p53 "A", and psiRF plasmids, respectively, and to Joachim Fandrey for kindly providing the REPC cells. We would like to thank Ana Ramos and Rafaela Lacerda for cloning the "pGL2-RLuc-empty", "pGL2-RLuc-β-globin" and "PGL2-RLuc-c-myc" plasmids. We would also like to thank Isabel Peixeiro and Alexandre Teixeira for critical reading of the manuscript. This research was partially supported by Fundação para a Ciência e a Tecnologia (PEst-OE/BIA/UI4046/2011, PTDC/BIM-MED/0352/2012 and SFRH/BD/63581/2009 to C.B.).

CHAPTER III – The role of the erythropoietin upstream open reading frame in the human neuronal tissue

Author's note

Manuscript in preparation.

III.1. Abstract

Beyond its role in erythropoiesis, erythropoietin (EPO) plays several other non-hematopoietic roles as a consequence of its expression in other tissues, such as the brain, where it acts as a neuroprotector. EPO expression is tightly regulated in order to maintain its correct expression in response to stress conditions. *EPO* transcript contains an upstream open reading frame (uORF) of 14 codons. We have previously shown that *EPO* uORF is functional in the liver and the kidney, the major production sites of EPO, and that its repression is released under hypoxia, proving the importance of EPO expression in response to stress. Here, we show that *EPO* uORF is also functional in the brain, using SW1088 cells. Our data demonstrate that the uORF AUG is recognized by the preinitiation complex, thus inhibiting the recognition of the main AUG. Yet, some ribosomes bypass the uAUG and others reinitiate after uORF translation, allowing the production of the downstream protein. Moreover, we prove that *EPO* uORF functions in a peptide-independent manner and independent of *EPO* 3' untranslated region (3'UTR). In addition, we observe a uORF-dependent induction of *EPO* translation under chemical ischemia. However, the underlying mechanism differs from those previously described. Actually, we show that protein levels are maintained, whereas mRNA levels decrease dramatically under ischemia, meaning that the efficiency of mRNA translation is greater in response to ischemia. The molecular basis underlying this process is still unclear, but these findings propose a specific regulation of EPO expression in the neuronal tissue.

III.2. Introduction

Translational control comprises a variety of mechanisms responsible for maintenance of homeostasis and for an accurate response of organisms to internal and external stimuli. Regulation of gene expression at this level accounts for quick and reversible changes on global translation or on a subset of selectively targeted messenger RNAs (mRNAs).

mRNAs have, both in the 5' leader sequence and in 3' untranslated region (UTR), evolutionary conserved features that may influence their translational rate and even their stability. Examples are regulatory upstream open reading frames (uORFs), internal

ribosomal entry sites (IRESs) and binding sites for proteins or microRNAs (Sonenberg and Hinnebusch, 2009).

uORFs are regulatory *cis*-acting elements present in the 5’ leader sequence of a transcript. These elements are common to genes that need to be tightly regulated, including oncogenes and genes involved in the control of cellular growth and differentiation (Morris and Geballe, 2000; Wethmar et al., 2010a). Although their presence throughout the genome has been demonstrated, their prevalence has been difficult to calculate (Mignone et al., 2002). The most recent studies estimate that about 49% of the human transcripts contain at least one uORF (Calvo et al., 2009).

In order to be functional, a uORF has to be recognized and translated. Its AUG is recognized by the scanning 40S ribosomal subunit and associated initiation factors depending on the context it is in (Hernández et al., 2010). The optimal context is GCC(A/G)CCAUGG, being the -3 and +4 the most important. An AUG in this context is putatively recognized by all the ribosomes that encountered it. However, differences on this sequence can modulate the strength of the AUG context resulting in the bypass of some or all preinitiation ribosomal complexes altering the translational efficiency of the uORF (Kozak, 2002). This mechanism is called leaky scanning and is also affected by the AUG proximity to the cap site and the presence of nearby secondary structures. Additionally, some uORFs promote ribosome stalling during elongation or termination phases, creating a blockade to additional ribosome scanning (Meijer and Thomas, 2002; Poyry et al., 2004). When the uORF is translated, the 40S ribosomal subunit, along with several initiation factors, can remain associated to the mRNA, resume scanning and reinitiate, at either a proximal or distal AUG codon. Reinitiation efficiency is dependent on the length/time taken to translate a uORF and on the length of the intercistronic region length. The probability of occurring reinitiation is greater when the uORF is short or has a higher rate of translation, because some initiation factors are still associated with the 40S ribosomal subunit, allowing the recognition of a downstream AUG (Kozak, 2001; Poyry et al., 2004; Rajkowitsch et al., 2004). On the other hand, the length of the intercistronic region is important since the eIF2/GTP/Met-$tRNA_i$ ternary complex was used to initiate uORF translation and hence has to be reacquired *de novo.* As a result, the eIF2α is one of the modulators of the reinitiation efficiency (Hinnebusch, 1997; Kozak, 2005; Sachs and Geballe, 2006). In fact, the protein kinases that phosphorylate

eIF2α are activated during stress conditions, resulting in global inhibition of translation (Sonenberg and Hinnebusch, 2009). However, the phosphorylation of eIF2α selectively promotes translational upregulation of a subset of mRNAs that contain uORFs, either by altering leaky scanning or, in the case of a transcript with multiples uORFs, reinitiation efficiency (Palam et al., 2011; Vattem and Wek, 2004; Watatani et al., 2008).

Human erythropoietin (EPO) has been the focus of many studies since it was discovered. Initially, the main function attributed to EPO was the stimulation of erythropoiesis. However, EPO has proven to be a more complex protein, having also non-hematopoietic functions, such as angiogenesis, stimulation of proliferation and anti-apoptosis (Ebert and Bunn, 1999; Gassmann and Soliz, 2009; Maiese et al., 2008). The first site known to produce and secrete EPO was the kidney in the adult. Indeed, it is responsible for the most part of circulating EPO. *EPO* mRNA expression has also been detected in the brain (neurons and glial cells), the lung, the heart, the bone marrow, the spleen, the hair follicles, and the reproductive tract (Dame et al., 1998; Fandrey and Bunn, 1993; Ghezzi and Brines, 2004; Hoch et al., 2011; Weidemann and Johnson, 2009; Yasuda et al., 1998). In these tissues, EPO has anti-inflammatory properties. For all this, EPO is known for its neuro and cardioprotective activities and has been used for the treatment of many disorders, such as cardiac and cerebral ischemia, and Alzheimer's disease (Arabpoor et al., 2012; Casals-Pascual et al., 2009; Undén et al., 2013).

Due to its complexity and differential expression in different organs, we can expect a tight regulation of EPO expression. Actually, EPO is known to be markedly up-regulated by hypoxia (Ebert and Bunn, 1999; Jelkmann, 1992). Both transcriptional and post-transcriptional mechanisms are able to change the expression, in order to increase its levels during stress conditions. Hypoxia inducible factor 1 (HIF1) is the most well-studied factor responsible for the increase of *EPO* transcription during hypoxia (Goldberg et al., 1991; McGary et al., 1997; Semenza, 2001; Wang et al., 1995).

We have previously characterized a translational mechanism controlling expression of EPO protein. We have shown that a 14-codon uORF, present in the 5' leader sequence of the *EPO* transcript, is recognized by the translational machinery, thus negatively affecting EPO expression. Also, we have observed that both leaky scanning and reinitiation are involved in the recognition of the main ORF. However, the uORF acts in

peptide-independent manner and does not trigger NMD. These mechanisms are conserved in liver and kidney cells.

Another conclusion from our work is that *EPO* 3'UTR is able to increase protein levels of the main ORF, nevertheless the mechanism underlying this phenomenon differs from HEK293 cells to HepG2 cells, in which mRNA levels remain unchanged, to REPC cells, in which mRNA levels increase significantly. Yet, when both uORF and 3'UTR are present in the transcrip, they seem to have independent roles on EPO translation. An interesting discovery was that the uAUG of *EPO* uORF is less recognized during hypoxia, increasing EPO production via eIF2α phosphorylation. This was observed only under hypoxia in REPC cells, which stand for a tissue- and stimuli-specific regulation of *EPO* uORF.

Bearing these data in mind and knowing that EPO protein is also expressed in neuronal tissue with neuroprotective functions, we were prompted to analyze whether *EPO* uORF plays a role in the regulation of EPO expression in neuronal tissue. Here, we report that uORF negatively regulates expression of the main ORF in the same extent to what we have observed in other cell lines. However, *EPO* mRNA levels decrease during chemical ischemia, whereas EPO protein expression is maintained, indicating that translational efficiency increases in this tissue.

III.3. Materials and Methods

III.3.1. Plasmid constructs

The pGL2-Luc, pGL2-WT, pGL2-no_AUG, pGL2-no_uSTOP, pGL2-optimal_uAUG, pGL2-frameshift, pGL2-Luc-3'UTR, pGL2-WT-3'UTR and pGL2-no_uAUG-3'UTR constructs were generated previously (Barbosa and Romão, 2013).

III.3.2. Cell culture and plasmid transfection

SW1088 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum. Cells were grown at 37°C in humidified incubator containing 5% CO_2. Transient reverse transfections were performed using Lipofectamine 2000 Transfection Reagent (Invitrogen), following the manufacturer's

instructions, in 35-mm plates. Cells were co-transfected with 750 ng of the test DNA construct corresponding to the pGL2-Luc, pGL2-WT, or its derivative plasmids, and 500 ng of the pRL-TK plasmid (Promega), which encodes *Renilla* luciferase as an internal control, and, then, harvested after 24h. To mimic chemical ischemia, 20h post-transfection, the cultures were changed to fresh medium supplemented with 10 µM 2-deoxy-D-glucose (Calbiochem) and 10 µM sodium azide (Sigma).

III.3.3. Luminometry assay

Lysis was performed in all cell lines with Passive Lysis Buffer (Promega). The cell lysates were used to determine luciferase activity with the Dual-Luciferase Reporter Assay System (Promega) and a Lucy 2 luminometer (Anthos Labtec), according to the manufacturer's standard protocol. One µg of extract was assayed for firefly and *Renilla* luciferase activities. Ratio is the unit of firefly luciferase after normalized with *Renilla* luciferase, and each value was derived from three independent experiments.

III.3.4. RNA isolation

Total RNA from transfected cells was isolated using the Nucleospin RNA extraction II kit (Marcherey-Nagel), following the manufacturer's instructions. Then, all RNA samples were treated with RNase-free DNase I (Ambion) and purified by phenol:chloroform extraction.

III.3.5. Reverse transcription-quantitative PCR (RT-qPCR)

Synthesis of cDNA was carried out using 1µg of total RNA and Superscript II Reverse Transcriptase (Invitrogen), according to the manufacturer's instructions. Real-time PCR was performed in ABI Prism 7000 Sequence Detection System, using SybrGreen Master Mix (Applied Biosystems). Primers specific for the firefly luciferase cDNA and *Renilla* luciferase cDNA were described in chapter II.3.5. Quantification was performed using the relative standard curve method (ΔΔCt, Applied Biosystems). The following cycling parameters were used: 10 min at 95°C and 40 cycles of 15 sec at 95°C and 1 min at 61°C.

Technical triplicates from three to four independent experiments were assessed in all cases.

III.3.6. Statistical analysis

Results are expressed as mean ± standard deviation. Student's *t* test was used for estimation of statistical significance. Significance for statistical analysis was defined as a $p < 0.05$.

III.4. Results

III.4.1. *EPO* uORF represses translation in neuronal cells

The majority of uORFs are regulatory elements with a negative influence on translation of the main ORF (Mignone et al., 2002). Human *EPO* transcript presents a 14-codon uORF conserved among species. We have previously shown that the human *EPO* uORF is functional and that it is able to decrease translation of the main ORF in about 3-fold in HEK293, HepG2 and REPC cell lines that mimic the major sites of production and secretion of EPO, the kidney in the adult and the liver in fetal life (Dame et al., 1998; Jelkmann, 2011). However, the *EPO* transcript has also been detected in other organs such as in neurons and glial cells (Ghezzi and Brines, 2004; Marti et al., 1996). This raised the question whether *EPO* uORF is also repressive in the neuronal cells. To test this hypothesis, the intact human *EPO* 5' leader sequence was cloned into the pGL2 expression vector, flanking the FLuc reporter gene to create the pGL2-WT construct (Figure III.1.A). In addition, the *EPO* uORF was disrupted by site directed mutagenesis of the uAUG (ATG→TTG), using the previous pGL2-WT construct as template, originating the pGL2-no_uAUG construct (Figure III.1.A). Expression of each of these reporter gene constructs was studied in a cell line derived from fibroblasts of the human brain (SW1088). Then, cellular extracts were prepared and assayed for luciferase activity and total RNA was isolated to quantify the relative luciferase mRNA levels by RT-qPCR (Figure III.1.B). FLuc activity of each construct was normalized to the activity units of RLuc expressed from the co-transfected pRL-TK plasmid. The relative luciferase activity

was compared to that of the empty pGL2-Luc vector (Figure III.1.A), arbitrary set as 1 (Figure III.1.B).

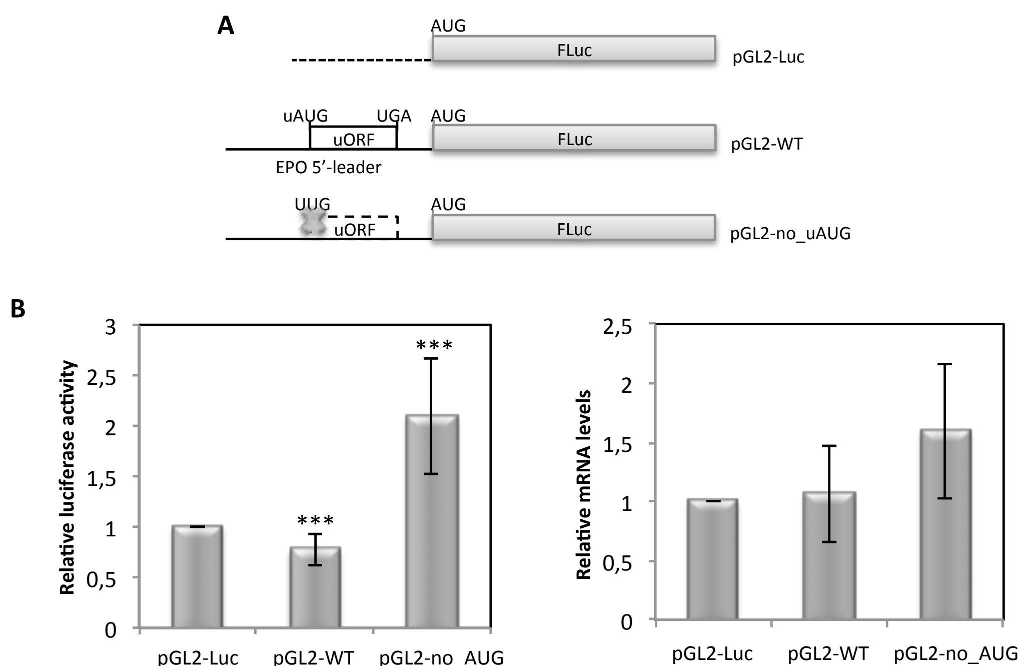


Figure III.1. The *EPO* uORF represses translation of the downstream main ORF in neuronal cells.
(A) Schematic representation of reporter constructs as in figure II.1. The human *EPO* 5' leader sequence encompassing its uORF (open box) with the intact initiation (uAUG) and termination (UGA) codons, was cloned into the empty vector (pGL2-Luc), upstream of the firefly luciferase coding region (FLuc; grey boxes) to create the pGL2-WT construct. In the pGL2-no_uAUG construct, the uORF initiation codon is mutated (AUG→UUG) (the cross represent the point mutation and the dashed lined box represent the non-functional uORF). **(B)** The *EPO* 5' leader sequence represses protein expression of the downstream reporter. SW1088 cells were transiently co-transfected with each one of the constructs described in (A) and with the pRL-TK plasmid encoding the *Renilla* luciferase (RLuc). Cells were lysed twenty-four hours later and the luciferase activity was measured by luminometry assays. FLuc activity values were normalized to RLuc activity to control for transfection efficiency. Relative luciferase activity of the pGL2-Luc was defined as one. In parallel, the luciferase mRNA levels were quantified by RT-qPCR. The FLuc mRNA levels were normalized to those of the RLuc mRNA and analyzed by the ΔΔCt method. The relative pGL2-Luc mRNA levels were also defined as one. Average values and standard deviation (SD) of three independent experiments are shown. Statistical analysis was performed using Student's *t* test (unpaired, two tailed); (∗) $p<0.05$; (∗∗) $p<0.01$; (∗∗∗) $p<0.001$.

Similarly to what we have demonstrated for the other studied cell lines, our results show that, in SW1088 cells, human *EPO* 5' leader sequence with the intact uORF induces a 3-fold repression of translation of the reporter transcript, when compared with the relative luciferase activity from the pGL2-no_uAUG construct with no uORF. Also, and as expected, the relative luciferase mRNA levels are not affected (Figure III.1.B). Thus,

intact *EPO* uORF induces a repression of protein expression at the translational level, in neuronal cells.

III.4.2. The mechanism by which the main ORF is recognized is maintained in liver, kidney and neuronal cells

Leaky scanning and translation reinitiation are the two described mechanisms that allow expression of the main ORF when a functional uORF is present (Morris and Geballe, 2000). Previously, we have observed that both uAUG leaky scanning and reinitiation, after translation of *EPO* uORF, are responsible for the translation of the main ORF in HEK293, HepG2 and REPC cells. Here, we intended to verify whether the same mechanisms act in SW1088. To evaluate these mechanisms we first mutated the stop codon of the uORF (TGA→AGA; pGL2-no_uSTOP construct), creating an extended uORF that terminates at the next in-frame stop codon, 83 nucleotides downstream from the FLuc initiation codon (pGL2-no_uSTOP construct; Figure III.2.A). This mutation allows evaluating the possibility of ribosome leaky scanning since it completely abrogates the possibility of FLuc to be produced by reinitiation after translation of the uORF. In addition, we mutated in the pGL2-WT vector, the context of the uAUG codon (gggAUGa→gccAUGg), to obtain the pGL2-optimal_uAUG construct (Figure III.2.A) with a uAUG sequence context shown by Kozak to yield maximum initiation frequency in higher eukaryotes (Kozak, 1997; Loughran et al., 2012; Wang and Rothnagel, 2004). In this case, the majority of the ribosomes load on the *EPO* mRNA 5' leader sequence are unable to leak past the uAUG codon and most likely they translate the uORF and may reinitiate at the downstream AUG codon. SW1088 cells were transiently co-transfected with pRL-TK and the pGL2-WT, pGL2-no_uSTOP or with pGL2-optimal_uAUG construct and translational efficiencies were monitored by dual luciferase assays, as before. Results were compared to those obtained from the pGL2-WT construct (Figure III.2.B). As shown in Figure III.2.B, mutation of the uORF stop codon (pGL2-no_uSTOP construct) reduces relative luciferase activity to approximately 20% of that of the pGL2-WT construct, without altering the mRNA levels. This suggests that the percentage of ribosomes that leak past the uORF is low and thus translation of the main ORF mostly occurs by reinitiation of the ribosomes after translation termination of the uORF. In fact, analysis

of the pGL2-optimal_uAUG expression allowed us to undestand that translation reinitiation at the main ORF can account for about 50% of relative luciferase activity, in comparison to the relative luciferase activity of the pGL2-WT construct, while the mRNA levels stay unchanged (Figure III.2.B).

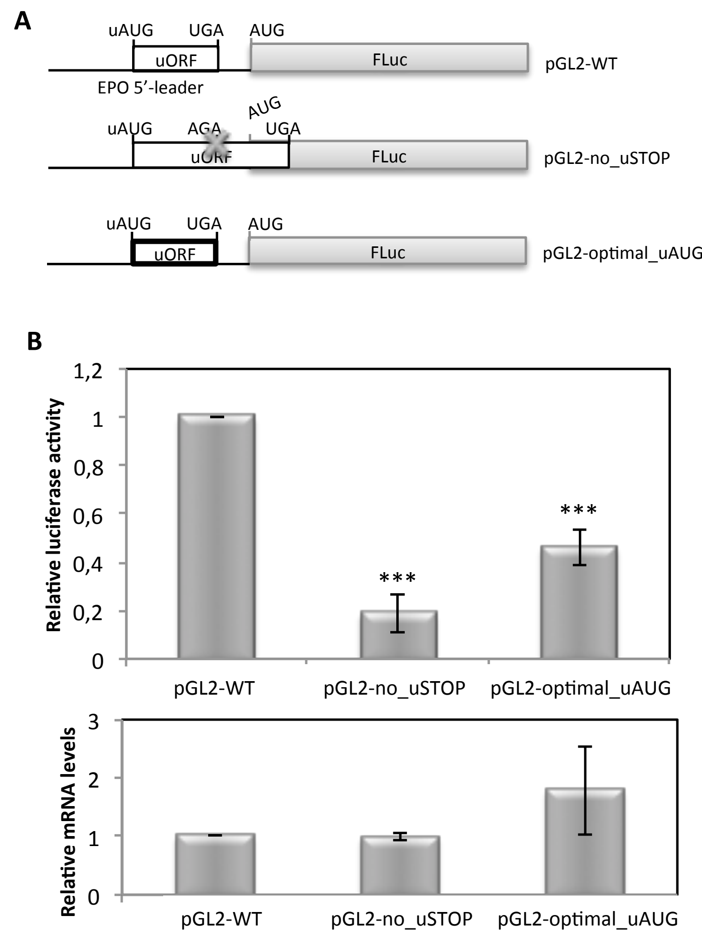


Figure III.2. Both translation reinitiation and uAUG leaky scanning are involved in the translational initiation at the main AUG codon.

(A) Schematic representation of reporter constructs, as in figure II.2. The pGL2-WT plasmid contains the wild-type human *EPO* 5' leader sequence, the pGL2-no_uSTOP construct presents the *EPO* 5' leader sequence with a mutation (UGA→AGA) at the uORF translation termination codon, which makes the uORF to overlap with the luciferase ORF (the cross represent the point mutation), and the pGL2-optimal_uAUG contains the *EPO* 5' leader sequence with a optimal uAUG sequence context (gggAUGa→gccAUGg; represented by a bold lined box). **(B)** SW1088 cells were transiently co-transfected with each one of the constructs described in (A) and with a plasmid encoding *Renilla* luciferase (pRL-TK) and analyzed as described in the legend to Figure III.1.B.

III.4.3. In neuronal cells, the translational machinery is not blocked by the *EPO* uORF-encoded peptide

Some uORFs have the ability to induce a blockade of the translational machinery increasing their inhibitory effect. These uORFs function in a peptide-dependent manner (Karagyozov et al., 2008; Wei et al., 2012). The sequence of the *EPO* uORF-encoded peptide is conserved among mammalian species, indicating a putative function of this region. However, when we frameshifted the nucleotide sequence of the *EPO* uORF, in order to produce a different peptide sequence, no blockade of the translational machinery in HEK293, HepG2 and REPC cell lines was observed. In spite of that, we have investigated whether this is preserved in the cell model used in this study. For that, we used the pGL2-frameshift construct (Figure III.3.A.) in which the uORF was modified by shifting the reading frame to generate a different amino acid sequence while preserving the uAUG context and most of the nucleotide sequence. The pGL2-frameshift construct was used to transiently transfect SW1088 cells. The corresponding relative luciferase activity and mRNA accumulation levels were analyzed as previously and the results were compared to those of the pGL2-WT construct (Figure III.3.B.). Our data show that the mutant uORF of pGL2-frameshift construct decreases the relative luciferase activity but not the mRNA levels (Figure III.3.B.). In chapter II, we propose that this is due to a decrease of the reinitiation efficiency, which might be a result of a rare codon in the altered uORF of pGL2-frameshift construct that might increase the time of translation. In this way, we can conclude that the native *EPO* uORF functions in a peptide sequence-independent manner in SW1088, which supports what we have observed in HEK293, HepG2 and REPC cell lines.

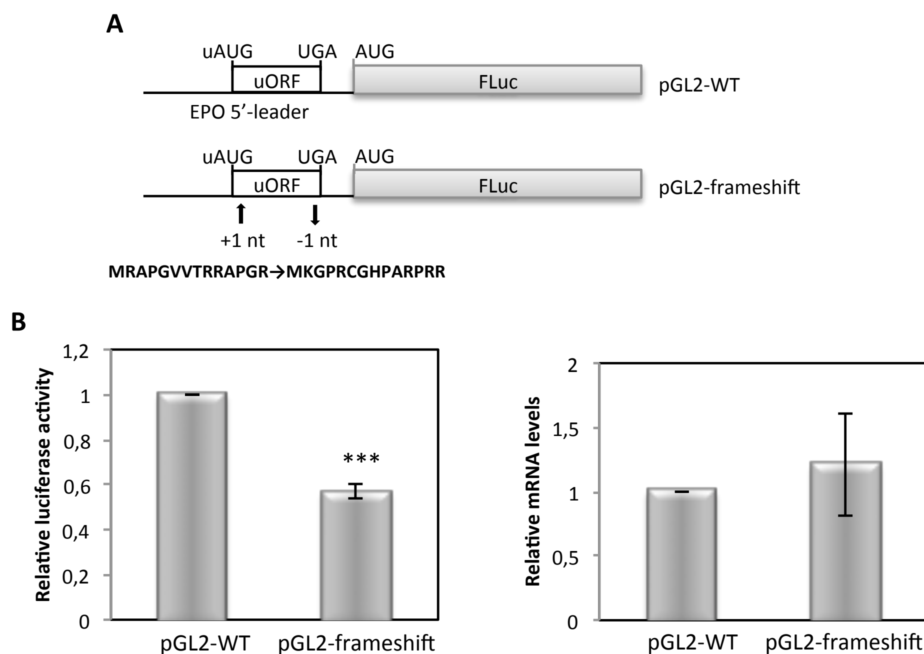


Figure III.3. In neuronal cells, the translational repression exerted by the *EPO* uORF is peptide sequence-independent.

(A) Schematic representation of the expression constructs, as in figure II.3. The pGL2-WT plasmid contains the human normal *EPO* 5' leader transcript sequence, the pGL2-frameshift vector carries a EPO uORF sequence modified by frameshift mutations, which consist in the insertion of one nucleotide in the second codon (+1 nt) and the deletion of one nucleotide in 13th codon (-1 nt). The resulting uORF-encoded peptide sequence is shown below. **(B)** SW1088 cells were transiently co-transfected with each one of the constructs described in (A) and with a plasmid encoding *Renilla* luciferase (pRL-TK) and analyzed as described in the legend to Figure III.1.B.

III.4.4. In neuronal cells, *EPO* 3'UTR has no impact on the inhibitory effect of the uORF

Circularization of the mRNA brings in close proximity the 5' leader sequence and the 3'UTR of a transcript. Indeed, some examples have reported that these two structures can interact with each other, altering the translational repression exerted by a uORF present in a transcript (Czyzyk-Krzeska and Bendixen, 1999; McGary et al., 1997; Medenbach et al., 2011). Since the *EPO* 3'UTR seems to be recognized by several proteins that regulate mRNA stability (Czyzyk-Krzeska and Bendixen, 1999; McGary et al., 1997; Ohigashi et al., 1999; Wang et al., 1995), we analysed whether the 3'UTR could in fact impact the repressive effect of the *EPO* uORF. In HEK293, HepG2 and REPC cell lines, we reported that the *EPO* 3'UTR alone is able to increase protein levels in the same extent. However, mRNA levels were different between HEK293 and HepG2, and

REPC cell lines, since in REPC cells the 3'UTR is able to increase the steady-state levels of mRNA. Based on these data, we aimed to prove the effect of the *EPO* 3'UTR in SW1088 cell line. The previously cloned pGL2-Luc-3'UTR construct (Figure III.4.A) was transiently transfected into SW1088 cells. Firefly luciferase activity was normalized to the activity units from co-transfected *Renilla* luciferase reporter construct, as before, and the relative luciferase activity of the pGL2-Luc-3'UTR was compared to that of the empty pGL2-Luc construct. The results show that the *EPO* 3'UTR alone induces about a 5-fold increase in relative luciferase activity, when compared to the relative luciferase activity of the pGL2-Luc control, whereas the mRNA levels remain unaltered (Figure III.4.B). Thus, the *EPO* 3'UTR-containing construct in SW1088 cell line has the same effect observed for HEK293 and HepG2 cell lines, highlighting the differential regulation of these structures in the REPC cell line.

Our results shown, nonetheless, that in HEK293, HepG2 and REPC cells the *EPO* uORF retains its repressive impact on the main ORF translation even in the presence of 3'UTR. To investigate whether the same mechanism of regulation occurs in SW1088 we have monitored the relative luciferase activity of the pGL2-WT reporter harboring both *EPO* 5' and 3'UTRs (pGL2-WT-3'UTR construct; Figure III.4.C), and we have compared it to the relative luciferase activity of the corresponding construct with the disrupted uORF (pGL2-no_uAUG-3'UTR; Figure III.5.C). To do that, each of these constructs was co-transfected with pRL-TK into SW1088 cells, as above, and luciferase activities and mRNA levels were obtained, as previously (Figure III.4.D). The results show that the insertion of the *EPO* 3'UTR into the construct pGL2-WT does not abrogate the ability of the *EPO* uORF to inhibit reporter translation (Figure III.4.D). Indeed, the intact *EPO* 5' leader sequence in the pGL2-WT-3'UTR construct allows a significant 3-fold decrease in relative luciferase activity when compared to that observed from the pGL2-no_uAUG-3'UTR construct with the disrupted uORF, while the relative mRNA levels remain unaltered (Figure III.4.D). Thus, the *EPO* 3'UTR fails to overcome translational repression induced by the *EPO* uORF also in the SW1088 cell line.

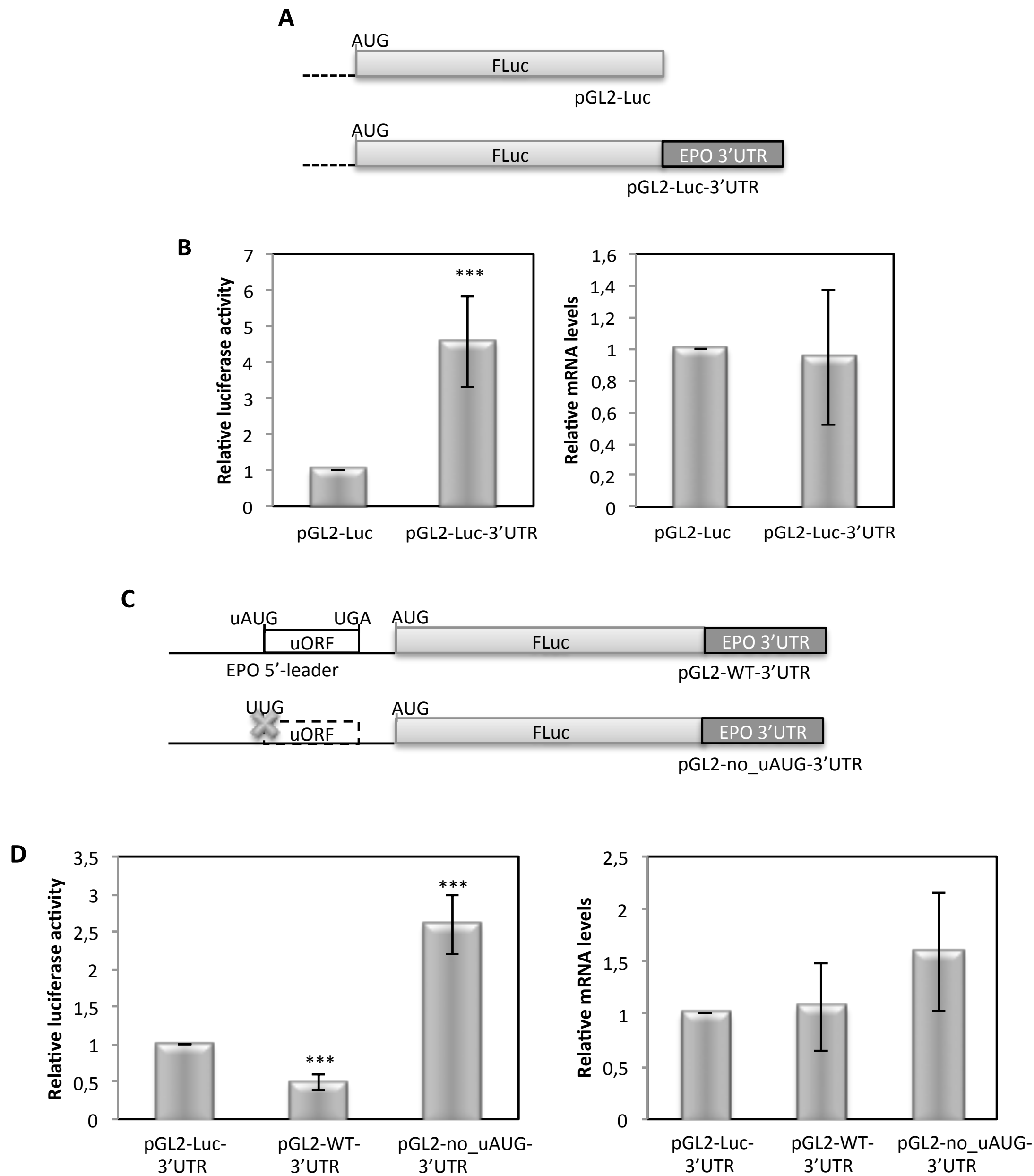


Figure III.4. In neuronal cells, the 3'UTR of the *EPO* mRNA has no influence in the inhibitory effect of the uORF.

(A) Schematic of the firefly luciferase (FLuc) reporter constructs containing the native luciferase 3'UTR (pGL2-Luc) or the 3'UTR sequence (dark grey box) of the human *EPO* transcript (pGL2-Luc-3'UTR). **(B)** SW1088 cells were transiently co-transfected with each one of the constructs described in (A) and with a plasmid encoding *Renilla* luciferase (pRL-TK) and analyzed as described in the legend to Figure 2B. **(C)** Schematic of the firefly luciferase (FLuc) reporter constructs containing the human *EPO* 5' leader sequence with the intact uORF and the 3'UTR sequence (dark grey box) of the human *EPO* transcript (pGL2-WT-3'UTR), or the *EPO* 5' leader sequence with a disrupted uORF due to the uAUG→UUG mutation (represented by a cross) and the EPO 3'UTR sequence (dark grey box; pGL2-no_uAUG-3'UTR). **(D)** SW1088 cells were transiently co-transfected with each one of the constructs described in (C) and with a plasmid encoding *Renilla* luciferase (pRL-TK) and analyzed as described in the legend to Figure III.1.B.

III.4.5. The repressive effect of the *EPO* uORF is inhibited during chemical ischemia

Stress conditions can lead to dramatic changes on the overall protein synthesis. In general, there is a global decrease on protein synthesis, but there is growing evidence that mRNAs can be specifically controlled in order to alter their expression patterns. Selective subsets of mRNAs that are shown to overcome this global pressure present in their sequence regulatory elements such as uORFs and IRES (Blais et al., 2004; Le Quesne et al., 2010; Yaman et al., 2003). We have demonstrated that the *EPO* uORF does not mediate any response to hypoxia neither to nutrient starvation in both HEK293 and HepG2 cells. However, in REPC cells we observed a derepression of the uORF negative effect in response to hypoxia but not in response to nutrient deprivation. In our search for the mechanisms involved in this effect we have shown that there is more ribosomes leaking past the uAUG and that the phosphorylation of eIF2α facilitates this bypass. In order to understand the neuronal relevance of the *EPO* uORF, we have decided to assess whether this structure alters the main ORF expression in response to ischemic conditions. For that, SW1088 cells were transiently transfected with the pGL2-WT and pGL2-no_uAUG constructs that carry the intact or disrupted *EPO* uORF, respectively; then, cells were treated with 10µM of 2-deoxy-D-glucose and 10 µM of sodium azide, to induce chemical ischemia. Six hours later, cells were lysed and protein and RNA were extracted and analysed by luciferase assays and RT-qPCR, as previously described. Our results show that the relative luciferase activity of pGL2-WT construct, when compared to the relative luciferase activity of the pGL2-no_uAUG construct, is not significantly altered in response to chemical ischemia in SW1088 cells (Figure III.5.A). However, we have observed a dramatic and significant decrease of the relative mRNA levels of the pGL2-WT construct during chemical ischemia (Fig. III.5.B). Furthermore, the normalization of the relative luciferase activity of the pGL2-WT construct to its corresponding relative mRNA levels revealed a 4-fold increase under chemical ischemia (Fig. III.5.C). This means that each mRNA molecule is translated with higher efficiency under chemical ischemia in an *EPO* uORF-dependent manner.

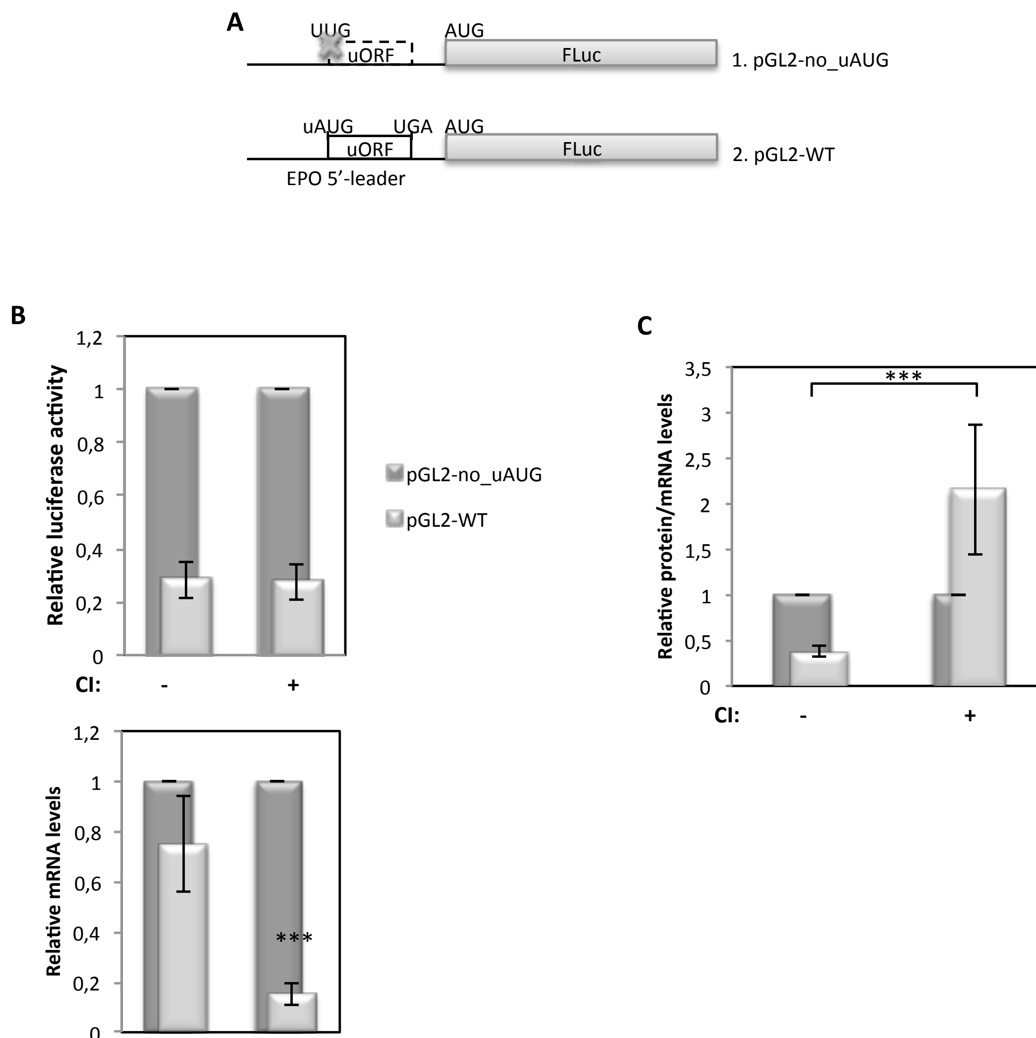


Figure III.5. EPO relative protein levels are enhanced in SW1088 cells in response to chemical ischemia. (A) Schematic representation of the pGL-WT (construct 1) and pGL2-no_uAUG (construct 2) vectors represented as in Figure III.1. These constructs were separately co-transfected with a plasmid encoding *Renilla* luciferase (pRL-TK) in SW1088 cells. Twenty hours later cells were untreated (-) or treated (+) with 10 µM 2-deoxy-D-glucose and 10 µM sodium azide. **(B)** Untreated (CI: -) and treated (CI: +) transfected cells were lysed and analyzed as described in the legend to Figure III.1.B. The dark bars correspond to the pGL2-no_uAUG construct and the light bars to the pGL2-WT. **(C)** Relative protein levels were normalized to the corresponding relative mRNA levels for the pGL2-no_uAUG construct (dark bars) and for the pGL2-WT (light bars) in untreated (CI: -) and treated (CI: +) transfected cells

III.5. Discussion

EPO is a complex protein that needs to be tightly regulated. EPO regulates the proliferation, differentiation and death of the erythroid cells (Fandrey, 2004). Due to its ability to promote cell survival and differentiation it acts in non-hematopoietic cells

modulating proliferation and cellular viability (Bunn, 2013; Maiese et al., 2008). Although, the major site of EPO production is the kidney in the adult, many organs have been described to express the *EPO* mRNA (Dame et al., 2001; Fandrey and Bunn, 1993; Hoch et al., 2011; Yasuda et al., 1998). In those organs, EPO seems to act locally not contributing to erythropoiesis. Since EPO is expressed in cardiac and neuronal cells it has been described as cardio and neuroprotective. Moreover, EPO is regulated at multiple levels to ensure its correct response to external stimuli in different tissues. *EPO* gene expression is best studied at transcriptional level.

Promoter silencing in the adult liver leads to a change of the site of production from the liver to the kidney (Dame et al., 2004). Many other mechanisms control EPO expression, such transcriptional activation of *EPO* gene by HIF1 during hypoxic conditions (Goldberg et al., 1991; Imagawa et al., 1991; Wang and Semenza, 1993; Warnecke et al., 2004). The higher levels of circulating EPO, as a result of this stimulation, increases the red blood cell mass in order to rise the oxygen-carrying capacity of the blood (Bunn, 2013). Although, this hypoxic activation has been described mainly in the kidney and liver, there is also an oxygen-dependent regulation of EPO expression in the brain (Marti et al., 1996). This fact, together with the neuronal expression of EPO and EPO receptor (EPOR) and the protection effects of this signalling pathway during stroke, brain injury or cerebral ischemia, suggests a paracrine function of EPO in the neuronal tissue and a regulation of EPO expression in these cells (Bunn, 2013; Chong et al., 2005; Ryou et al., 2012).

Previously we have shown that the *EPO* transcript presents in its 5’ leader sequence a highly conserved 14-codon uORF (chapter II) that acts as a negative regulatory element able to decrease the expression of the main ORF in about 3-fold in kidney and liver model cell lines (HEK293, HepG2 and REPC cell lines). Since, the EPO protein is expressed in the neuronal tissue and seems to have specific neuroprotective functions, we aimed to study *EPO* uORF-mediated regulation in these cells lines. We found that the *EPO* uORF is functional also in neuronal cell lines decreasing the main ORF expression in about 3-fold (Figure III.1.) as seen in kidney and liver cells, which suggests that the repressive effect of the *EPO* uORF is maintained in the studied tissues.

Leaky scanning and reinitiation are the two mechanisms capable of promoting the expression of the main ORF when a functional uORF is present (Geballe and Morris,

1994). We have reported that both mechanisms are implicated in the recognition of the *EPO* AUG and here, we observed, that these mechanisms are preserved in neuronal cells (Figure III.2.). We have also shown that the repressive effect exerted by the *EPO* uORF in neuronal as well as in liver and kidney cells does not depend on the encoded peptide.

One interesting finding of our prior study was the influence of the *EPO* 3'UTR on the main ORF expression. The *EPO* 3'UTR increases the protein expression in all cell lines but only in REPC cells these levels are a result of the increased mRNA levels. The regulation of the *EPO* mRNA stability due to the binding of several proteins was already described (Czyzyk-Krzeska and Bendixen, 1999; Madan et al., 1995; McGary et al., 1997), leading us to analyse whether this effect was also maintained in the present model. We observed that similarly to what happened in HEK293 and HepG2 cells the *EPO* 3'UTR increases protein levels, maintaining the mRNA levels unaltered (Figure III.4.A and B). This highlights the tissue specific regulation by this structure on REPC cells. However, *EPO* 3'UTR does not affect the repressive ability of the uORF (Figure III.4.C and D). Taking this into consideration, we conclude that these two elements influence EPO translation independently.

As EPO is a pleiotropic protein that responds to different cell stress stimuli and tissue injuries (Arcasoy, 2008; Brines et al., 2008; Ruifrok et al., 2008; Ryou et al., 2012), we were prompted to verify whether the *EPO* uORF repression was relieved during cerebral ischemia. For that, we used chemical ischemia to stimulate the neuronal cells and observed that protein levels are the same under both normal and stressed conditions; however, there is a sharp decrease on the mRNA levels during chemical ischemia stimulus when the *EPO* uORF is functional (Figure III.5.). This means that each molecule of mRNA is more efficiently translated after ischemia, leading to an increase of about 4-fold of the translation efficiency (Figure III.5.C). Since translation of the main ORF is more effective, it might indicate that *EPO* uORF repression is abrogated. This suggests the existence of another completely different mechanism for translational regulation from the one observed in REPC cells (Chapter II).

This is a striking result since no other study have yet reported this type of mRNA control by the uORF under stress conditions. In the future it would be interesting to determine whether the decreased mRNA levels are due to transcriptional inhibition or due to a

decrease in the mRNA stability. Also, we would like to inquire whether phosphorylated eIF2α is mediating uORF derepression as seen before.

Overall, we have explained the regulatory mechanism of the human *EPO* uORF in neuronal tissue. The fact that the basic mechanism is preserved reveals that this structure thoroughly regulated the human *EPO* expression, although its response to chemical ischemia shows a different regulatory mechanism.

Further studies might bring new insights on the modulation of human *EPO* expression, particularly in the brain, encouraging the development of new forms of therapy for many neurodegenerative diseases.

III.6. Acknowledgements

We are grateful to Margarida Gama Carvalho, for supplying the pRL-TK plasmid, and to the Oncology group at *Instituto Nacional de Saúde Dr. Ricardo Jorge* for providing the SW1088 cell line. We would also like to thank Isabel Peixeiro, Cláudia Onofre, João Lavinha e Rafaela Lacerda for critical reading of the manuscript. This research was partially supported by Fundação para a Ciência e a Tecnologia (PEst-OE/BIA/UI4046/2011, PTDC/BIM-MED/0352/2012 and SFRH/BD/63581/2009 to C.B.).

CHAPTER IV – The translation reinitiation mechanism of the human erythropoietin transcript

Author's note

This chapter arose from an on-going work in our lab already published, where my contribution in stated:

Peixeiro I, Inácio A, **Barbosa C**, Silva AL, Liebhaber SA and Romão L. (2012) Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations. Nucleic Acids Research *40*, 1160–1173. doi:10.1093/nar/gkr820.

IV.1. Abstract

Eukaryotic initiation factor 3 (eIF3) is a protein complex composed of 13 subunits. Due to the interaction of specific subunits with several other factors and ribosomal subunits it can impact translation initiation, termination, recycling and translation deregulation. Upstream open reading frames (uORFs) are *cis*-acting elements present in the 5' leader sequence of the transcript that negatively regulate the expression of the main ORF. However, after translation of a small uORF, the 40S ribosomal subunit might remain associated with the mRNA, thus resuming scanning and subsequent translation of a downstream ORF. Multiple studies in yeast and plants have shown that eIF3 is involved in translation. The human erythropoietin (*EPO*) transcript has a conserved and functional 14-codon uORF that allows reinitiation to a certain extent. Here, we have used this uORF as an experimental model to study the molecular basis of reinitiation efficiency after its translation.

In this way, we have analyzed the effect of the *EPO* uORF and how different eIF3 subunits contribute to the reinitiation mechanism. Our results demonstrate that reinitiation efficiency is directly related to the size of the *EPO* uORF. In addition, depletion of eIF3h, f, and e subunits decrease translation of the main ORF due to reinitiation after *EPO* uORF translation. However, and contrary to what we expected, eIF3a and c have no impact on reinitiation. Our data contribute to the clarification of the basis of the reinitiation mechanism in mammalian cells.

IV.2. Introduction

Translation initiation is a rate-limiting step that involves several proteins, the eukaryotic initiation factors (eIFs). During this process, the eukaryotic initiation factor 4F (eIF4F) complex binds to the 5' end of the mRNA. eIF4F encompasses the cap-binding protein eIF4E, the helicase eIF4A and eIF4G, a scaffolding protein with a binding site for eIF4E and for poly(A)-binding protein (PABP), resulting in mRNA circularization (Holcik and Pestova, 2007; Morino et al., 2000; Sonenberg and Hinnebusch, 2009).

The 43S preinitiation complex includes the small ribosomal subunit, the eukaryotic initiation factors 1, 1A and 3, and the ternary complex eIF2-GTP-Met-tRNA$_i^{Met}$ (Gebauer

and Hentze, 2004; Sonenberg and Hinnebusch, 2009). It is recruited to the 5' end of the mRNA and scans in a 5' to 3' direction until an AUG is recognized by the anticodon of Met-tRNA$_i^{Met}$, in a process involving the concerted action of eIFs 1, 1A, 2 and 5. In this step there is the release of eIF2-GDP and probably other 40S-bound eIFs. After this release eIF5B catalyzes the recruitment of the 60S ribosomal subunit, forming the 80S ribosome, and elongation can start (Gebauer and Hentze, 2004; Holcik and Pestova, 2007; Kozak, 1999; Sonenberg and Hinnebusch, 2009).

The eIF3 is one important factor that can act in almost all the steps of translation serving as a target for translational control. It is involved not only in translation initiation but also in termination phase, where it is implicated in the dissociation of the translational machinery, and also in the recycling and reinitiation mechanisms ((Hinnebusch, 2006; Pisarev et al., 2007; Roy et al., 2010; Szamecz et al., 2008), thus being considered a good candidate for the regulation of the overall outcome of translation. The 750 kDa-eIF3 is the most complex initiation factor comprising 13 non-identical subunits designated from eIF3a to eIF3m in mammalian cells (Herrmannová et al., 2012; Hinnebusch, 2006). Many studies have tried to reassemble this complex in mammalian cells, but, due to its complexity, its actual composition is still poorly understood. In budding yeast, eIF3 comprises five core essential subunits – a, b, c, g and i – and one noncore subunit – eIF3j (Herrmannová et al., 2012). Although, the mammalian eIF3 includes all the corresponding orthologs found in yeast, the presence of seven additional subunits emphasizes its higher complexity. *In vitro* studies suggested that the functional core of the mammalian eIF3 comprises subunits eIF3a, b, c, e, f and h (Masutani et al., 2007). Conversely other study based on tandem mass spectrometry and solution disruption assays identified three stable modules: one composed of a, b, i, and g subunits, resembling the yeast eIF3 core: a second one including subunits c, d, e, l, and k; and the third one consisting of subunits f, h, and m (Zhou et al., 2008b).

The subunit-subunit network of the mammalian eIF3 and its interaction with other proteins still needs further clarification. Yet it is known that eIF3 interacts with eIF4G through eIF3e and eIF3f (LeFebvre et al., 2006; Masutani et al., 2013) and that it also contacts with the 40S ribosomal subunit through eIF3a, b, c and j (Fraser et al., 2007; Hinnebusch, 2006). Additionally, it promotes mRNA recruitment, assembly of the

preinitiation complex, and translation initiation (Chiu et al., 2010; Hinnebusch, 2006; Sokabe et al., 2011; Valásek, 2012).

Recently, we provided evidence that eIF3 is also implicated in the inhibition of nonsense-mediated mRNA decay (NMD). NMD is a surveillance mechanism that degrades transcripts bearing premature translation termination codons (PTCs). Our data revealed that human eIF3h and eIF3f subunits are involved in the efficient translation termination required for the NMD-resistance of mRNAs containing PTCs in close proximity to the corresponding AUG codon. This suggests that these subunits might be bridging the interaction amongst poly(A)-binding protein cytoplasmic 1 (PABPC1), eIF4G and the ribosome, into the vicinity of this PTC. On the contrary, our results show that eIF3e has the opposite function, and may be required for NMD-commitment (Peixeiro et al., 2012).

In addition to all of these functions, eIF3 has been also implicated in translation reinitiation. Typically, translation reinitiation is thought to be an ineffective mechanism that occurs after translation of a short upstream open reading frame (uORF) (Meijer and Thomas, 2002). In this case, after the translation termination step the 40S ribosomal subunit can remain associated with the mRNA, resume scanning, and initiate translation at a downstream AUG (Kozak, 2001). Reinitiation is dependent on (i) the time required for the uORF translation, which is determined by the relative length of the uORF and the translation elongation rate; (ii) the translation initiation factors involved in the translation initiation event; and (iii) the length of the intercistronic region (Kozak, 2002; Poyry et al., 2004). A key factor for translation reinitiation is the reacquisition of a new ternary complex (eIF2-GTP-Met-$tRNA_i$) so that the ribosome can recognize a further downstream AUG (Kozak, 2005). Also, several initiation factors need to remain associated with the ribosome during translation and even after the termination event so that reinitiation can occur (Child et al., 1999; Roy et al., 2010). eIF3 is a good candidate to remain associated to the ribosome during the elongation step, and even after termination, since it is bound to the solvent side of the 40S subunit, suggesting that its dissociation is not essential for subunit joining prior to elongation (Szamecz et al., 2008; Valásek et al., 2002). Supporting this idea, in yeast, eIF3 remains associated during several rounds of elongation and enhances translation reinitiation (Szamecz et al., 2008). Furthermore, it has been shown that eIF3a and g are implicated in this process in yeast

(Cuchalová et al., 2010; Szamecz et al., 2008; Valásek et al., 2002) and eIF3h subunit promotes reinitiation after uORF translation in plants (Roy et al., 2010).

Several other subunits may be involved in reinitiation efficiency, such as, for instance, eIF3c, which contacts directly with eIF1 and 5, thus serving as a critical regulator of AUG recognition. Consequently, its maintenance during elongation and 40S subunit scanning after termination can be essential for recognition of the downstream AUG (Karásková et al., 2012; Valásek et al., 2002).

Bearing in mind the importance of eIF3 for reinitiation efficiency, here, we aim to study the reinitiation event after translation of the erythropoietin (*EPO*) uORF. The *EPO* uORF is totally located on the 5’ leader of the transcript, it is composed of 14 codons and its termination codon is 22 nucleotides upstream of the *EPO* initiation translation site. We previously shown that this uORF is functional and that reinitiation accounts for about 60% of the main ORF translation. Consequently, we decided to investigate the features that modulate reinitiation efficiency, specifically how length of the uORF and the presence of eIF3 are implicated in reinitiation after translation of the *EPO* uORF.

Our data indicate that reinitiation efficiency depends on the length of the uORF. Moreover, depletion of eIF3h, f, and e affects translation reinitiation, suggesting a role for these subunits in this process. However, depletion of eIF3a and c had no effect on the translation of the downstream ORF.

IV.3. Materials and Methods

IV.3.1. Plasmid constructs

The constructs pGL2-no_uAUG and pGL2-WT were described previously (Barbosa and Romão, 2013). The pGL2-39codons construct was obtained by introducing a 75bp nucleotide sequence from the ampilicin resistance gene into the ApaI restriction site of the uORF sequence resulting in a 39-codon uORF.

IV.3.2. Cell culture, plasmid and siRNA transfection

HeLa cells were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum. The short RNA interference (siRNA) duplexes (Table III.2) were designed as 19-mers with 3'-dTdT overhangs and purchased from Thermo. For the Luciferase assay transfections of cells with siRNAs were carried out using Lipofectamine 2000 Transfection Reagent (Invitrogen), following the manufacturer's instructions, in 35-mm plates using 200 pmol of siRNA oligonucleotides and 4µl of transfection reagent. Twenty-four hours later, 750 ng of pGL2-no_AUG, pGL2-WT or pGL2-39codons were co--transfected with 750ng of the pRL-TK plasmid.

Table IV.1. Sequences of the siRNAs used in the current work.

siRNA	Sequence (5' → 3')	References
eIF3h	ACUGCCCAAGGAUCUCUCU	(Peixeiro et al., 2012)
eIF3f	GUGAAGGAGAAAUGGGUUU	(Peixeiro et al., 2012)
eIF3e	CCAGGGAUGGUAGGAUGCU	(Peixeiro et al., 2012)
eIF3a	CGAACCAAUUAUGUUGAAA	(Xu et al., 2012)
eIF3c	UGACCUAGAGGACUAUCUU	(Choe et al., 2012)
GFP	GGCUACGUCCAGGAGCGCAC	

IV.3.3. RNA isolation

Total RNA from transfected cells was prepared using the Nucleospin RNA extraction II (Marcherey-Nagel) following the manufacturer's instructions.

IV.3.4. Semi-quantitative RT-PCR

1000 ng of total mRNA were reverse-transcribed with Superscript II Reverse Transcriptase (Invitrogen) according to the manufacturer's standard protocol and using 250 ng of Random Primers (Invitrogen) in a final volume of 20 µl. The PCR reactions for eIF3e, eIF3a or eIF3c and histone deacetylase 1 (*HDAC1*) cDNAs were performed in parallel at similar conditions: 3 µl of the RT product was amplified in a 50-µl reaction volume using 0.2 mM dNTPs, 1.5 mM $MgCl_2$, 15 pmol of each primer (primers #1 and #2 for eIF3e, primers #3 and #4 for eIF3a, primers #5 and #6 for eIF3c and primers #7 and #8 for HDAC1; Table IV.2), 0.75 U of Amplitaq (Promega), and 1X PCR buffer (Promega).

Thermocycler conditions were 95°C for 4 min followed by 26 cycles of 95°C for 45 sec, 56°C for 45 sec, and 72°C for 45 sec followed by a final extension of 72°C for 10 min. Ten-microliter aliquots from each RT-PCR sample were analyzed by electrophoresis on 1.8% agarose gels.

Table IV.2. DNA oligonucleotides used in the current work.

Primer	Sequence (5' → 3')
#1	GGACAAGCATGGTTTTAGGCA
#2	TGCTGCTCCTGAGTAATTCCC
#3	ACAGGCAGTGTTTGGAC
#4	GAGAATAGCCCGTGAATA
#5	ACCAAGAGAGTTGTCCGCAGT
#6	TCATGGCATTACGGATGGTCC
#7	ATGGCGCAGACGCAGGG
#8	CCGCACTAGGCTGGAACATC

IV.3.5. Dual luciferase assay

Co-transfected HeLa cells were lysed with Passive lysis buffer (Promega) and luminescence was measured in Lucy 2 Luminometer (Anthos Labtec) with the Dual Luciferase Assay System (Promega) according to the manufacturer's indications.

IV.3.8. SDS-PAGE and Western blotting

Protein lysates were resolved, according to standard protocols, in 10% SDS-PAGE and transferred to PVDF membranes (Bio-Rad). Membranes were probed using mouse monoclonal anti-α-tubulin (Sigma) at 1:10000 dilution (as a loading control), goat polyclonal anti-hUPF1 (Bethyl Labs), rabbit monoclonal anti-eIF3h (Cell Signaling) and rabbit monoclonal anti-eIF3f (Abcam), at 1:500 dilution. Detection was carried out using secondary peroxidase-conjugated anti-mouse IgG (Bio-Rad), anti-rabbit IgG (Bio-Rad) or anti-goat IgG (Sigma) antibodies followed by chemiluminescence.

IV.4. Results

IV.4.1. The size of *EPO* uORF influences translation reinitiation efficiency

Translation reinitiation depends on the time taken to translate the uORF. Thus, if there are no major differences on translation rates, a shorter uORF will retain more ability to reinitiate then a longer uORF (Kozak, 2002; Poyry et al., 2004). To test whether reinitiation efficiency after translation of *EPO* uORF depends on its length, we have generated a pGL2-39codons construct by introducing the nucleotide sequence from the ampilicin resistance gene into the *EPO* uORF from the pGL2-WT construct (see chapter II). Then pGL2-39codons, pGL2-WT or pGL2-no_uAUG (previously described in chapter II; Figure IV.1.A) were co-transfected with pRL-TK plasmid that expresses the RLuc and serves as an internal control. Then, cellular extracts were prepared and assayed for luciferase activity (Figure IV.1.B.). FLuc activity of each construct was normalized to the activity units from RLuc. The relative luciferase activity was compared to that of the empty pGL2-no_uAUG vector, arbitrary set to 1 (Figure IV.1.B.).

Our results show that the luciferase activity obtained from the pGL2-WT construct is lower than the one from pGL2-no_uAUG (Figure IV.1.B), demonstrating the inhibitory effect of the *EPO* uORF (Figure IV.1.B and chapter II). Additionally, relative protein levels given by pGL2-39codons construct are significantly lower when compared to those from pGL2-no_uAUG and pGL2-WT (Figure IV.1.B). This implies that the uORF with 39 codons is even more repressive then the one with 14 codons. Assuming that leaky scanning past the uAUG is maintained, since its context is not altered, we propose that it is due to lower reinitiation efficiency caused by longer uORFs.

NMD is a mechanism that can regulate the steady-state level of a set of wild-type transcripts, such as those presenting uORFs in their 5’ leader sequence (Mendell et al., 2004; Wittmann et al., 2006; Yepiskoposyan et al., 2011). In analogy to what is seen with transcripts carrying PTCs, small uORFs are NMD-resistant due to the proximity of the uORF stop codon to the uAUG. On the other hand, transcripts with longer uORFs are NMD-sensitive, similarly to transcripts carrying a PTC in a more distal position (Inácio et al., 2004; Peixeiro et al., 2012). These data raised the question whether low expression levels of pGL2-39codons construct could be explained by NMD-triggering. To test this

hypothesis, we used short interfering RNA (siRNA)-mediated depletion of UPF1 in HeLa cells. All results were compared to those obtained in NMD-competent cells transfected with nonspecific control (GFP) siRNAs. Twenty-four hours after siRNA transfection cells were transiently co-transfected with the pRL-TK plasmid and each reporter pGL2-no_uAUG, pGL2-WT and pGL2-39codons. The extracts obtained twenty-four hours after plasmid transfection were used to monitor the endogenous levels of UPF1 and to measure the relative luciferase activity.

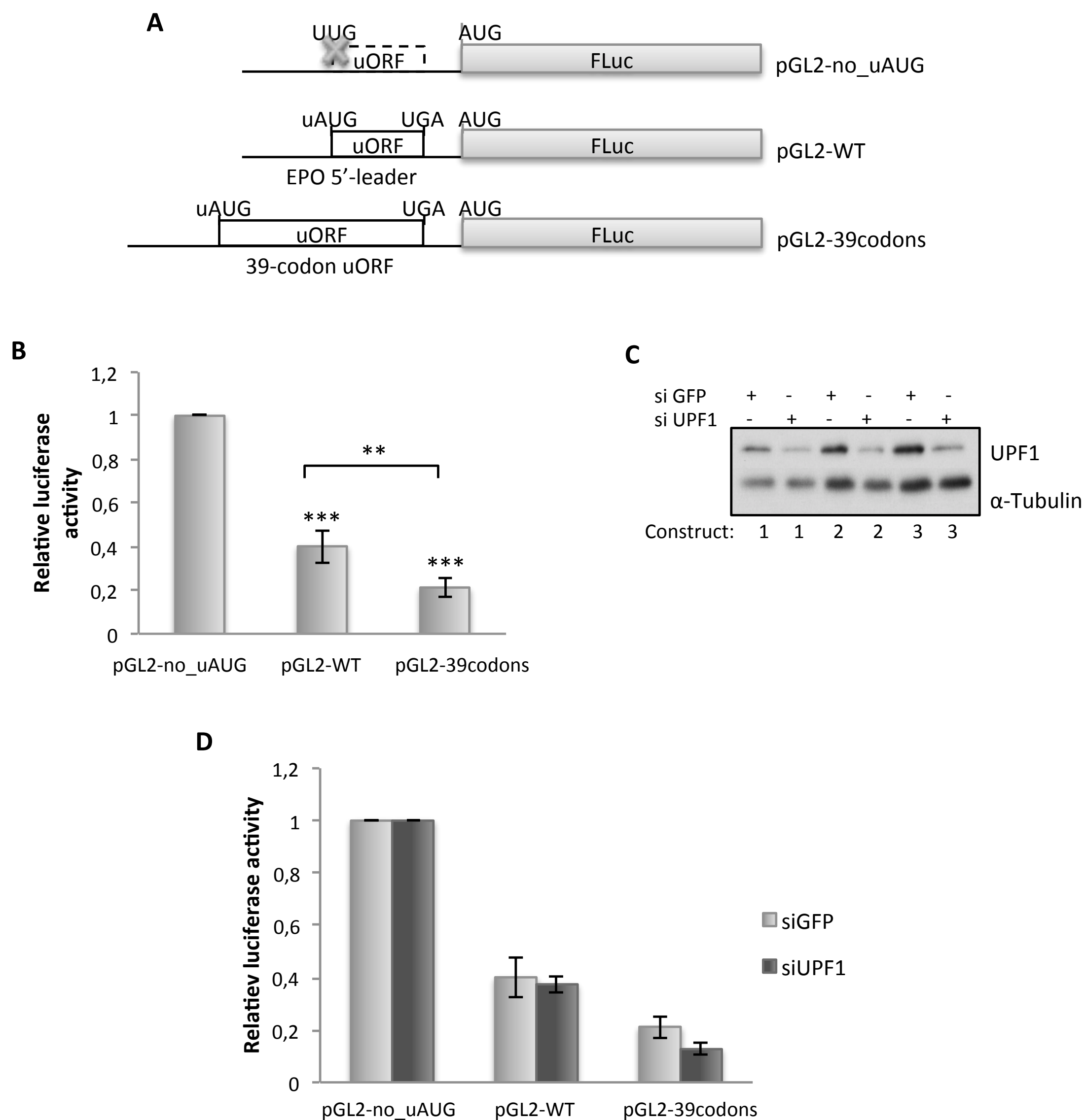


Figure IV.1. The size of the uORF influences the translation reinitiation efficiency.
(A) Schematic representation of reporter constructs. The human *EPO* 5' leader sequence encompassing its uORF (open box) with the intact initiation (uAUG) and termination (UGA) codons, was cloned into the empty vector (pGL2-Luc), upstream of the firefly luciferase coding region (FLuc; grey boxes) to create the pGL2-WT construct. In the pGL2-no_uAUG construct, the uORF initiation codon is mutated (AUG→UUG) (the cross represent the point mutation and the dashed lined box represent the non-functional uORF). The

pGL2-39codons constructs was originated by introducing a nucleotide sequence of the ampicilin resistance gene so that the termination codon of the uORF was located 39 codons downstream of the uAUG.**(B)** The size of the EPO uORF influences the main ORF translation repression, being the longer uORF more repressive. HeLa cells were transiently co-transfected with each one of the constructs described in (A) and with the pRL-TK plasmid encoding the *Renilla* luciferase (RLuc). Cells were lysed twenty-four hours later and the luciferase activity was measured by luminometry assays. FLuc activity values were normalized to RLuc activity to control for transfection efficiency. Relative luciferase activity of the pGL2-no_AUG was defined as one. **(C)** Representative Western blot analysis of HeLa cells extracts transfected with human UPF1 siRNA or a control siRNA target (GFP siRNA). HeLa cells treated with a control (GFP) siRNA or eIF3a, or eIF3c-specific siRNAs were transiently co-transfected with the pGL2-no_uAUG (construct 1), pGL2-WT (construct 2) or pGL2-39codons (construct 3) reporters.Immunoblotting was performed using a human UPF1 specific antibody and an α-tubulin specific antibody to control for variations in protein loading. **(D)** Neither of the uORFs trigger NMD. HeLa cells transfected with GFP or UPF1 siRNAs were also transfected with the reporters described in (A), luciferase activity was measured by luminometry assays and analysis was preformed as described in (B). Average values and standard deviation (SD) of three independent experiments are shown. Statistical analysis was performed using Student's *t* test (unpaired, two tailed); (∗) p<0.05; (∗∗) p<0.01; (∗∗∗) p<0.001.

Western blot analysis demonstrated a decrease in UPF1 protein levels induced by siRNA of about 60%, when compared with results obtained after treatment with GFP siRNAs (Figure IV.1.C). Under these conditions, no significant changes were seen in relative luciferase activity of the reported constructs under UPF1 depletion (Figure IV.1.D). These results are in agreement to the model where NMD is dependent on the deposition of EJC, leading to the resistance to NMD of intronless transcripts (Chapter I).

IV.4.2. eIF3h, f and e affect the efficiency of translation reinitiation

It has been shown that some eIF3 subunits are involved in reinitiation ability during translation of mRNAs harbouring uORFs. Based on these data, we decided to investigate how eIF3 subunits affect reinitiation. For that, we first depleted HeLa cells from eIF3h, f and e subunits, by siRNA transfection using siRNA to GFP as a control. Twenty-four hours later the plasmids pGL2-no_uAUG, pGL2-WT and pGL2-39codons were transiently expressed on those cells for another twenty-four hours. Then, cells were lysed and protein and RNA extracts were obtained.

Efficient knock-down of eIF3h and eIF3f was confirmed by Western blot with specific anti-body against eIF3h and eIF3f, respectively, using α-tubulin as an loading control (Figure IV.2.A and C). In addition, eIF3e depletion was confirmed at the mRNA level, by normalization to mRNA levels of the histone deacetylase 1 (*HDAC1*) (Figure IV.2.E).

HDAC1 mRNA was chosen as an internal control for these analyses since it is constitutively expressed (de Ruijter et al., 2003).

Under these knock-down conditions, we further demonstrated that the relative luciferase activity obtained from the pGL2-WT expression is significantly lower under knock-down of the eIF3h, f and e subunits, in comparison to that at control conditions (Figure IV.2.B, D and F). On the contrary, levels of pGL2-39codons expression are not affected by eIF3h, f and e depletion (Figure IV.2.B, D and F).

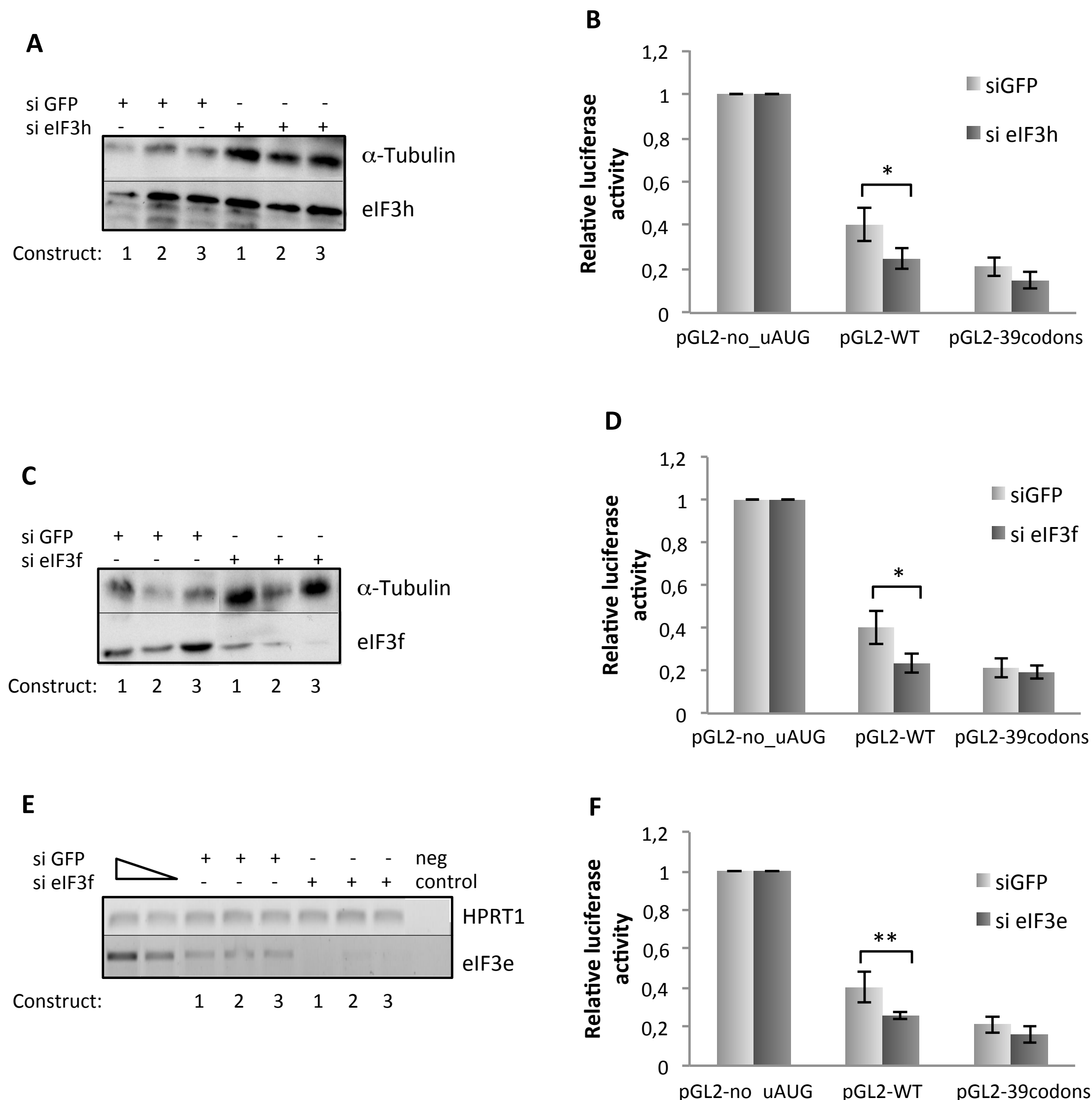


Figure IV.2. Depletion of eIF3h, f, and e, alters the reinitiation efficiency after the translation of the 14--codon uORF.

(A) and **(C)** Representative Western blot analysis of HeLa cells extracts transfected with human eIF3h siRNA (A), eIF3f si RNA (B) or a control siRNA target (GFP siRNA). HeLa cells treated with a control (GFP) siRNA or eIF3a, or eIF3c-specific siRNAs were transiently co-transfected with the pGL2-no_uAUG (construct 1), pGL2-WT (construct 2) or pGL2-39codons (construct 3) reporters. Immunoblotting was

performed using a human eIF3h (A), or human eIF3f (B) specific antibody and an α-tubulin specific antibody to control for variations in protein loading. **(B)**, **(D)** and **(F)** Effect of eIF3h (B), eIF3f (D), or eIF3e (F) depletion on reinitiation downstream from *EPO* uORF and from the 39-codon uORF. HeLa cells treated with a control (GFP) siRNA or eIF3h, eIF3f, or eIF3e-specific siRNAs were transiently co-transfected with the pGL2-no_uAUG, pGL2-WT or pGL2-39codons reporters and with a plasmid encoding *Renilla* luciferase. Twenty- -four hours later cells were lysed, luciferase activity was measured by luminometry assays and analysis was preformed as described in (Figure IV.1.B). **(E)** Representative RT-PCR analyses of RNAs extracted from GFP or eIF3e siRNAs-treated HeLa cells. RT-PCRs were carried out with eIF3e mRNA specific primers to monitor endogenous eIF3e knockdown. The eIF3e mRNA levels were normalized to those of HDAC1 mRNA level. In each panel, the left two lanes correspond to serial dilutions of RNA, demonstrating semi-quantitative conditions used for RT-PCR.

Altogether, these data suggest that all these three subunits are involved in the reinitiation event. The fact that no alteration is observed for the pGL2-39codons expression can be explained by the inefficient reinitiation due to the longer uORF.

IV.4.3. eIF3a and c do not affect the efficiency of translation reinitiation

Considering the putative influence of each eIF3 subunits on translation reinitiation, we decided to broaden our search and therefore test the influence of depleting eIF3a and c in reinitiation efficiency. It is known that both subunits interact with the 40S ribosomal subunit (Fraser et al., 2007; Hinnebusch, 2006). Thus they are good candidates for being involved in translation reinitiation. To test this hypothesis we depleted HeLa cells from each one of these subunits by siRNA transfection using siRNA to GFP as a control. Twenty-four hours later, the plasmids pGL2-no_uAUG, pGL2-WT and pGL2-39codons, together with pRL-TK, were transiently expressed in those cells for another twenty-four hours. Then, cells were lysed and protein and RNA extracts were obtained. Depletion of eIF3a and c was confirmed at the mRNA level, by RT-PCR using the mRNA levels of the *HDAC1* as normalizer (Figure IV.3.A and C). Under these conditions, the relative luciferase activity of each construct measured under depletion of the referred subunits presents no significant changes when compared to the corresponding control GFP siRNA (Figure IV.3.B and D). This indicates that neither of these subunits is involved in the reinitiation efficiency after translation of the *EPO* uORF or the 39-codon uORF.

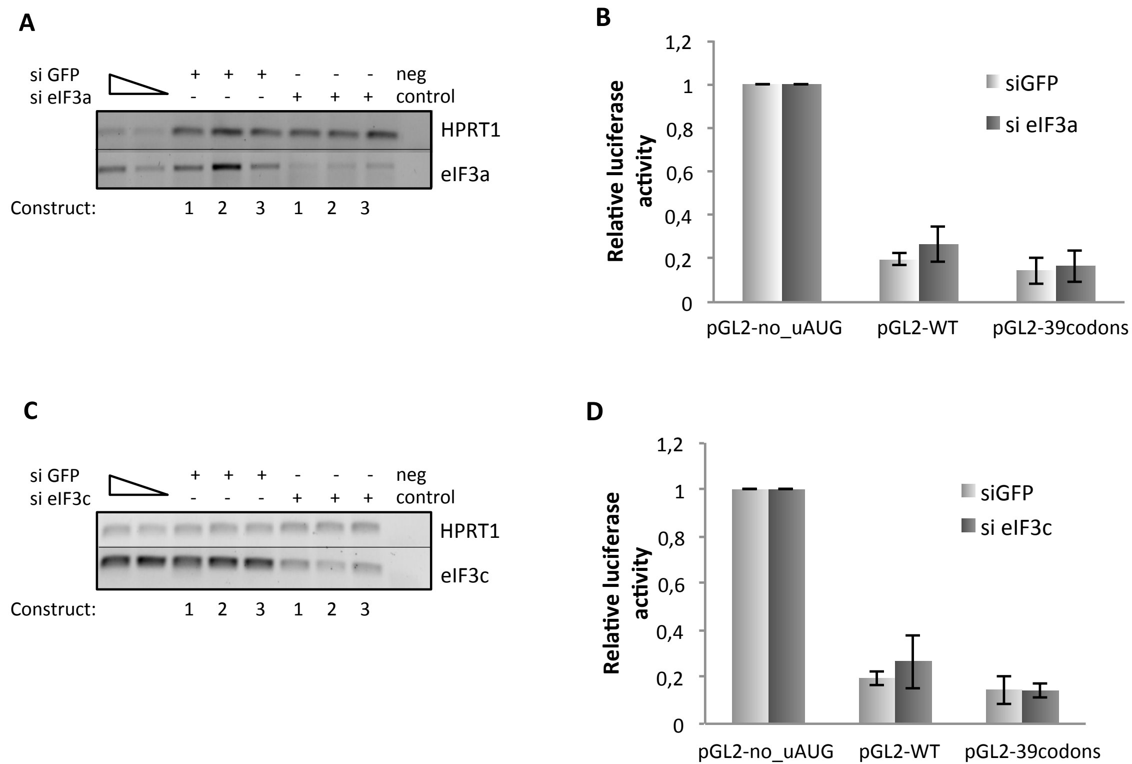


Figure IV.3. Depletion of eIF3a and c does not affect the reinitiation efficiency.
(A) and **(C)** Representative RT-PCR analyses of eIF3a (A) and eIF3c (C) RNA levels extracted from GFP or eIF3a (A), or eIF3c (C) siRNAs-treated HeLa cells. HeLa cells treated with a control (GFP) siRNA or eIF3a, or eIF3c-specific siRNAs were transiently co-transfected with the pGL2-no_uAUG (construct 1), pGL2-WT (construct 2) or pGL2-39codons (construct 3) reporters. RT-PCRs were carried out with eIF3a or eIF3c mRNA specific primers to monitor endogenous knockdown. The eIF3a (A) or eIF3c (C) mRNA levels were normalized to those of *HDAC1* mRNA level. In each panel, the left two lanes correspond to serial dilutions of RNA, demonstrating semi-quantitative conditions used for RT-PCR and the last lane corresponds to the negative control. **(B)** and **(D)** Effect of eIF3a (B), or eIF3c (D) depletion on reinitiation downstream from EPO uORF and from the 39-codon uORF. HeLa cells treated with a control (GFP) siRNA, or with eIF3a-, or eIF3c-specific siRNAs were transiently co-transfected with the pGL2-no_uAUG, pGL2-WT or pGL2-39codons reporter constructs and with a plasmid encoding *Renilla* luciferase, pRL-TK. Twenty-four hours later cells were lysed, luciferase activity was measured by luminometry assays and analysis was preformed as described in (Figure IV.1.B).

IV.5. Discussion

The eIF3 complex is a very important and versatile factor that is able to regulate the translation process. During the translation initiation step, eIF3 promotes the binding of the ternary complex and other eIFs to the 40S ribosomal subunit, recruits the mRNA, promotes scanning and AUG recognition, and functions as a bridge between the 40S

subunit and the mRNA by interacting with eIF4G (Hinnebusch, 2006; Holcik and Pestova, 2007; Valásek, 2012). In the course of elongation, dissociation of eIF3 seems prone to delays, since it is not essential for subunit joining. Consequently, eIF3 can remain associated to the 40S ribosomal subunit throughout the first elongation steps (Szamecz et al., 2008; Valásek et al., 2002). During the termination event, eIF3 is implicated in the dissociation of the translational machinery, recycling, NMD commitment, and translation reinitiation (Hinnebusch, 2006; Pisarev et al., 2007; Roy et al., 2010; Szamecz et al., 2008).

In previous studies, we have shown how eIF3h, f and e influence NMD triggering. We have shown that NMD resistance or induction is modulated by the PTC position along the coding sequence and proved that short ORFs are NMD resistant (Inácio et al., 2004; Peixeiro et al., 2012; Silva and Romão, 2009; Silva et al., 2008). This occurs because in such a short ORF, as it is our model, the PABPC1/eIF4G/eIF3 complex might be still bound to the ribosome when it reaches the stop codon and thus PABPC1 is in a favored position to inhibit NMD (Peixeiro et al., 2012). To be more precise, we have shown that the human eIF3h and eIF3f subunits are involved in the mechanism by which transcripts with an AUG-proximal PTC are NMD-resistant. However, eIF3e has the opposite function, being required for NMD triggering (Peixeiro et al., 2012).

In this study our aim was to study the involvement of these factors in translation reinitiation. For that, we used the *EPO* uORF as experimental model as it allows translation reinitiation to occur (Chapter II).

First, we investigated how the uORF size influences the reinitiation efficiency. For that, we generated a 39-codon uORF from the *EPO* uORF without altering the AUG context so that the leaky scanning mechanism was unaltered (Figure IV.1.A). In these conditions, we observed that the 39-codon uORF allows less translation reinitiation of the main ORF (Figure IV.1.B), meaning that, in uORFs with similar translation rates, reinitiation at a downstream ORF is greater after translation of a shorter uORF as some initiation factors need to stay associated with 40S ribosomal subunit during the translation process. For that matter, eIF3 seems to be a good candidate to determine the reinitiation mechanism.

We observed that depletion of eIF3h, f and e subunits significantly decreases the relative luciferase activity given from the mRNA with the 14-codon uORF but does not affect the

low levels of reinitiation after the 39-codon uORF translation (Figure IV.2.), indicating that these subunits are involved in reinitiation efficiency. This was as expected because the involvement of eIF3h with the reinitiation process after uORF translation was previously described in plants (Roy et al., 2010). Furthermore, since eIF3f and e interact with eIF4G they can allow for the maintenance of the PABPC1/eIF4G/eIF3 complex binding to the 40S ribosomal subunit when it reaches the stop codon, allowing reinitiation to occur (LeFebvre et al., 2006; Masutani et al., 2013; Peixeiro et al., 2012).

As eIF3a and c subunits are conserved among different species, contrary to what happens to eIF3h, f and e, we have also studied their role in reinitiation.

In our model, we show that depletion of eIF3a and c subunits are not involved in translation reinitation (Figure IV.3). This was not expected, since these subunits interact directly with the 40S ribosomal subunit (Fraser et al., 2007; Hinnebusch, 2006), which might indicate their association during the elongation phase and after termination. However, these results might be due an insufficient knock-down of these subunits.

The yeast eIF3a subunit has already been described to impact reinitiation in the GCN4 model, by stabilizing the association of the 40S subunit to the mRNA after dissociation of the 60S subunit (Szamecz et al., 2008; Valásek et al., 2002). Moreover, eIF3c is a critical regulator of AUG recognition (Karásková et al., 2012; Valásek et al., 2002). And hence it was expected that the depletion of eIF3c would decrease the downstream AUG recognition after the 40S subunit resume scanning. However, the yeast eIF3 complex presents a lower degree of complexity than the mammalian eIF3, as it comprises only six subunits. This means that during the course of evolution some subunits could have lost the function of their yeast homologous thus being replaced by another non-homologous subunit.

In conclusion, our work demonstrates that the size of the *EPO* uORF allows for reinitiation and provides an insight of the molecular basis of the reinitiation process by showing that the eIF3h, f and e subunits, but not eIF3a and c subunits, support the reinitiation process.

IV.6. Acknowledgements

We are grateful to Margarida Gama Carvalho for supplying the pRL-TK plasmid. We would also like to thank Isabel Peixeiro, Cláudia Onofre, João Lavinha and Rafaela Lacerda for critical reading of the manuscript. This research was partially supported by Fundação para a Ciência e a Tecnologia (PEst-OE/BIA/UI4046/2011, PTDC/BIM-MED/0352/2012 and SFRH/BD/63581/2009 to C.B.).

CHAPTER V – General Discussion

V.1. General Discussion and Future Perspectives

Human EPO is more than just a hormone responsible for stimulating erythropoiesis. EPO is a multifaceted protein able to promote differentiation, angiogenesis, proliferation and anti-apoptotic activities in several tissues ((Alnaeeli et al., 2012; Marti et al., 1996; Noguchi et al., 2008), turning out to be a protein of the highest interest as a therapeutic agent for several human disorders beyond anemic conditions. However, there is still an insufficient knowledge of how EPO is regulated and how the many regulatory pathways described so far are integrated. It is our belief that a deep understanding of the EPO regulatory mechanisms can provide insights for the development of new therapies, maybe through new therapeutic targets.

In our study, we show how the *EPO* transcript is regulated by a highly conserved uORF (Figure II.1). In fact, we show that *EPO* uORF represses translation in about 3-fold in all the cell lines studied: HEK293, derived from human embryonic kidney; HepG2, from human liver; REPC, from human adult kidney; and SW1088, from fibroblast of the human brain (Chapter II and III). Our results suggest that the *EPO* uORF is responsible for the low levels of EPO expression in several human tissues and that it does not present a tissue specific effect.

In our search for the mechanistic basis behind the *EPO* uORF repression we show that both leaky scanning and reinitiation are able to promote translation of the main ORF. This was expected since the AUG of the *EPO* uORF is in a good but not optimal context and also since the uORF length seems to allow some reinitiation. In addition, the *EPO* uORF functions in a peptide-independent manner and is not able to trigger NMD. Again, our results suggest that these mechanisms are conserved since no alterations were observed between the different tissues in study (Chapter II and III).

With these results we provide a thorough characterization of the *EPO* uORF function. We would like to emphasize that not all studies regarding uORFs present such characterization. In fact, many studies that relate mutations in uORFs with pathologies do not provide specific evidences of the uORF function. In this matter, we believe that such study should be applied more widely in order to increase our knowledge of these elements.

Another surprising aspect was the influence of the *EPO* 3'UTR on the expression of the main ORF. In all cell lines, protein expression from the main ORF increases without altering the mRNA levels, except in REPC cells were there is also an increase in the mRNA levels due to the presence of the 3'UTR. Although that was not the aim of our study, this fact demonstrates how the regulatory mechanisms of EPO expression are still poorly understood. Indeed, the studies on the *EPO* 3'UTR have described the existence of binding sites of several proteins, not entirely described, that might increase the mRNA stability (Choi et al., 2007; Czyzyk-Krzeska and Bendixen, 1999; McGary et al., 1997), which agrees solely with what we have observed in REPC cells. However, the *EPO* 3'UTR did not alleviate the inhibitory effect of the *EPO* uORF as reported for other transcripts (Medenbach et al., 2011; Mehta, 2006). On the contrary, in REPC cells, the uORF repressive effect seems to be potentiated since the mRNA levels increased in the presence of both structures, but no concomitant increase of the protein levels has been observed.

Even with these results showing that the *EPO* uORF is functional in repressing translation of the main ORF, we still wonder about the relevance of this structure. Does it have a canonical function and its presence is just for negatively controlling the expression of EPO or does it respond to cellular stress as reported for other uORF-containing transcripts?

To answer this question, hypoxia arose has a putative stress since it is the main stress able to increase the levels of EPO through HIF1 transcriptional activation (Bunn, 2013; Jelkmann, 2011; Semenza et al., 1990). For that, we stimulate HEK293, HepG2 and REPC cells with chemical hypoxia or with nutrient starvation as a control. Our results show that *EPO* uORF releases its repressive effect during hypoxia only in REPC cells. This suggests a tissue specific derepression during hypoxia. Indeed, it seems that the *EPO* uORF represses translation in all tissues, but responds to different stimuli in a tissue-specific manner. Transcriptional regulation is less prone to respond to sudden changes and hence it takes more time to increase the protein levels, which account for the possibility that regulation at the translational level could provide a more rapid and reversible response to stress conditions. In the present model, the transcriptional activation by HIF1 and the higher efficiency of the *EPO* main ORF translation might act

together, in a coordinated mode, to maximize the increase of EPO protein levels in response to hypoxia.

The hypoxic response of the *EPO* uORF prompted us to understand the mechanisms behind this effect. One possibility was the existence of an IRES structure in the *EPO* transcript 5' leader sequence. However, despite its extremely stable secondary structure (Figure II.8), we have ruled out that possibility both in normoxia and hypoxia. Not many examples have emerged in which a uORF can interact with an IRES on the same transcript to alter the corresponding protein expression (Park et al., 2005; Yaman et al., 2003). Yet we believe that further and more detailed studies will strengthen this perspective and will improve our knowledge on the cooperation of these elements and their relevance to specific regulation of protein expression. Another possible explanation for the alleviated inhibitory effect of the *EPO* uORF was an increase of the leaky scanning mechanism past the uAUG, and/or an increase of translation reinitiation efficiency. In fact, our results show that the uAUG of the *EPO* uORF is less recognized during hypoxia while reinitiation efficiency is not altered. In addition, we show that this is also observed upon stress-induced phosphorylation of eIF2α. This is in agreement with what has been observed for other transcripts bearing only one uORF that is derepressed during stress conditions, nevertheless the underlying mechanism still needs further clarification. So far, it is hypothesized that under phosphorylation of eIF2α the translational machinery will recognize less efficiently AUGs that are not in the optimal context for translation initiation (Palam et al., 2011). In our model, we suppose that this mechanism may rely on other specific agent in order to explain the tissue specificity of the *EPO* uORF derepression. Otherwise, if it was just the translational machinery acting directly on the uAUG recognition, the effect of hypoxia would be observed in all the cell types studied. In this way, we propose that further studies might unravel a tissue-specific protein able to regulate the *EPO* uORF, which eventually might become a putative target for the development of new therapies involving the hematopoietic function of EPO.

In the present dissertation, we also suggest that the repressive effect of the *EPO* uORF is released in neuronal cells in response to ischemia. On the contrary to what has been observed in REPC cells, in SW1088 cells, we show that the mRNA levels are decreased in a uORF dependent manner under stress. Thus, under the same conditions, the protein levels are maintained when compared to unstressed cells, which means that the

translation efficiency of the main ORF has increased. This is a completely different effect from the one observed in REPC cells and, moreover, has not yet been described for other uORF-containing transcripts. Meanwhile, further studies will be crucial to understand whether we are in the presence of transcription inactivation or of mRNA destabilization. Nevertheless, we hypothesized that in the neuronal tissue the cooperation between mechanisms of transcriptional and translational regulation would be necessary to increase the EPO production during ischemia.

Our experimental model also allowed us to study the mechanism of translation reinitiation in what concerns the effect of the uORF length and the involvement of eIF3. Our results show that the length of the uORF is inversely related to the reinitiation efficiency. Also, we show that the eIF3h, f and e subunits are involved in the modulation of translation reinitiation efficiency, and that eIF3a and c subunits do not seem to have an impact on this mechanism (Chapter IV).

The results obtained to eIF3h, f and e subunits were as expected since these subunits have been previously reported to be involved in the interaction amongst the ribosome, eIF4G and the termination machinery, which might be necessary for reinitiation to occur (LeFebvre et al., 2006; Masutani et al., 2013; Peixeiro et al., 2012). Even more, eIF3h was described to directly impact reinitiation in plants (Roy et al., 2010). On the contrary, the data obtained for eIF3a and c subunits were not predictable since eIF3a is implicated in reinitiation in yeast (Szamecz et al., 2008; Valásek et al., 2002), and eIF3c is a critical regulator of AUG recognition, whose depletion could influence recognition of the downstream AUG (Karásková et al., 2012; Valásek et al., 2002). Nonetheless, we cannot forget that the yeast eIF3 complex differs from the human one, resulting in different functions of the homologous subunits, which might explain our results. Also, it is important to note that the knock-down levels might not be enough to observe an effect in reinitiation.

In conclusion, the work from the present dissertation reports a new mechanism involved in the regulation of human EPO gene. Specifically, we dissected the basic mechanisms of the *EPO* uORF function and reinitiation efficiency and show the biological relevance for the *EPO* translational control during stress conditions in renal cells. Furthermore, it also sheds light on the possible regulation of EPO production in the brain during ischemia. These findings might present the start point for the development of therapies for

numerous disorders, using as therapeutic target the modulation of hematopoietic or non-hematopoietic expression of EPO through its uORF. Also, it could present a way to more accurately modulate EPO expression during gene therapy in a tissue specific manner, thus solving many of the problems experienced so far with the usage of the rhEPO administration (for more details see chapter I.3.4.).

In the future, besides the questions that still need to be answered, we believe that it would be interesting to confirm whether the *EPO* uORF is functional in a living model and how it modulates the translation of EPO during stress or tissue injuries. We propose such a study to be performed in model organisms like mice or zebrafish. Furthermore, since EPO is such a multifaceted protein, it would be fascinating to investigate whether the *EPO* uORF is important in cardiac tissue, or others, in which its expression can be detected and regulated. Additionally, knowing that the disruption or alteration of a uORF can result in disease, and that the disturbance of the expression of EPO can result in clinical disorders, we wonder whether there is a SNP or a mutation on the *EPO* uORF nucleotide sequence involved in any disease. We believe that these and other questions about the role of *EPO* uORF will be subject of study in a close future.

CHAPTER VI – References

Abbott, C., and Proud, C. (2004). Translation factors: in sickness and in health. Trends Biochem. Sci. *29*, 25–31.

Alkalaeva, E.Z., Pisarev, A.V., Frolova, L.Y., Kisselev, L.L., and Pestova, T.V. (2006). In vitro reconstitution of eukaryotic translation reveals cooperativity between release factors eRF1 and eRF3. Cell *125*, 1125–1136.

Alnaeeli, M., Wang, L., Piknova, B., Rogers, H., Li, X., and Noguchi, C.T. (2012). Erythropoietin in brain development and beyond. Anat. Res. Int. *2012*, 953264.

Amrani, N., Ganesan, R., Kervestin, S., Mangus, D.A., Ghosh, S., and Jacobson, A. (2004). A faux 3'-UTR promotes aberrant termination and triggers nonsense-mediated mRNA decay. Nature *432*, 112–118.

Amrani, N., Dong, S., He, F., Ganesan, R., Ghosh, S., Kervestin, S., Li, C., Mangus, D.A., Spatrick, P., and Jacobson, A. (2006). Aberrant termination triggers nonsense-mediated mRNA decay. Biochem. Soc. Trans. *34*, 39–42.

Arabpoor, Z., Hamidi, G., Rashidi, B., Shabrang, M., Alaei, H., Sharifi, M.R., Salami, M., Dolatabadi, H.R.D., and Reisi, P. (2012). Erythropoietin improves neuronal proliferation in dentate gyrus of hippocampal formation in an animal model of Alzheimer's disease. Adv. Biomed. Res. *1*, 50.

Arcasoy, M.O. (2008). The non-haematopoietic biological effects of erythropoietin. Br. J. Haematol. *141*, 14–31.

Bach, J., Endler, G., Winkelmann, B.R., Boehm, B.O., Maerz, W., Mannhalter, C., and Hellstern, P. (2008). Coagulation factor XII (FXII) activity, activated FXII, distribution of FXII C46T gene polymorphism and coronary risk. J. Thromb. Haemost. JTH *6*, 291–296.

Baek, I.C., Kim, J.K., Cho, K.-H., Cha, D.-S., Cho, J.-W., Park, J.K., Song, C.-W., and Yoon, S.K. (2009). A novel mutation in Hr causes abnormal hair follicle morphogenesis in hairpoor mouse, an animal model for Marie Unna Hereditary Hypotrichosis. Mamm. Genome *20*, 350–358.

Barbosa, C., and Romão, L. (2013). Translation of the human erythropoietin transcript is regulated by an upstream open reading frame in response to hypoxia. RNA Rev.

Barbosa, C., Peixeiro, I., and Romão, L. (2013). Gene Expression Regulation by Upstream Open Reading Frames and Human Disease. PLoS Genet. *9*, e1003529.

Bastide, A., Karaa, Z., Bornes, S., Hieblot, C., Lacazette, E., Prats, H., and Touriol, C. (2008). An upstream open reading frame within an IRES controls expression of a specific VEGF-A isoform. Nucleic Acids Res. *36*, 2434–2445.

Beffagna, G., Occhi, G., Nava, A., Vitiello, L., Ditadi, A., Basso, C., Bauce, B., Carraro, G., Thiene, G., Towbin, J.A., et al. (2005). Regulatory mutations in transforming growth factor-beta3 gene cause arrhythmogenic right ventricular cardiomyopathy type 1. Cardiovasc. Res. *65*, 366–373.

Behm-Ansmant, I., and Izaurralde, E. (2006). Quality control of gene expression: a stepwise assembly pathway for the surveillance complex that triggers nonsense-mediated mRNA decay. Genes Dev. *20*, 391–398.

Bennett, C.L., Silver, S.M., Djulbegovic, B., Samaras, A.T., Blau, C.A., Gleason, K.J., Barnato, S.E., Elverman, K.M., Courtney, D.M., McKoy, J.M., et al. (2008). Venous thromboembolism and mortality associated with recombinant erythropoietin and darbepoetin administration for the treatment of cancer-associated anemia. JAMA J. Am. Med. Assoc. *299*, 914–924.

Bersano, A., Ballabio, E., Bresolin, N., and Candelise, L. (2008). Genetic polymorphisms for the study of multifactorial stroke. Hum. Mutat. *29*, 776–795.

Besarab, A., Frinak, S., and Yee, J. (2009). What is so bad about a hemoglobin level of 12 to 13 g/dL for chronic kidney disease patients anyway? Adv. Chronic Kidney Dis. *16*, 131–142.

Van den Beucken, T., Magagnin, M.G., Savelkouls, K., Lambin, P., Koritzinsky, M., and Wouters, B.G. (2007). Regulation of Cited2 expression provides a functional link between translational and transcriptional responses during hypoxia. Radiother. Oncol. *83*, 346–352.

Bisio, A., Nasti, S., Jordan, J.J., Gargiulo, S., Pastorino, L., Provenzani, A., Quattrone, A., Queirolo, P., Bianchi-Scarrà, G., Ghiorzo, P., et al. (2010). Functional analysis of CDKN2A/p16INK4a 5′-UTR variants predisposing to melanoma. Hum. Mol. Genet. *19*, 1479–1491.

Bittorf, T., Seiler, J., Lüdtke, B., Büchse, T., Jaster, R., and Brock, J. (2000). Activation of STAT5 during EPO-directed suppression of apoptosis. Cell. Signal. *12*, 23–30.

Blais, J.D., Filipenko, V., Bi, M., Harding, H.P., Ron, D., Koumenis, C., Wouters, B.G., and Bell, J.C. (2004). Activating transcription factor 4 is translationally regulated by hypoxic stress. Mol. Cell. Biol. *24*, 7469–7482.

Blanchard, K.L., Acquaviva, A.M., Galson, D.L., and Bunn, H.F. (1992). Hypoxic induction of the human erythropoietin gene: cooperation between the promoter and enhancer, each of which contains steroid receptor response elements. Mol. Cell. Biol. *12*, 5373–5385.

Braverman, N., Chen, L., Lin, P., Obie, C., Steel, G., Douglas, P., Chakraborty, P.K., Clarke, J.T.R., Boneh, A., Moser, A., et al. (2002). Mutation analysis of PEX7 in 60 probands with rhizomelic chondrodysplasia punctata and functional correlations of genotype with phenotype. Hum. Mutat. *20*, 284–297.

Brines, M., Grasso, G., Fiordaliso, F., Sfacteria, A., Ghezzi, P., Fratelli, M., Latini, R., Xie, Q.-W., Smart, J., Su-Rick, C.-J., et al. (2004). Erythropoietin mediates tissue protection through an erythropoietin and common beta-subunit heteroreceptor. Proc. Natl. Acad. Sci. U. S. A. *101*, 14907–14912.

Brines, M., Patel, N.S.A., Villa, P., Brines, C., Mennini, T., De Paola, M., Erbayraktar, Z., Erbayraktar, S., Sepodes, B., Thiemermann, C., et al. (2008). Nonerythropoietic, tissue-protective peptides derived from the tertiary structure of erythropoietin. Proc. Natl. Acad. Sci. U. S. A. *105*, 10925–10930.

Brown, C.Y., Mize, G.J., Pineda, M., George, D.L., and Morris, D.R. (1999). Role of two upstream open reading frames in the translational control of oncogene mdm2. Oncogene *18*, 5631–5637.

Brunn, G.J., Hudson, C.C., Sekulić, A., Williams, J.M., Hosoi, H., Houghton, P.J., Lawrence, J.C., Jr, and Abraham, R.T. (1997). Phosphorylation of the translational repressor PHAS-I by the mammalian target of rapamycin. Science *277*, 99–101.

Bunn, H.F. (1990). Erythropoietin: current status. Yale J. Biol. Med. *63*, 381–386.

Bunn, H.F. (2013). Erythropoietin. Cold Spring Harb. Perspect. Med. *3*, a011619.

Bushuev, V.I., Miasnikova, G.Y., Sergueeva, A.I., Polyakova, L.A., Okhotin, D., Gaskin, P.R., Debebe, Z., Nekhai, S., Castro, O.L., and Prchal, J.T. (2006). Endothelin-1, vascular endothelial growth factor and systolic pulmonary artery pressure in patients with Chuvash polycythemia. Haematologica *91*, 744–749.

Calvo, S.E., Pagliarini, D.J., and Mootha, V.K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proc. Natl. Acad. Sci. U. S. A. *106*, 7507–7512.

Candeias, M.M., Powell, D.J., Roubalova, E., Apcher, S., Bourougaa, K., Vojtesek, B., Bruzzoni-Giovanelli, H., and Fåhraeus, R. (2006). Expression of p53 and p53/47 are controlled by alternative mechanisms of messenger RNA translation initiation. Oncogene *25*, 6936–6947.

Caprara, C., and Grimm, C. (2012). From oxygen to erythropoietin: Relevance of hypoxia for retinal development, health and disease. Prog. Retin. Eye Res. *31*, 89–119.

Casals-Pascual, C., Idro, R., Picot, S., Roberts, D.J., and Newton, C.R.J.C. (2009). Can erythropoietin be used to prevent brain damage in cerebral malaria? Trends Parasitol. *25*, 30–36.

Cazzola, M., and Skoda, R.C. (2000). Translational pathophysiology: a novel molecular mechanism of human disease. Blood *95*, 3280–3288.

Chateauvieux, S., Grigorakaki, C., Morceau, F., Dicato, M., and Diederich, M. (2011). Erythropoietin, erythropoiesis and beyond. Biochem. Pharmacol. *82*, 1291–1303.

Chatterjee, S., and Pal, J. (2009). Role of 5'-and 3'-untranslated regions of mRNAs in human diseases. Biol. Cell *101*, 251–262.

Chen, C., and Sytkowski, A.J. (2004). Erythropoietin regulation of Raf-1 and MEK: evidence for a Ras-independent mechanism. Blood *104*, 73–80.

Chen, Y.-J., Tan, B.C.-M., Cheng, Y.-Y., Chen, J.-S., and Lee, S.-C. (2010). Differential regulation of CHOP translation by phosphorylated eIF4E under stress conditions. Nucleic Acids Res. *38*, 764–777.

Child, S.J., Miller, M.K., and Geballe, A.P. (1999). Translational control by an upstream open reading frame in the HER-2/neu transcript. J. Biol. Chem. *274*, 24335–24341.

Chin, K., Yu, X., Beleslin-Cokic, B., Liu, C., Shen, K., Mohrenweiser, H.W., and Noguchi, C.T. (2000). Production and processing of erythropoietin receptor transcripts in brain. Mol. Brain Res. *81*, 29–42.

Chiu, W.-L., Wagner, S., Herrmannová, A., Burela, L., Zhang, F., Saini, A.K., Valásek, L., and Hinnebusch, A.G. (2010). The C-terminal region of eukaryotic translation initiation factor 3a (eIF3a) promotes mRNA recruitment, scanning, and, together with eIF3j and the eIF3b RNA recognition motif, selection of AUG start codons. Mol. Cell. Biol. *30*, 4415–4434.

Choe, J., Oh, N., Park, S., Lee, Y.K., Song, O.-K., Locker, N., Chi, S.-G., and Kim, Y.K. (2012). Translation initiation on mRNAs bound by nuclear cap-binding protein complex CBP80/20 requires interaction between CBP80/20-dependent translation initiation factor and eukaryotic translation initiation factor 3g. J. Biol. Chem. *287*, 18500–18509.

Choi, B.H., Ha, Y., Ahn, C.H., Huang, X., Kim, J.M., Park, S.R., Park, H., Park, H.C., Kim, S.W., and Lee, M. (2007). A hypoxia-inducible gene expression system using erythropoietin 3'untranslated region for the gene therapy of rat spinal cord injury. Neurosci. Lett. *412*, 118–122.

Chong, Z.Z., Li, F., and Maiese, K. (2005). Erythropoietin requires NF-kappaB and its nuclear translocation to prevent early and late apoptotic neuronal injury during beta-amyloid toxicity. Curr. Neurovasc. Res. *2*, 387–399.

Cobbold, L.C., Spriggs, K.A., Haines, S.J., Dobbyn, H.C., Hayes, C., de Moor, C.H., Lilley, K.S., Bushell, M., and Willis, A.E. (2008). Identification of internal ribosome entry segment (IRES)-trans-acting factors for the Myc family of IRESs. Mol. Cell. Biol. *28*, 40–49.

Cohen, H.Y., Miller, C., Bitterman, K.J., Wall, N.R., Hekking, B., Kessler, B., Howitz, K.T., Gorospe, M., de Cabo, R., and Sinclair, D.A. (2004). Calorie restriction promotes mammalian cell survival by inducing the SIRT1 deacetylase. Science *305*, 390–392.

Cuchalová, L., Kouba, T., Herrmannová, A., Dányi, I., Chiu, W.-L., and Valásek, L. (2010). The RNA recognition motif of eukaryotic translation initiation factor 3g (eIF3g) is required for resumption of scanning of posttermination ribosomes for reinitiation on GCN4 and together with eIF3i stimulates linear scanning. Mol. Cell. Biol. *30*, 4671–4686.

Czyzyk-Krzeska, M.F., and Bendixen, A.C. (1999). Identification of the poly (C) binding protein in the complex associated with the 3′ untranslated region of erythropoietin messenger RNA. Blood *93*, 2111–2120.

Dame, C., Fahnenstich, H., Freitag, P., Hofmann, D., Abdul-Nour, T., Bartmann, P., and Fandrey, J. (1998). Erythropoietin mRNA expression in human fetal and neonatal tissue. Blood *92*, 3218–3225.

Dame, C., Juul, S.E., and Christensen, R.D. (2001). The biology of erythropoietin in the central nervous system and its neurotrophic and neuroprotective potential. Biol. Neonate *79*, 228–235.

Dame, C., Sola, M.C., Lim, K.-C., Leach, K.M., Fandrey, J., Ma, Y., Knöpfle, G., Engel, J.D., and Bungert, J. (2004). Hepatic erythropoietin gene regulation by GATA-4. J. Biol. Chem. *279*, 2955–2961.

Dang Do, A.N., Kimball, S.R., Cavener, D.R., and Jefferson, L.S. (2009). eIF2alpha kinases GCN2 and PERK modulate transcription and translation of distinct sets of mRNAs in mouse liver. Physiol. Genomics *38*, 328–341.

Davies, W.L., Vandenberg, J.I., Sayeed, R.A., and Trezise, A.E.O. (2004). Post-transcriptional regulation of the cystic fibrosis gene in cardiac development and hypertrophy. Biochem. Biophys. Res. Commun. *319*, 410–418.

Digicaylioglu, M., and Lipton, S.A. (2001). Erythropoietin-mediated neuroprotection involves cross-talk between Jak2 and NF-κB signalling cascades. Nature *412*, 641–647.

Ebert, B.L., and Bunn, H.F. (1998). Regulation of transcription by hypoxia requires a multiprotein complex that includes hypoxia-inducible factor 1, an adjacent transcription factor, and p300/CREB binding protein. Mol. Cell. Biol. *18*, 4089–4096.

Ebert, B.L., and Bunn, H.F. (1999). Regulation of the erythropoietin gene. Blood *94*, 1864–1877.

Egrie, J.C., Browne, J., Lai, P., and Lin, F.K. (1985). Characterization of recombinant monkey and human erythropoietin. Prog. Clin. Biol. Res. *191*, 339–350.

Fandrey, J. (2004). Oxygen-dependent and tissue-specific regulation of erythropoietin gene expression. AJP Regul. Integr. Comp. Physiol. *286*, R977–R988.

Fandrey, J., and Bunn, H. (1993). In vivo and in vitro regulation of erythropoietin mRNA: measurement by competitive polymerase chain reaction. Blood *81*, 617 –623.

Fasken, M.B., and Corbett, A.H. (2005). Process or perish: quality control in mRNA biogenesis. Nat. Struct. Mol. Biol. *12*, 482–488.

Fernandez, J., Yaman, I., Sarnow, P., Snider, M.D., and Hatzoglou, M. (2002). Regulation of internal ribosomal entry site-mediated translation by phosphorylation of the translation initiation factor eIF2alpha. J. Biol. Chem. *277*, 19198–19205.

Fraser, C.S., Berry, K.E., Hershey, J.W.B., and Doudna, J.A. (2007). eIF3j Is Located in the Decoding Center of the Human 40S Ribosomal Subunit. Mol. Cell *26*, 811–819.

Frede, S., Freitag, P., Geuting, L., Konietzny, R., and Fandrey, J. (2011). Oxygen-regulated expression of the erythropoietin gene in the human renal cell line REPC. Blood *117*, 4905.

Frietsch, T., Maurer, M.H., Vogel, J., Gassmann, M., Kuschinsky, W., and Waschke, K.F. (2007). Reduced cerebral blood flow but elevated cerebral glucose metabolic rate in erythropoietin overexpressing transgenic mice with excessive erythrocytosis. J. Cereb. Blood Flow Metab. Off. J. Int. Soc. Cereb. Blood Flow Metab. *27*, 469–476.

Frolova, L.Y., Merkulova, T.I., and Kisselev, L.L. (2000). Translation termination in eukaryotes: polypeptide release factor eRF1 is composed of functionally and structurally distinct domains. RNA New York N *6*, 381–390.

Gardner, L.B. (2008). Hypoxic inhibition of nonsense-mediated RNA decay regulates gene expression and the integrated stress response. Mol. Cell. Biol. *28*, 3729–3741.

Gassmann, M., and Soliz, J. (2009). Erythropoietin modulates the neural control of hypoxic ventilation. Cell. Mol. Life Sci. *66*, 3575–3582.

Geballe, A.P., and Morris, D.R. (1994). Initiation codons within 5'-leaders of mRNAs as regulators of translation. Trends Biochem. Sci. *19*, 159–164.

Gebauer, F., and Hentze, M.W. (2004). Molecular mechanisms of translational control. Nat. Rev. Mol. Cell Biol. *5*, 827–835.

Ghezzi, P., and Brines, M. (2004). Erythropoietin as an antiapoptotic, tissue-protective cytokine. Cell Death Differ. *11*, S37–S44.

Ghilardi, N., and Skoda, R.C. (1999). A single-base deletion in the thrombopoietin (TPO) gene causes familial essential thrombocythemia through a mechanism of more efficient translation of TPO mRNA. Blood *94*, 1480–1482.

Ghilardi, N., Wiestner, A., Kikuchi, M., Ohsaka, A., and Skoda, R.C. (1999). Hereditary thrombocythaemia in a Japanese family is caused by a novel point mutation in the thrombopoietin gene. Br. J. Haematol. *107*, 310–316.

Goldberg, M., Gaut, C., and Bunn, H. (1991). Erythropoietin mRNA levels are governed by both the rate of gene transcription and posttranscriptional events. Blood *77*, 271 –277.

Göpfert, U., Kullmann, M., and Hengst, L. (2003). Cell cycle-dependent translation of p27 involves a responsive element in its 5'-UTR that overlaps with a uORF. Hum. Mol. Genet. *12*, 1767–1779.

Grobe, K., and Esko, J.D. (2002). Regulated translation of heparan sulfate N-acetylglucosamine N-deacetylase/n-sulfotransferase isozymes by structured 5'-untranslated regions and internal ribosome entry sites. J. Biol. Chem. *277*, 30699–30706.

Gu, Z., Harrod, R., Rogers, E.J., and Lovett, P.S. (1994). Anti-peptidyl transferase leader peptides of attenuation-regulated chloramphenicol-resistance genes. Proc. Natl. Acad. Sci. U. S. A. *91*, 5612–5616.

Haase, V.H. (2013). Regulation of erythropoiesis by hypoxia-inducible factors. Blood Rev. *27*, 41–53.

Hambley, H., and Mufti, G.J. (1990). Erythropoietin: an old friend revisited. BMJ *300*, 621–622.

Hansen, M.B., Mitchelmore, C., Kjaerulff, K.M., Rasmussen, T.E., Pedersen, K.M., and Jensen, N.A. (2002). Mouse Atf5: molecular cloning of two novel mRNAs, genomic organization, and odorant sensory neuron localization. Genomics *80*, 344–350.

Harding, H.P., Zeng, H., Zhang, Y., Jungries, R., Chung, P., Plesken, H., Sabatini, D.D., and Ron, D. (2001). Diabetes mellitus and exocrine pancreatic dysfunction in perk-/- mice reveals a role for translational control in secretory cell survival. Mol. Cell *7*, 1153–1163.

Hay, N., and Sonenberg, N. (2004). Upstream and downstream of mTOR. Genes Dev. *18*, 1926–1945.

Hayden, C.A., and Jorgensen, R.A. (2007). Identification of novel conserved peptide uORF homology groups in Arabidopsis and rice reveals ancient eukaryotic origin of select groups and preferential association with transcription factor-encoding genes. BMC Biol. *5*, 32.

Hernández, G., Altmann, M., and Lasko, P. (2010). Origins and evolution of the mechanisms regulating translation initiation in eukaryotes. Trends Biochem. Sci. *35*, 63–73.

Herrmannová, A., Daujotyte, D., Yang, J.-C., Cuchalová, L., Gorrec, F., Wagner, S., Dányi, I., Lukavsky, P.J., and Valásek, L.S. (2012). Structural analysis of an eIF3 subcomplex reveals conserved interactions required for a stable and proper translation pre-initiation complex assembly. Nucleic Acids Res. *40*, 2294–2311.

Hinnebusch, A.G. (1997). Translational regulation of yeast GCN4. J. Biol. Chem. *272*, 21661–21664.

Hinnebusch, A.G. (2005). Translational regulation of GCN4 and the general amino acid control of yeast. Annu. Rev. Microbiol. *59*, 407–450.

Hinnebusch, A.G. (2006). eIF3: a versatile scaffold for translation initiation complexes. Trends Biochem. Sci. *31*, 553–562.

Hinnebusch, A., Dever, T., and Asano, K. (2007). Mechanisms of translation initiation in the yeast Saccharomyces cerevisiae (In: Mathews M, Sonenberg N, Hershey JWB, editors. Translational Control in Biology and Medicine. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. pp 225-268).

Hino, M., Miyazono, K., Urabe, A., and Takaku, F. (1989). Effects of recombinant human erythropoietin on hematopoietic progenitors of chronic hemodialysis patients in vitro and in vivo. Int. J. Cell Cloning *7*, 257–263.

Le Hir, H., Izaurralde, E., Maquat, L.E., and Moore, M.J. (2000). The spliceosome deposits multiple proteins 20-24 nucleotides upstream of mRNA exon-exon junctions. EMBO J. *19*, 6860–6869.

Hoch, M., Fischer, P., Stapel, B., Missol-Kolka, E., Sekkali, B., Scherr, M., Favret, F., Braun, T., Eder, M., Schuster-Gossler, K., et al. (2011). Erythropoietin preserves the endothelial differentiation capacity of cardiac progenitor cells and reduces heart failure during anticancer therapies. Cell Stem Cell *9*, 131–143.

Holcik, M., and Pestova, T.V. (2007). Translation mechanism and regulation: old players, new concepts. Meeting on Translational Control and Non-Coding RNA. EMBO Rep. *8*, 639–643.

Hood, H.M., Neafsey, D.E., Galagan, J., and Sachs, M.S. (2009). Evolutionary roles of upstream open reading frames in mediating gene regulation in fungi. Annu. Rev. Microbiol. *63*, 385–409.

Hoshino, S., Imai, M., Kobayashi, T., Uchida, N., and Katada, T. (1999). The eukaryotic polypeptide chain releasing factor (eRF3/GSPT) carrying the translation termination signal to the 3’-Poly(A) tail of mRNA. Direct association of erf3/GSPT with polyadenylate-binding protein. J. Biol. Chem. *274*, 16677–16680.

Lee-Huang, S., Lin, J.J., Kung, H., Huang, P.L., Lee, L., and Huang, P.L. (1993). The human erythropoietin-encoding gene contains a CAAT box, TATA boxes and other transcriptional regulatory elements in its 5’flanking region. Gene *128*, 227–236.

Huopio, H., Jääskeläinen, J., Komulainen, J., Miettinen, R., Kärkkäinen, P., Laakso, M., Tapanainen, P., Voutilainen, R., and Otonkoski, T. (2002). Acute insulin response tests for the differential diagnosis of congenital hyperinsulinism. J. Clin. Endocrinol. Metab. *87*, 4502–4507.

Iacono, M., Mignone, F., and Pesole, G. (2005). uAUG and uORFs in human and rodent 5’untranslated mRNAs. Gene *349*, 97–105.

Imagawa, S., Goldberg, M., Doweiko, J., and Bunn, H. (1991). Regulatory elements of the erythropoietin gene. Blood *77*, 278 –285.

Imagawa, S., Izumi, T., and Miura, Y. (1994). Positive and negative regulation of the erythropoietin gene. J. Biol. Chem. *269*, 9038–9044.

Imagawa, S., Yamamoto, M., and Miura, Y. (1997). Negative regulation of the erythropoietin gene expression by the GATA transcription factors. Blood *89*, 1430–1439.

Inácio, A., Silva, A.L., Pinto, J., Ji, X., Morgado, A., Almeida, F., Faustino, P., Lavinha, J., Liebhaber, S.A., and Romão, L. (2004). Nonsense mutations in close proximity to the initiation codon fail to trigger full nonsense-mediated mRNA decay. J. Biol. Chem. *279*, 32170–32180.

Janzen, D.M., Frolova, L., and Geballe, A.P. (2002). Inhibition of translation termination mediated by an interaction of eukaryotic release factor 1 with a nascent peptidyl-tRNA. Mol. Cell. Biol. *22*, 8562–8570.

Jelkmann, W. (1992). Erythropoietin: structure, control of production, and function. Physiol. Rev. *72*, 449–489.

Jelkmann, W. (2011). Regulation of erythropoietin production. J. Physiol. *589*, 1251–1258.

Jousse, C., Bruhat, A., Carraro, V., Urano, F., Ferrara, M., Ron, D., and Fafournoux, P. (2001). Inhibition of CHOP translation by a peptide encoded by an open reading frame localized in the chop 5’UTR. Nucleic Acids Res. *29*, 4341–4351.

Jubinsky, P.T., Krijanovski, O.I., Nathan, D.G., Tavernier, J., and Sieff, C.A. (1997). The beta chain of the interleukin-3 receptor functionally associates with the erythropoietin receptor. Blood *90*, 1867–1873.

Kanaji, T., Okamura, T., Osaki, K., Kuroiwa, M., Shimoda, K., Hamasaki, N., and Niho, Y. (1998). A common genetic polymorphism (46 C to T substitution) in the 5’-untranslated region of the coagulation factor XII gene is associated with low translation efficiency and decrease in plasma factor XII level. Blood *91*, 2010–2014.

Kapp, L.D., and Lorsch, J.R. (2004). The molecular mechanics of eukaryotic translation. Annu. Rev. Biochem. *73*, 657–704.

Karagyozov, L., Godfrey, R., Böhmer, S.A., Petermann, A., Hölters, S., Östman, A., and Böhmer, F.D. (2008). The structure of the 5’-end of the protein-tyrosine phosphatase PTPRJ mRNA reveals a novel mechanism for translation attenuation. Nucleic Acids Res. *36*, 4443–4453.

Karásková, M., Gunišová, S., Herrmannová, A., Wagner, S., Munzarová, V., and Valášek, L.S. (2012). Functional characterization of the role of the N-terminal domain of the c/Nip1 subunit of eukaryotic initiation factor 3 (eIF3) in AUG recognition. J. Biol. Chem. *287*, 28420–28434.

Kikuchi, M., Tayama, T., Hayakawa, H., Takahashi, I., Hoshino, H., and Ohsaka, A. (1995). Familial thrombocytosis. Br. J. Haematol. *89*, 900–902.

King, V.R., Averill, S.A., Hewazy, D., Priestley, J.V., Torup, L., and Michael-Titus, A.T. (2007). Erythropoietin and carbamylated erythropoietin are neuroprotective following spinal cord hemisection in the rat. Eur. J. Neurosci. *26*, 90–100.

Kisselev, L., Ehrenberg, M., and Frolova, L. (2003). Termination of translation: interplay of mRNA, rRNAs and release factors? EMBO J. *22*, 175–182.

Knabe, W., Sirén, A.-L., Ehrenreich, H., and Kuhn, H.-J. (2005). Expression patterns of erythropoietin and its receptor in the developing spinal cord and dorsal root ganglia. Anat. Embryol. (Berl.) *210*, 209–219.

Kochetov, A.V., Ahmad, S., Ivanisenko, V., Volkova, O.A., Kolchanov, N.A., and Sarai, A. (2008). uORFs, reinitiation and alternative translation start sites in human mRNAs. FEBS Lett. *582*, 1293–1297.

Komar, A.A., and Hatzoglou, M. (2011). Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states. Cell Cycle *10*, 229–240.

Kondo, T., Okabe, M., Sanada, M., Kurosawa, M., Suzuki, S., Kobayashi, M., Hosokawa, M., and Asaka, M. (1998). Familial essential thrombocythemia associated with one-base deletion in the 5'-untranslated region of the thrombopoietin gene. Blood *92*, 1091–1096.

Kontos, H., Napthine, S., and Brierley, I. (2001). Ribosomal pausing at a frameshifter RNA pseudoknot is sensitive to reading phase but shows little correlation with frameshift efficiency. Mol. Cell. Biol. *21*, 8657–8670.

Koritzinsky, M., Rouschop, K.M.A., van den Beucken, T., Magagnin, M.G., Savelkouls, K., Lambin, P., and Wouters, B.G. (2007). Phosphorylation of eIF2α is required for mRNA translation inhibition and survival during moderate hypoxia. Radiother. Oncol. *83*, 353–361.

Koschmieder, S., D'Alo, F., Radomska, H., Schoneich, C., Chang, J.S., Konopleva, M., Kobayashi, S., Levantini, E., Suh, N., Di Ruscio, A., et al. (2007). CDDO induces granulocytic differentiation of myeloid leukemic blasts through translational up-regulation of p42 CCAAT enhancer binding protein alpha. Blood *110*, 3695–3705.

Koumenis, C., Naczki, C., Koritzinsky, M., Rastani, S., Diehl, A., Sonenberg, N., Koromilas, A., and Wouters, B.G. (2002). Regulation of protein synthesis by hypoxia via activation of the endoplasmic reticulum kinase PERK and phosphorylation of the translation initiation factor eIF2alpha. Mol. Cell. Biol. *22*, 7405–7416.

Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell *44*, 283–292.

Kozak, M. (1987). An analysis of 5'-noncoding sequences from 699 vertebrate messenger RNAs. Nucleic Acids Res. *15*, 8125–8148.

Kozak, M. (1991). An analysis of vertebrate mRNA sequences: intimations of translational control. J. Cell Biol. *115*, 887–903.

Kozak, M. (1997). Recognition of AUG and alternative initiator codons is augmented by G in position+ 4 but is not generally affected by the nucleotides in positions+ 5 and+ 6. EMBO J. *16*, 2482–2492.

Kozak, M. (1999). Initiation of translation in prokaryotes and eukaryotes. Gene *234*, 187–208.

Kozak, M. (2001). Constraints on reinitiation of translation in mammals. Nucleic Acids Res. *29*, 5226–5232.

Kozak, M. (2002). Pushing the limits of the scanning mechanism for initiation of translation. Gene *299*, 1–34.

Kozak, M. (2005). Regulation of translation via mRNA structure in prokaryotes and eukaryotes. Gene *361*, 13–37.

Krantz, S.B. (1991). Erythropoietin. Blood *77*, 419–434.

Krude, H., Biebermann, H., Luck, W., Horn, R., Brabant, G., and Gruters, A. (1998). Severe early-onset obesity, adrenal insufficiency and red hair pigmentation caused by POMC mutations in humans. Nat. Genet. *19*, 155–157.

Kumral, A., Tüzün, F., Oner, M.G., Genç, S., Duman, N., and Özkan, H. (2011). Erythropoietin in neonatal brain protection: The past, the present and the future. Brain Dev. *33*, 632–643.

Lammich, S., Schöbel, S., Zimmer, A.-K., Lichtenthaler, S.F., and Haass, C. (2004). Expression of the Alzheimer protease BACE1 is suppressed via its 5'-untranslated region. EMBO Rep. *5*, 620–625.

Lee, M., Choi, D., Choi, M.J., Jeong, J.H., Kim, W.J., Oh, S., Kim, Y.H., Bull, D.A., and Kim, S.W. (2006). Hypoxia-inducible gene expression system using the erythropoietin enhancer and 3'-untranslated region

for the VEGF gene therapy. J. Controlled Release *115*, 113–119.

Lee, Y.-Y., Cevallos, R.C., and Jan, E. (2009). An upstream open reading frame regulates translation of GADD34 during cellular stresses that induce eIF2alpha phosphorylation. J. Biol. Chem. *284*, 6661–6673.

LeFebvre, A.K., Korneeva, N.L., Trutschl, M., Cvek, U., Duzan, R.D., Bradley, C.A., Hershey, J.W.B., and Rhoads, R.E. (2006). Translation initiation factor eIF4G-1 binds to eIF3 through the eIF3e subunit. J. Biol. Chem. *281*, 22917–22932.

Lewerenz, J., Sato, H., Albrecht, P., Henke, N., Noack, R., Methner, A., and Maher, P. (2012). Mutation of ATF4 mediates resistance of neuronal cell lines against oxidative stress by inducing xCT expression. Cell Death Differ. *19*, 847–858.

Leyland-Jones, B. (2003). Breast cancer trial with erythropoietin terminated unexpectedly. Lancet Oncol. *4*, 459–460.

Lipsic, E., Schoemaker, R.G., van der Meer, P., Voors, A.A., van Veldhuisen, D.J., and van Gilst, W.H. (2006). Protective effects of erythropoietin in cardiac ischemia: from bench to bedside. J. Am. Coll. Cardiol. *48*, 2161–2167.

Liu, L., Dilworth, D., Gao, L., Monzon, J., Summers, A., Lassam, N., and Hogg, D. (1999). Mutation of the CDKN2A 5' UTR creates an aberrant initiation codon and predisposes to melanoma. Nat. Genet. *21*, 128–132.

Livingstone, M., Atas, E., Meller, A., and Sonenberg, N. (2010). Mechanisms governing the control of mRNA translation. Phys. Biol. *7*, 021001.

Lohse, I., Reilly, P., and Zaugg, K. (2011). The CPT1C 5'UTR contains a repressing upstream open reading frame that is regulated by cellular energy availability and AMPK. PLoS ONE *6*, e21486.

Loughran, G., Sachs, M.S., Atkins, J.F., and Ivanov, I.P. (2012). Stringency of start codon selection modulates autoregulation of translation initiation factor eIF5. Nucleic Acids Res. *40*, 2898–2906.

Lovett, P.S., and Rogers, E.J. (1996). Ribosome regulation by the nascent peptide. Microbiol. Rev. *60*, 366–385.

Lu, P.D., Harding, H.P., and Ron, D. (2004). Translation reinitiation at alternative open reading frames regulates gene expression in an integrated stress response. J. Cell Biol. *167*, 27–33.

Lukowski, S.W., Bombieri, C., and Trezise, A.E.O. (2011). Disrupted posttranscriptional regulation of the cystic fibrosis transmembrane conductance regulator (CFTR) by a 5'UTR mutation is associated with a cftr-related disease. Hum. Mutat. *32*, E2266–E2282.

Luo, Z., Freitag, M., and Sachs, M.S. (1995). Translational regulation in response to changes in amino acid availability in Neurospora crassa. Mol. Cell. Biol. *15*, 5235–5245.

Madan, A., Lin, C., and Curtin, P.T. (1995). Regulated basal, inducible, and tissue-specific human erythropoietin gene expression in transgenic mice requires multiple cis DNA sequences. Blood *85*, 2735–2741.

Maiese, K., Chong, Z.Z., and Shang, Y.C. (2008). Raves and risks for erythropoietin. Cytokine Growth Factor Rev. *19*, 145–155.

Maquat, L.E., Kinniburgh, A.J., Rachmilewitz, E.A., and Ross, J. (1981). Unstable beta-globin mRNA in mRNA-deficient beta o thalassemia. Cell *27*, 543–553.

Marcotrigiano, J., Gingras, A.C., Sonenberg, N., and Burley, S.K. (1999). Cap-dependent translation

initiation in eukaryotes is regulated by a molecular mimic of eIF4G. Mol. Cell *3*, 707–716.

Marti, H.H., Wenger, R.H., Rivas, L.A., Straumann, U., Digicaylioglu, M., Henn, V., Yonekawa, Y., Bauer, C., and Gassmann, M. (1996). Erythropoietin gene expression in human, monkey and murine brain. Eur. J. Neurosci. *8*, 666–676.

Martins, R., Proença, D., Silva, B., Barbosa, C., Silva, A.L., Faustino, P., and Romão, L. (2012). Alternative polyadenylation and nonsense-mediated decay coordinately regulate the human HFE mRNA levels. PLoS ONE *7*, e35461.

Masutani, M., Sonenberg, N., Yokoyama, S., and Imataka, H. (2007). Reconstitution reveals the functional core of mammalian eIF3. EMBO J. *26*, 3373–3383.

Masutani, M., Machida, K., Kobayashi, T., Yokoyama, S., and Imataka, H. (2013). Reconstitution of eukaryotic translation initiation factor 3 by co-expression of the subunits in a human cell-derived in vitro protein synthesis system. Protein Expr. Purif. *87*, 5–10.

Mata, J., Marguerat, S., and Bähler, J. (2005). Post-transcriptional control of gene expression: a genome-wide perspective. Trends Biochem. Sci. *30*, 506–514.

Matsui, M., Yachie, N., Okada, Y., Saito, R., and Tomita, M. (2007). Bioinformatic analysis of post-transcriptional regulation by uORF in human and mouse. FEBS Lett. *581*, 4184–4188.

Maxwell, P.H., Wiesener, M.S., Chang, G.W., Clifford, S.C., Vaux, E.C., Cockman, M.E., Wykoff, C.C., Pugh, C.W., Maher, E.R., and Ratcliffe, P.J. (1999). The tumour suppressor protein VHL targets hypoxia-inducible factors for oxygen-dependent proteolysis. Nature *399*, 271–275.

Maynard, M.A., Qi, H., Chung, J., Lee, E.H.L., Kondo, Y., Hara, S., Conaway, R.C., Conaway, J.W., and Ohh, M. (2003). Multiple splice variants of the human HIF-3 alpha locus are targets of the von Hippel-Lindau E3 ubiquitin ligase complex. J. Biol. Chem. *278*, 11032–11040.

McGary, E.C., Rondon, I.J., and Beckman, B.S. (1997). Post-transcriptional regulation of erythropoietin mRNA stability by erythropoietin mRNA-binding protein. J. Biol. Chem. *272*, 8628–8634.

McGlincy, N.J., Tan, L.-Y., Paul, N., Zavolan, M., Lilley, K.S., and Smith, C.W.J. (2010). Expression proteomics of UPF1 knockdown in HeLa cells reveals autoregulation of hnRNP A2/B1 mediated by alternative splicing resulting in nonsense-mediated mRNA decay. BMC Genomics *11*, 565.

Medenbach, J., Seiler, M., and Hentze, M.W. (2011). Translational control via protein-regulated upstream open reading frames. Cell *145*, 902–913.

Mehta, A. (2006). Derepression of the Her-2 uORF is mediated by a novel post-transcriptional control mechanism in cancer cells. Genes Dev. *20*, 939–953.

Meijer, H.A., and Thomas, A.A.M. (2002). Control of eukaryotic protein synthesis by upstream open reading frames in the 5’-untranslated region of an mRNA. Biochem. J. *367*, 1–11.

Mendell, J.T., Sharifi, N.A., Meyers, J.L., Martinez-Murillo, F., and Dietz, H.C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nat. Genet. *36*, 1073–1078.

Mignone, F., Gissi, C., Liuni, S., and Pesole, G. (2002). Untranslated regions of mRNAs. Genome Biol *3*, 1–0004.

Mihailovich, M., Thermann, R., Grohovaz, F., Hentze, M.W., and Zacchetti, D. (2007). Complex translational regulation of BACE1 involves upstream AUGs and stimulatory elements within the 5’ untranslated region. Nucleic Acids Res. *35*, 2975–2985.

Mole, D.R., and Ratcliffe, P.J. (2007). Cellular oxygen sensing in health and disease. Pediatr. Nephrol. *23*, 681–694.

Moore, M.J. (2005). From birth to death: the complex lives of eukaryotic mRNAs. Science *309*, 1514–1518.

Moriconi, F., Ramadori, P., Schultze, F.C., Blaschke, M., Amanzada, A., Khan, S., and Ramadori, G. (2013). Characterization of the erythropoietin/erythropoietin receptor axis in a rat model of liver damage and cholangiocarcinoma development. Histochem. Cell Biol. *139*, 473–485.

Morino, S., Imataka, H., Svitkin, Y.V., Pestova, T.V., and Sonenberg, N. (2000). Eukaryotic translation initiation factor 4E (eIF4E) binding site and the middle one-third of eIF4GI constitute the core domain for cap-dependent translation, and the C-terminal one-third functions as a modulatory region. Mol. Cell. Biol. *20*, 468–477.

Morris, D.R. (1995). Growth control of translation in mammalian cells. Prog. Nucleic Acid Res. Mol. Biol. *51*, 339–363.

Morris, D.R. (1997). Cis-acting mRNA structures in gene-specific translational control in post-transcriptional gene regulation (Wiley-Liss Inc., New York).

Morris, D.R., and Geballe, A.P. (2000). Upstream open reading frames as regulators of mRNA translation. Mol. Cell. Biol. *20*, 8635–8642.

Mouton-Liger, F., Paquet, C., Dumurgier, J., Bouras, C., Pradier, L., Gray, F., and Hugon, J. (2012). Oxidative stress increases BACE1 protein levels through activation of the PKR-eIF2α pathway. Biochim. Biophys. Acta *1822*, 885–896.

Mueller, P.P., and Hinnebusch, A.G. (1986). Multiple upstream AUG codons mediate translational control of GCN4. Cell *45*, 201–207.

Mühlemann, O. (2008). Recognition of nonsense mRNA: towards a unified model. Biochem. Soc. Trans. *36*, 497–501.

Munzarová, V., Pánek, J., Gunišová, S., Dányi, I., Szamecz, B., and Valášek, L.S. (2011). Translation reinitiation relies on the interaction between eIF3a/TIF32 and progressively folded cis-acting mRNA elements preceding short uORFs. PLoS Genet. *7*, e1002137.

Nagy, E., and Maquat, L.E. (1998). A rule for termination-codon position within intron-containing genes: when nonsense affects RNA abundance. Trends Biochem. Sci. *23*, 198–199.

Nett, J.H., Gomathinayagam, S., Hamilton, S.R., Gong, B., Davidson, R.C., Du, M., Hopkins, D., Mitchell, T., Mallem, M.R., Nylen, A., et al. (2012). Optimization of erythropoietin production with controlled glycosylation-PEGylated erythropoietin produced in glycoengineered Pichia pastoris. J. Biotechnol. *157*, 198–206.

Nielsen, K.H., Szamecz, B., Valá\vsek, L., Jivotovskaya, A., Shin, B.S., and Hinnebusch, A.G. (2004). Functions of eIF3 downstream of 48S assembly impact AUG recognition and GCN4 translational control. EMBO J. *23*, 1166–1177.

Niesler, B., Flohr, T., Nöthen, M.M., Fischer, C., Rietschel, M., Franzek, E., Albus, M., Propping, P., and Rappold, G.A. (2001). Association between the 5’ UTR variant C178T of the serotonin receptor gene HTR3A and bipolar affective disorder. Pharmacogenetics *11*, 471–475.

Noguchi, C.T., Wang, L., Rogers, H.M., Teng, R., and Jia, Y. (2008). Survival and proliferative roles of erythropoietin beyond the erythroid lineage. Expert Rev. Mol. Med. *10*, e36–e36.

Nyikó, T., Sonkoly, B., Mérai, Z., Benkovics, A.H., and Silhavy, D. (2009). Plant upstream ORFs can trigger

nonsense-mediated mRNA decay in a size-dependent manner. Plant Mol. Biol. *71*, 367–378.

O'Connor, T., Sadleir, K.R., Maus, E., Velliquette, R.A., Zhao, J., Cole, S.L., Eimer, W.A., Hitt, B., Bembinster, L.A., Lammich, S., et al. (2008). Phosphorylation of the translation initiation factor eIF2alpha increases BACE1 levels and promotes amyloidogenesis. Neuron *60*, 988–1009.

Ohigashi, T., Mallia, C.S., McGary, E., Scandurro, A.B., Rondon, I., Fisher, J.W., and Beckman, B.S. (1999). Protein kinase C alpha protein expression is necessary for sustained erythropoietin production in human hepatocellular carcinoma (Hep3B) cells exposed to hypoxia. Biochim. Biophys. Acta *1450*, 109–118.

Oliner, J.D., Kinzler, K.W., Meltzer, P.S., George, D.L., and Vogelstein, B. (1992). Amplification of a gene encoding a p53-associated protein in human sarcomas. Nature *358*, 80–83.

Örd, T., Örd, D., Kõivomägi, M., Juhkam, K., and Örd, T. (2009). Human TRB3 is upregulated in stressed cells by the induction of translationally efficient mRNA containing a truncated 5'-UTR. Gene *444*, 24–32.

Palam, L.R., Baird, T.D., and Wek, R.C. (2011). Phosphorylation of eIF2 facilitates ribosomal bypass of an inhibitory upstream ORF to enhance CHOP translation. J. Biol. Chem. *286*, 10939–10949.

Paliege, A., Rosenberger, C., Bondke, A., Sciesielski, L., Shina, A., Heyman, S.N., Flippin, L.A., Arend, M., Klaus, S.J., and Bachmann, S. (2010). Hypoxia-inducible factor-2alpha-expressing interstitial fibroblasts are the only renal cells that express erythropoietin under hypoxia-inducible factor stabilization. Kidney Int. *77*, 312–318.

Palii, S.S., Kays, C.E., Deval, C., Bruhat, A., Fafournoux, P., and Kilberg, M.S. (2008). Specificity of amino acid regulated gene expression: analysis of genes subjected to either complete or single amino acid deprivation. Amino Acids *37*, 79–88.

Park, E.-H., Lee, J.M., Blais, J.D., Bell, J.C., and Pelletier, J. (2005). Internal translation initiation mediated by the angiogenic factor Tie2. J. Biol. Chem. *280*, 20945–20953.

Pasaje, C.F.A., Bae, J.S., Park, B.-L., Cheong, H.S., Kim, J.-H., Uh, S.-T., Park, C.-S., and Shin, H.D. (2012). WDR46 is a Genetic Risk Factor for Aspirin-Exacerbated Respiratory Disease in a Korean Population. Allergy Asthma Immunol. Res. *4*, 199–205.

Peixeiro, I., Inácio, Â., Barbosa, C., Silva, A.L., Liebhaber, S.A., and Romão, L. (2012). Interaction of PABPC1 with the translation initiation complex is critical to the NMD resistance of AUG-proximal nonsense mutations. Nucleic Acids Res. *40*, 1160–1173.

Pentecost, B., Song, R., Luo, M., DePasquale, J., and Fasco, M. (2005). Upstream regions of the estrogen receptor alpha proximal promoter transcript regulate ER protein expression through a translational mechanism. Mol. Cell. Endocrinol. *229*, 83–94.

Phrommintikul, A., Haas, S.J., Elsik, M., and Krum, H. (2007). Mortality and target haemoglobin concentrations in anaemic patients with chronic kidney disease treated with erythropoietin: a meta-analysis. Lancet *369*, 381–388.

Pisarev, A.V., Hellen, C.U.T., and Pestova, T.V. (2007). Recycling of eukaryotic posttermination ribosomal complexes. Cell *131*, 286–299.

Poyry, T.A.A., Kaminski,, A., and Jackson, R.J., R.J. (2004). What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? Genes Dev. *18*, 62–75.

Proud, C.G. (1994). Peptide-chain elongation in eukaryotes. Mol. Biol. Rep. *19*, 161–170.

Le Quesne, J.P.C., Spriggs, K.A., Bushell, M., and Willis, A.E. (2010). Dysregulation of protein synthesis and

disease. J. Pathol. *220*, 140–151.

Rajkowitsch, L., Vilela, C., Berthelot, K., Ramirez, C.V., and McCarthy, J.E.G. (2004). Reinitiation and recycling are distinct processes occurring downstream of translation termination in yeast. J. Mol. Biol. *335*, 71–85.

Raney, A., Law, G.L., Mize, G.J., and Morris, D.R. (2002). Regulated translation termination at the upstream open reading frame in s-adenosylmethionine decarboxylase mRNA. J. Biol. Chem. *277*, 5988–5994.

Raught, B., and Gingras, A. (2007). Signaling to translation initiation (In: Mathews M, Sonenberg N, Hershey JWB, editors. Translational Control in Biology and Medicine. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. pp 369-400).

Raveh-Amit, H., Maissel, A., Poller, J., Marom, L., Elroy-Stein, O., Shapira, M., and Livneh, E. (2009). Translational control of protein kinase Ceta by two upstream open reading frames. Mol. Cell. Biol. *29*, 6140–6148.

Re, A., Hamel, N., Fu, C., Bush, H., McCaffrey, T., and Asch, A.S. (2001). A link between diabetes and atherosclerosis: glucose regulates expression of CD36 at the level of translation. Nat. Med. *7*, 840–846.

Rehwinkel, J., Raes, J., and Izaurralde, E. (2006). Nonsense-mediated mRNA decay: Target genes and functional diversification of effectors. Trends Biochem. Sci. *31*, 639–646.

Rideau, A., Mangeat, B., Matthes, T., Trono, D., and Beris, P. (2007). Molecular mechanism of hepcidin deficiency in a patient with juvenile hemochromatosis. Haematologica *92*, 127–128.

Rogers Jr, G.W., Edelman, G.M., and Mauro, V.P. (2004). Differential utilization of upstream AUGs in the β-secretase mRNA suggests that a shunting mechanism regulates translation. Proc. Natl. Acad. Sci. U. S. A. *101*, 2794–2799.

Romão, L., Inácio, A., Santos, S., Avila, M., Faustino, P., Pacheco, P., and Lavinha, J. (2000). Nonsense mutations in the human beta-globin gene lead to unexpected levels of cytoplasmic mRNA accumulation. Blood *96*, 2895–2901.

Ron, D., and Harding, H. (2007). eIF2α phosphorylation in cellular stress responses and disease (In: Mathews M, Sonenberg N, Hershey JWB, editors. Translational Control in Biology and Medicine. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press. pp 345-368).

Rondon, I.J., MacMillan, L.A., Beckman, B.S., Goldberg, M.A., Schneider, T., Bunn, H.F., and Malter, J.S. (1991). Hypoxia up-regulates the activity of a novel erythropoietin mRNA binding protein. J. Biol. Chem. *266*, 16594–16598.

Roy, B., Vaughn, J.N., Kim, B.-H., Zhou, F., Gilchrist, M.A., and Von Arnim, A.G. (2010). The h subunit of eIF3 promotes reinitiation competence during translation of mRNAs harboring upstream open reading frames. RNA *16*, 748–761.

Ruifrok, W.P.T., de Boer, R.A., Westenbrink, B.D., van Veldhuisen, D.J., and van Gilst, W.H. (2008). Erythropoietin in cardiac disease: new features of an old drug. Eur. J. Pharmacol. *585*, 270–277.

De Ruijter, A.J.M., van Gennip, A.H., Caron, H.N., Kemp, S., and van Kuilenburg, A.B.P. (2003). Histone deacetylases (HDACs): characterization of the classical HDAC family. Biochem. J. *370*, 737–749.

Ryou, M.-G., Liu, R., Ren, M., Sun, J., Mallet, R.T., and Yang, S.-H. (2012). Pyruvate protects the brain against ischemia-reperfusion injury by activating the erythropoietin signaling pathway. Stroke J. Cereb. Circ. *43*, 1101–1107.

Sachs, M.S., and Geballe, A.P. (2006). Downstream control of upstream open reading frames. Genes Dev.

20, 915–921.

Sathirapongsasuti, J.F., Sathira, N., Suzuki, Y., Huttenhower, C., and Sugano, S. (2011). Ultraconserved cDNA segments in the human transcriptome exhibit resistance to folding and implicate function in translation and alternative splicing. Nucleic Acids Res. *39*, 1967–1979.

Scandurro, A.B., Rondon, I.J., Wilson, R.B., Tenenbaum, S.A., Garry, R.F., and Beckman, B.S. (1997). Interaction of erythropoietin RNA binding protein with erythropoietin RNA requires an association with heat shock protein 70. Kidney Int. *51*, 579–584.

Schepens, B., Tinton, S.A., Bruynooghe, Y., Beyaert, R., and Cornelis, S. (2005). The polypyrimidine tract-binding protein stimulates HIF-1alpha IRES-mediated translation during hypoxia. Nucleic Acids Res. *33*, 6884–6894.

Scoppetta, C., and Grassi, F. (2004). Erythropoietin: a new tool for muscle disorders? Med. Hypotheses *63*, 73–75.

Semenza, G.L. (2001). HIF-1 and mechanisms of hypoxia sensing. Curr. Opin. Cell Biol. *13*, 167–171.

Semenza, G.L., Dureza, R.C., Traystman, M.D., Gearhart, J.D., and Antonarakis, S.E. (1990). Human erythropoietin gene expression in transgenic mice: multiple transcription initiation sites and cis-acting regulatory elements. Mol. Cell. Biol. *10*, 930–938.

Sherry, S.T., Ward, M.-H., Kholodov, M., Baker, J., Phan, L., Smigielski, E.M., and Sirotkin, K. (2001). dbSNP: the NCBI database of genetic variation. Nucleic Acids Res. *29*, 308–311.

Shoemaker, C.B., and Mitsock, L.D. (1986). Murine erythropoietin gene: cloning, expression, and human gene homology. Mol. Cell. Biol. *6*, 849–858.

Shyu, A.-B., Wilkinson, M.F., and van Hoof, A. (2008). Messenger RNA regulation: to translate or to degrade. EMBO J. *27*, 471–481.

Silva, A.L., and Romão, L. (2009). The mammalian nonsense-mediated mRNA decay pathway: To decay or not to decay! Which players make the decision? FEBS Lett. *583*, 499–505.

Silva, A.L., Pereira, F.J.C., Morgado, A., Kong, J., Martins, R., Faustino, P., Liebhaber, S.A., and Romao, L. (2006). The canonical UPF1-dependent nonsense-mediated mRNA decay is inhibited in transcripts carrying a short open reading frame independent of sequence context. RNA *12*, 2160–2170.

Silva, A.L., Ribeiro, P., Inácio, A., Liebhaber, S.A., and Romão, L. (2008). Proximity of the poly(A)-binding protein to a premature termination codon inhibits mammalian nonsense-mediated mRNA decay. RNA *14*, 563–576.

Singh, G., Rebbapragada, I., and Lykke-Andersen, J. (2008). A competition between stimulators and antagonists of Upf complex recruitment governs human nonsense-mediated mRNA decay. PLoS Biol. *6*, e111.

Sivagnanasundaram, S., Morris, A.G., Gaitonde, E.J., McKenna, P.J., Mollon, J.D., and Hunt, D.M. (2000). A cluster of single nucleotide polymorphisms in the 5′-leader of the human dopamine D3 receptor gene (DRD3) and its relationship to schizophrenia. Neurosci. Lett. *279*, 13–16.

Sokabe, M., Fraser, C.S., and Hershey, J.W.B. (2011). The human translation initiation multi-factor complex promotes methionyl-tRNAi binding to the 40S ribosomal subunit. Nucleic Acids Res. *40*, 905–913.

Sonenberg, N., and Hinnebusch, A.G. (2009). Regulation of translation initiation in eukaryotes: mechanisms and biological targets. Cell *136*, 731–745.

Sözen, M.M., Whittall, R., Öner, C., Tokatlı, A., Kalkanoğlu, H.S., Dursun, A., Coşkun, T., Öner, R., and Humphries, S.E. (2005). The molecular basis of familial hypercholesterolaemia in Turkish patients. Atherosclerosis *180*, 63–71.

Spriggs, K.A., Bushell, M., and Willis, A.E. (2010). Translational Regulation of Gene Expression during Conditions of Cell Stress. Mol. Cell *40*, 228–237.

Stalder, L., and Mühlemann, O. (2008). The meaning of nonsense. Trends Cell Biol. *18*, 315–321.

Stein, R.S., Abels, R.I., and Krantz, S.B. (1991). Pharmacologic doses of recombinant human erythropoietin in the treatment of myelodysplastic syndromes. Blood *78*, 1658–1663.

Steinmann, K., Richter, A.M., and Dammann, R.H. (2011). Epigenetic silencing of erythropoietin in human cancers. Genes Cancer *2*, 65–73.

Stoneley, M., Paulin, F.E., Le Quesne, J.P., Chappell, S.A., and Willis, A.E. (1998). C-Myc 5′ untranslated region contains an internal ribosome entry segment. Oncogene *16*, 423–428.

Suzuki, N., Obara, N., Pan, X., Watanabe, M., Jishage, K.-I., Minegishi, N., and Yamamoto, M. (2011). Specific contribution of the erythropoietin gene 3′ enhancer to hepatic erythropoiesis after late embryonic stages. Mol. Cell. Biol. *31*, 3896–3905.

Suzuki, Y., Ishihara, D., Sasaki, M., Nakagawa, H., Hata, H., Tsunoda, T., Watanabe, M., Komatsu, T., Ota, T., Isogai, T., et al. (2000). Statistical analysis of the 5′ untranslated region of human mRNA using "Oligo-Capped" cDNA libraries. Genomics *64*, 286–297.

Szamecz, B., Rutkai, E., Cuchalová, L., Munzarová, V., Herrmannová, A., Nielsen, K.H., Burela, L., Hinnebusch, A.G., and Valásek, L. (2008). eIF3a cooperates with sequences 5′ of uORF1 to promote resumption of scanning by post-termination ribosomes for reinitiation on GCN4 mRNA. Genes Dev. *22*, 2414–2425.

Tahiri-Alaoui, A., Smith, L.P., Baigent, S., Kgosana, L., Petherbridge, L.J., Lambeth, L.S., James, W., and Nair, V. (2009). Identification of an intercistronic internal ribosome entry site in a Marek's disease virus immediate-early gene. J. Virol. *83*, 5846–5853.

Tassin, J., Dürr, A., Bonnet, A.M., Gil, R., Vidailhet, M., Lücking, C.B., Goas, J.Y., Durif, F., Abada, M., Echenne, B., et al. (2000). Levodopa-responsive dystonia. GTP cyclohydrolase I or parkin mutations? Brain J. Neurol. *123 (Pt 6)*, 1112–1121.

Tian, H., McKnight, S.L., and Russell, D.W. (1997). Endothelial PAS domain protein 1 (EPAS1), a transcription factor selectively expressed in endothelial cells. Genes Dev. *11*, 72–82.

Uddin, S., Kottegoda, S., Stigger, D., Platanias, L.C., and Wickrema, A. (2000). Activation of the Akt/FKHRL1 pathway mediates the antiapoptotic effects of erythropoietin in primary human erythroid progenitors. Biochem. Biophys. Res. Commun. *275*, 16–19.

Undén, J., Sjölund, C., Länsberg, J.-K., Wieloch, T., Ruscher, K., and Romner, B. (2013). Post-ischemic continuous infusion of erythropoeitin enhances recovery of lost memory function after global cerebral ischemia in the rat. BMC Neurosci. *14*, 27.

Valásek, L.S. (2012). "Ribozoomin"-translation initiation from the perspective of the ribosome-bound eukaryotic initiation factors (eIFs). Curr. Protein Pept. Sci. *13*, 305–330.

Valásek, L., Nielsen, K.H., and Hinnebusch, A.G. (2002). Direct eIF2-eIF3 contact in the multifactor complex is important for translation initiation in vivo. EMBO J. *21*, 5886–5898.

Vattem, K.M., and Wek, R.C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA

translation in mammalian cells. Proc. Natl. Acad. Sci. U. S. A. *101*, 11269–11274.

Wang, G.L., and Semenza, G.L. (1993). Characterization of hypoxia-inducible factor 1 and regulation of DNA binding activity by hypoxia. J. Biol. Chem. *268*, 21513–21518.

Wang, X.-Q., and Rothnagel, J.A. (2004). 5'-untranslated regions with multiple upstream AUG codons can support low-level translation via leaky scanning and reinitiation. Nucleic Acids Res. *32*, 1382–1391.

Wang, Z., and Sachs, M.S. (1997). Ribosome stalling is responsible for arginine-specific translational attenuation in Neurospora crassa. Mol. Cell. Biol. *17*, 4904–4913.

Wang, G.L., Jiang, B.-H., Rue, E.A., and Semenza, G.L. (1995). Hypoxia-inducible factor 1 is a basic-helix-loop-helix-PAS heterodimer regulated by cellular O2 tension. Proc. Natl. Acad. Sci. *92*, 5510–5514.

Warnecke, C., Zaborowska, Z., Kurreck, J., Erdmann, V.A., Frei, U., Wiesener, M., and Eckardt, K.-U. (2004). Differentiating the functional role of hypoxia-inducible factor (HIF)-1alpha and HIF-2alpha (EPAS-1) by the use of RNA interference: erythropoietin is a HIF-2alpha target gene in Hep3B and Kelly cells. FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. *18*, 1462–1464.

Watatani, Y., Ichikawa, K., Nakanishi, N., Fujimoto, M., Takeda, H., Kimura, N., Hirose, H., Takahashi, S., and Takahashi, Y. (2008). Stress-induced translation of ATF5 mRNA is regulated by the 5'-untranslated region. J. Biol. Chem. *283*, 2543–2553.

Watowich, S.S., Hilton, D.J., and Lodish, H.F. (1994). Activation and inhibition of erythropoietin receptor function: role of receptor dimerization. Mol. Cell. Biol. *14*, 3535–3549.

Wei, J., Wu, C., and Sachs, M.S. (2012). The arginine attenuator peptide interferes with the ribosome peptidyl transferase center. Mol. Cell. Biol. *32*, 2396–2406.

Weidemann, A., and Johnson, R.S. (2009). Nonrenal regulation of EPO synthesis. Kidney Int *75*, 682–688.

Wen, Y., Liu, Y., Xu, Y., Zhao, Y., Hua, R., Wang, K., Sun, M., Li, Y., Yang, S., Zhang, X.-J., et al. (2009). Loss-of-function mutations of an inhibitory upstream ORF in the human hairless transcript cause Marie Unna hereditary hypotrichosis. Nat. Genet. *41*, 228–233.

Wethmar, K., Smink, J.J., and Leutz, A. (2010a). Upstream open reading frames: Molecular switches in (patho)physiology. BioEssays *32*, 885–893.

Wethmar, K., Begay, V., Smink, J.J., Zaragoza, K., Wiesenthal, V., Dorken, B., Calkhoven, C.F., and Leutz, A. (2010b). C/EBP uORF mice--a genetic model for uORF-mediated translational control in mammals. Genes Dev. *24*, 15–20.

Wiesener, M.S., Jürgensen, J.S., Rosenberger, C., Scholze, C.K., Hörstrup, J.H., Warnecke, C., Mandriota, S., Bechmann, I., Frei, U.A., Pugh, C.W., et al. (2003). Widespread hypoxia-inducible expression of HIF-2alpha in distinct cell populations of different organs. FASEB J. Off. Publ. Fed. Am. Soc. Exp. Biol. *17*, 271–273.

Wiestner, A., Schlemper, R.J., van der Maas, A.P., and Skoda, R.C. (1998). An activating splice donor mutation in the thrombopoietin gene causes hereditary thrombocythaemia. Nat. Genet. *18*, 49–52.

Wittmann, J., Hol, E.M., and Jäck, H.-M. (2006). hUPF2 silencing identifies physiologic substrates of mammalian nonsense-mediated mRNA decay. Mol. Cell. Biol. *26*, 1272–1287.

Xu, T.-R., Lu, R.-F., Romano, D., Pitt, A., Houslay, M.D., Milligan, G., and Kolch, W. (2012). Eukaryotic translation initiation factor 3, subunit a, regulates the extracellular signal-regulated kinase pathway. Mol. Cell. Biol. *32*, 88–95.

Yaman, I., Fernandez, J., Liu, H., Caprara, M., Komar, A.A., Koromilas, A.E., Zhou, L., Snider, M.D.,

Scheuner, D., Kaufman, R.J., et al. (2003). The zipper model of translational control: a small upstream ORF is the switch that controls structural remodeling of an mRNA leader. Cell *113*, 519–531.

Yasuda, Y., Masuda, S., Chikuma, M., Inoue, K., Nagao, M., and Sasaki, R. (1998). Estrogen-dependent production of erythropoietin in uterus and its implication in uterine angiogenesis. J. Biol. Chem. *273*, 25381–25387.

Yasuda, Y., Maeda, Y., Koike, E., Watanabe, Y., Masuda, S., Yamasaki, H., Okumoto, K., Horiuchi, Y., and Hoshiai, H. (2010). Cancer cell lines' growth is promoted through individual responsiveness to autocrine and/or exogenous erythropoietin in vitro. In Recent Advances in Clinical Medicine (Anninos P, M Rossi, Pham TD, Falugi C, Bussing A, Koukkou M. Eds.), Proceedings of the International Conference on Oncology, University of Cambridge, UK, pp. 337–348.

Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M., and Muhlemann, O. (2011). Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. RNA *17*, 2108–2118.

Yin, H., and Blanchard, K.L. (2000). DNA methylation represses the expression of the human erythropoietin gene by two different mechanisms. Blood *95*, 111–119.

Zhang, L., Smit-McBride, Z., Pan, X., Rheinhardt, J., and Hershey, J.W.B. (2008). An oncogenic role for the phosphorylated h-subunit of human translation initiation factor eIF3. J. Biol. Chem. *283*, 24047–24060.

Zhao, C., Datta, S., Mandal, P., Xu, S., and Hamilton, T. (2010). Stress-sensitive regulation of IFRD1 mRNA decay is mediated by an upstream open reading frame. J. Biol. Chem. *285*, 8552–8562.

Zhou, W., and Song, W. (2006). Leaky scanning and reinitiation regulate BACE1 gene expression. Mol. Cell. Biol. *26*, 3353–3364.

Zhou, D., Palam, L.R., Jiang, L., Narasimhan, J., Staschke, K.A., and Wek, R.C. (2008a). Phosphorylation of eIF2 directs ATF5 translational control in response to diverse stress conditions. J. Biol. Chem. *283*, 7064–7073.

Zhou, F., Roy, B., and von Arnim, A.G. (2010). Translation reinitiation and development are compromised in similar ways by mutations in translation initiation factor eIF3h and the ribosomal protein RPL24. BMC Plant Biol. *10*, 193.

Zhou, M., Sandercock, A.M., Fraser, C.S., Ridlova, G., Stephens, E., Schenauer, M.R., Yokoi-Fong, T., Barsky, D., Leary, J.A., and Hershey, J.W. (2008b). Mass spectrometry reveals modularity and a complete subunit interaction map of the eukaryotic translation factor eIF3. Proc. Natl. Acad. Sci. *105*, 18139–18144.

Zhou, Q.-H., Hui, E.K.-W., Lu, J.Z., Boado, R.J., and Pardridge, W.M. (2011). Brain penetrating IgG-erythropoietin fusion protein is neuroprotective following intravenous treatment in Parkinson's disease in the mouse. Brain Res. *1382*, 315–320.

Zhu, Y., Sun, Y., Mao, X., Jin, K., and Greenberg, D. (2002). Expression of poly (C)-binding proteins is differentially regulated by hypoxia and ischemia in cortical neurons. Neuroscience *110*, 191–198.