

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE BIOLOGIA VEGETAL



**ESTUDO DO CONTROLO DA EXPRESSÃO GÉNICA DA ERITROPOIETINA  
HUMANA POR UMA PEQUENA GRELHA DE LEITURA A MONTANTE DA  
GRELHA PRINCIPAL**

**Cristina Maria Botelho da Rocha Barbosa**

MESTRADO EM BIOLOGIA MOLECULAR HUMANA

2009

UNIVERSIDADE DE LISBOA  
FACULDADE DE CIÊNCIAS  
DEPARTAMENTO DE BIOLOGIA VEGETAL



**ESTUDO DO CONTROLO DA EXPRESSÃO GÉNICA DA ERITROPOIETINA  
HUMANA POR UMA PEQUENA GRELHA DE LEITURA A MONTANTE DA  
GRELHA PRINCIPAL**

Dissertação orientada por:  
Doutora Luísa Romão  
(Departamento de Genética, Instituto Nacional de Saúde Dr. Ricardo Jorge)  
Prof. Doutora Rita Zilhão  
(Departamento de Biologia Vegetal, Faculdade de Ciências da Universidade de Lisboa)

**Cristina Maria Botelho da Rocha Barbosa**

MESTRADO EM BIOLOGIA MOLECULAR HUMANA

2009

## Agradecimentos

Ao Doutor João Lavinha que, na qualidade de Assessor do ex-Centro de Genética Humana e actual responsável da Unidade de Investigação e Desenvolvimento do Departamento de Genética, autorizou a realização deste trabalho, assim como ao Prof. Doutor Luís Nunes, coordenador do Departamento de Genética e ao Prof. Doutor José Pereira Miguel, Presidente do Conselho Directivo do Instituto Nacional de Saúde Dr. Ricardo Jorge.

À Doutora Luísa Romão por ter me ter recebido, por ter tornado possível este trabalho, e pela sua orientação e disponibilidade.

À Prof. Doutora Rita Zilhão que na sua qualidade de orientadora interna demonstrou uma elevada disponibilidade e ajuda para além do interesse no presente projecto.

Aos meus colegas de laboratório, Isabel Peixeiro, Filipa Tomé, Ana Morgado, Rute Martins, Ana Luísa Silva, Francisco Pereira e Alexandre Teixeira, pela constante ajuda e discussões científicas e por me terem incluído no seu seio laboratorial. Um agradecimento também aos colegas mais distantes mas igualmente presentes das Hemoglobinopatias (II), Bruno Silva, Andreia Coelho, Taciana Moniz, Isabel Moutinho e Paula Faustino. Um obrigado a todos por se terem tornado mais do que simples colegas de laboratório, principalmente à Isabel Peixeiro por ter sido uma autêntica orientadora a nível prático e por ter tornado os meus dias mais divertidos juntamente com o Tiago.

À Ângela Inácio por me ter iniciado neste projecto e em toda a vida científica do INSA, para além de me ter ajudado a dar os primeiros passos.

Aos meus amigos Ulrica, MaC, Joana Cruz, Tolas, Thomas, Rui, Acetona, Filipa, Rita, Lara e Marta Carvalho, que foram tolerantes em relação a todas as minhas ausências e pacientes em ter que me ouvir a falar de trabalho em todas as presenças. Um agradecimento especial ao Pedro Neto por ter aturado tudo e mais alguma coisa.

Dedico esta tese aos meus pais pelo apoio, pela compreensão demonstrada independentemente do meu estado de espírito e pela mente aberta que tiveram para aceitar esta minha escolha de profissão.

## Resumo

Vários elementos do mRNA, que actuam em *cis*, participam na regulação da síntese proteica, incluindo codões AUGs na região líder do transcrito, ou, em alguns casos, grelhas de leitura a estes associadas (*upstream open reading frames* – uORFs). Embora aproximadamente 15% dos mRNAs humanos (principalmente, transcritos codificantes de factores de crescimento ou hormonas) contenham uORFs, na maioria dos casos, o princípio pelo qual a uORF medeia o controlo traducional é pouco compreendido.

A eritropoietina (EPO) é uma glicoproteína sintetizada e excretada principalmente pelo rim, cuja função hormonal clássica envolve um papel chave na eritropoiese. No entanto, diversas funções não-hematopoiéticas foram recentemente descritas. Consequentemente, a EPO pode ser utilizada como alvo terapêutico para diversas disfunções humanas. A compreensão dos mecanismos moleculares do controlo traducional do mRNA da EPO pode ser útil na determinação dessas terapias. Sabendo que o transcrito da EPO humana apresenta uma região 5' líder com 181 nucleótidos contendo uma uORF de 14 codões e dado que parece ser conservada entre espécies diferentes, o que parece indicar a sua função na regulação traducional, o objectivo foi investigar esta hipótese.

De forma a analisar o efeito desta uORF, as linhas celulares HepG2 e HEK293 foram transfectadas com diversas construções usando o gene repórter da luciferase com a uORF intacta ou inactiva. A actividade da luciferase foi medida por ensaios de luminescência e normalizada aos correspondentes níveis de mRNA para obter as eficiências traducionais. Os níveis de mRNA foram quantificados por RT-PCR quantitativo. Os resultados revelaram que a uORF consegue diminuir a eficiência traducional da ORF principal em cerca de 70%. Adicionalmente, os resultados sustentam a conclusão de que o *leaky scanning* reconhece o codão AUG principal. Para além disso, a uORF é capaz de responder à falta de nutrientes aumentando a quantidade de proteína produzida, de uma forma específica de tecido.

Palavras-chave: expressão génica; tradução; grelha de leitura a montante da grelha de leitura principal (*upstream open reading frames* – uORF); *leaky scanning*; eritropoietina (EPO)

## Abstract

Among the various *cis*-acting elements in mRNAs that participate in regulating protein synthesis are AUG codons within transcript leaders and, in some cases, associated upstream open reading frames (uORFs). Although about 15% of the human mRNAs (mainly, transcripts encoding growth factors or hormones) present uORFs, for the vast majority of the cases, the principles by which uORFs participate in translational control are still poorly understood.

In its classical hormonal role, erythropoietin (EPO) is a glycoprotein synthesized and released mainly from the kidney, which has a key role in hematopoiesis. However, many other non-hematopoietic functions have recently been reported. Consequently, it might be used as a therapeutic target for the treatment of several human disorders. Understanding the molecular mechanisms of translational control of EPO mRNA may be valuable in the determination of these therapies. Knowing that human EPO transcript presents a 5' leader with 181 nucleotides containing a 14-codon-uORF and given that it seems to be conserved among different species, which might indicate its role in translational regulation, we aimed to prove this hypothesis.

To analyze the effect of EPO's uORF, HepG2 and HEK293 cells were transfected with several constructs carrying the luciferase reporter gene with the intact or disrupted uORF. Luciferase activity was measured by luminometry and normalized to the corresponding mRNA levels to obtain translation efficiencies. The mRNA levels were quantified by real-time RT-PCR. Results show that the uORF can decrease the main ORF translation efficiency by 70%. In addition, results support the conclusion that the native AUG codon is recognized by leaky scanning. Furthermore, the uORF is able to respond to the lack of nutrients with an increased amount of protein produced, in a tissue-specific way.

Keywords: genetic expression; translation; upstream open reading frames (uORF); leaky scanning; erythropoietin (EPO)

## Abreviaturas

A	Adenina
Amp	Ampicilina
bp	pares de bases ( <i>base pairs</i> )
C	Citosina
<i>cis</i>	No mesmo elemento
DH5 $\alpha$	Estirpe de bactérias <i>E. coli</i>
DMEM	<i>Dulbecco's Modified Eagle Medium</i>
DNA	Ácido desoxirribonucleico
ECL	<i>Enhanced chemiluminescence</i>
eIF	Factor de iniciação eucariota
EPO	Eritropoietina
ERBP	Proteína de ligação ao mRNA da eritropoietina ( <i>EPO mRNA-Binding Protein</i> )
FBS	Soro fetal bovino ( <i>Bovine Fetal Serum</i> )
G	Guanosina
GDP	Guanosina difosfato
GTP	Guanosina trifosfato
h	Hora(s)
H <sub>2</sub> O <sub>2</sub>	Peróxido de hidrogénio
HBS	Local de ligação do HIF-1 ( <i>HIF-<math>\alpha</math>-binding site</i> )
HEK293	Linha celular derivada de rim embrionário
HepG2	Linha celular derivada de carcinoma hepatocelular
HIF-1	Factor induzido por hipoxia 1 ( <i>Hipoxia-Inducible Factor 1</i> )
i.e.	isto é ( <i>id est</i> )
IRES	Sequência interna de entrada do ribossoma ( <i>Internal Ribosome Entry Site</i> )
kDa	quiloDalton
LB	Luria-Bertani
Met	Metionina
min	Minuto(s)
mRNA	RNA mensageiro
n	Qualquer nucleótido

---

N-terminal	Região terminal amínica
NaCl	Cloreto de sódio
NMD	Decaimento do mRNA mediada por mutação <i>nonsense</i> ( <i>Nonsense Mediated mRNA Decay</i> )
ORF	Grelha de leitura ( <i>Open Reading Frame</i> )
PABP	Proteína de ligação à cauda poli(A) ( <i>Poly(A)-Binding Protein</i> )
PCR	Reacção de polimerização em cadeia ( <i>Polymerase chain reaction</i> )
PLB	Tampão de lise passiva ( <i>Passive Lysis Buffer</i> )
RNA	Ácido ribonucleico
rpm	Rotações por minuto
RPMI	Meio rico ( <i>Roswell Park Memorial Institute</i> )
seg	Segundo(s)
SDS	Dodecil sulfato de sódio ( <i>Sodium Dodecyl Sulfate</i> )
T	Timina
TBST	<i>Tris-Buffered Saline Triton X-100</i>
TEMED	N,N,N',N'-Tetrametiletilenodiamina ( <i>N,N,N',N'-Tetramethylethylenediamine</i> )
Tris	tris(hidroximetil)aminometano ( <i>tris(hydroxymethyl)aminomethane</i> )
tRNA	RNA transportador
U	Uracilo
uORF	Grelha de leitura a montante da grelha de leitura principal ( <i>upstream Open Reading Frame</i> )
UTR	Região não traduzida ( <i>Untranslated Region</i> )

## Índice

Agradecimentos.....	iii
Resumo.....	iv
Abstract.....	v
Abreviaturas.....	vi
Índice.....	viii
I. Introdução.....	1
I.1. Regulação da expressão génica.....	1
I.2. Grelhas de leitura a montante da grelha de leitura principal (uORFs).....	4
I.3. Eritropoietina.....	8
II. Metodologia experimental.....	11
II.1 Construção de vectores de expressão.....	11
II.2 Culturas celulares.....	14
II.3 Transfecção de células em cultura.....	15
II.4 Transfecção de células em cultura com privação de soro.....	15
II.5 Ensaio de luminescência.....	16
II.6 RT-PCR quantitativo.....	16
II.7 Imuno-deteção por <i>Western Blot</i> .....	17
III. Resultados.....	19
III.1 A uORF da EPO inibe a eficiência de tradução da ORF principal.....	19
III.2 A inibição da expressão da EPO, exercida pela uORF, responde a estímulos externos.....	21
III.3 A iniciação da tradução no AUG principal da EPO ocorre pelo mecanismo de <i>leaky scanning</i> .....	23
Discussão.....	26
Bibliografia.....	29
Anexo A – Oligonucleótidos, Vectores e Sequências.....	31
Anexo B – Soluções para imuno-deteção por <i>Western Blot</i> .....	34



## I. Introdução

### I.1. Regulação da expressão génica

A expressão génica é indispensável para a definição das características de todos os organismos, sejam eles procariotas ou eucariotas. Sobretudo, é fundamental a dinâmica entre o DNA, o RNA e a maquinaria de regulação da expressão para que seja possível integrar e responder a toda a informação intrínseca e extrínseca a que as células são sujeitas <sup>(1)</sup>.

Diversos mecanismos de controlo da expressão da informação genética têm sido descritos. Com efeito, recentes análises revelam que somente 1,5% do genoma de eucariotas é codificante para proteínas, estando o restante envolvido na regulação da expressão génica <sup>(2)</sup>. Um destes mecanismos ocorre a um nível transcricional e tem merecido elevado ênfase nos últimos anos. Entre os seus efectores incluem-se a DNA polimerase, elementos localizados no DNA que actuam em *cis*, como promotores e *enhancers*, factores de transcrição e mais recentemente a remodelação global da cromatina <sup>(1, 2)</sup>. Contudo, elevada complexidade tem vindo a ser introduzida como resultado da descoberta de processos de regulação pós-transcricional, que por sua vez, incluem o controlo do processamento, exportação, localização, *turnover* e tradução do mRNA. Adicionalmente, verificou-se que o controlo da tradução permite uma resposta rápida dos níveis de proteína sintetizada permitindo manter a homeostase, contrastando com a regulação transcricional que modula alterações mais permanentes na fisiologia e destino celular <sup>(3)</sup>.

A regulação da tradução pode ocorrer em qualquer uma das fases que a compõem: iniciação, alongamento e terminação. Deste processo pode resultar um controlo a nível global ou a nível específico de uma determinada proteína ou grupo de proteínas. O primeiro é relevante aquando das primeiras etapas da embriogénese ou durante determinadas fases do ciclo celular e envolve essencialmente factores de iniciação da tradução ou proteínas presentes no ribossoma <sup>(4)</sup>. O segundo é principalmente mediado por elementos nas regiões 5' líder do transcrito ou nas regiões não traduzidas (UTR - *untranslated region*) a 3' <sup>(5)</sup>.

A maioria dos mRNA maduros estudados em eucariotas obedece a uma estrutura típica que consiste numa estrutura *cap* a 5' (ligação no sentido 5'-5' de uma guanosina que é posteriormente metilada) seguida de uma sequência líder na extremidade 5' (tipicamente designada por 5' UTR), um codão de iniciação, uma única grelha de leitura, cujo final ocorre num codão de terminação, que por sua vez é seguido por uma região a 3' UTR e uma cauda poli(A) <sup>(6)</sup>.

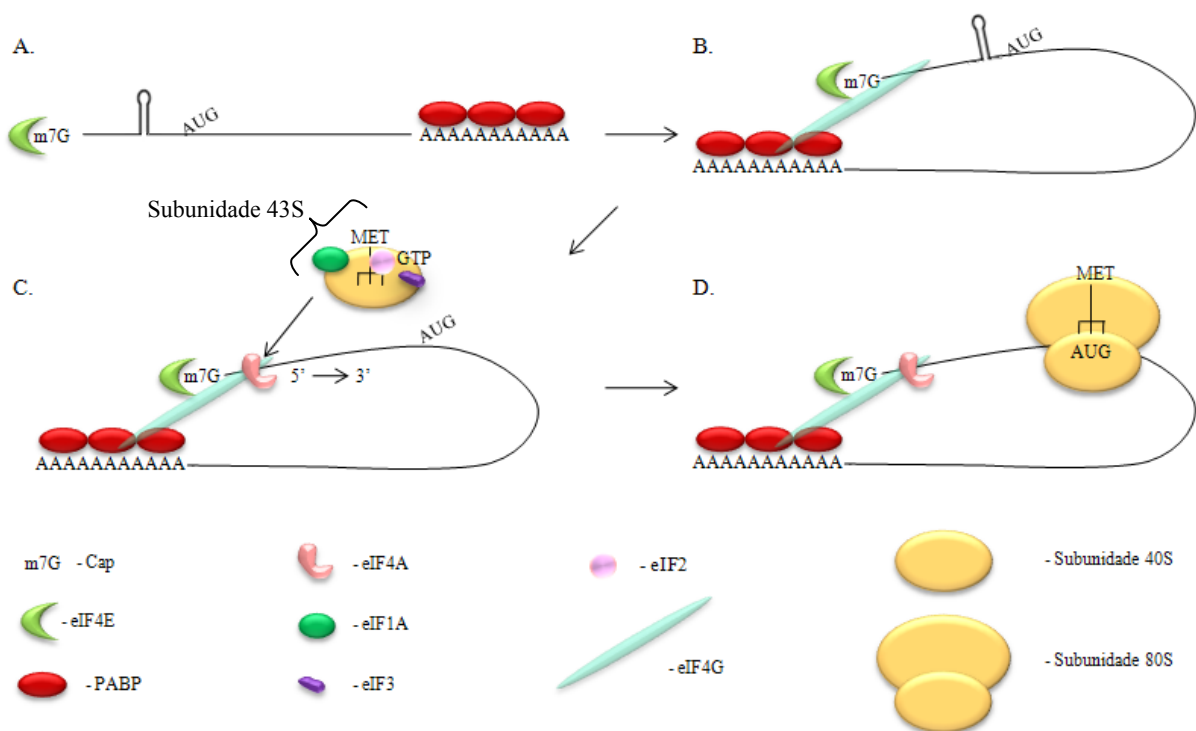
A tradução é baseada no movimento do ribossoma, no sentido 5'→3', do mRNA que está a ser traduzido. O movimento inicial do ribossoma designa-se por *scanning*. O primeiro passo deste processo é mediado pelo factor de iniciação eIF4E. Em células em repouso este factor está associado à proteína de ligação eIF4E-BP (eIF4E-binding protein), no entanto, aquando da fosforilação desta, o eIF4E é libertado tornando-se disponível para se ligar à estrutura *cap* e a outro factor, o eIF4G. (Fig.1A). Por sua vez o eIF4G parece ser o agente envolvido na circularização do mRNA, visto conter um local de ligação para o eIF4E e outro para a proteína de ligação à cauda poli(A) (PABP – *poly(A)-binding protein*) (Fig.1B) <sup>(7)</sup>.

O segundo passo no mecanismo de iniciação da tradução envolve a ligação de uma helicase dependente de ATP, a eIF4A, juntamente com dois factores que estimulam a sua actividade, o eIF4B e o eIF4H, ao eIF4G. Desta forma, a região 5' líder do mRNA é desprovida de estruturas secundárias, pela actividade helicásica da eIF4A, o que permite a posterior ligação da subunidade ribossomal 40S (Fig.1C). A associação de eIF1A e eIF3 à subunidade 40S do ribossoma, ainda antes da sua ligação ao mRNA, impede a associação da subunidade 60S, ao mesmo tempo que permite a ligação do complexo terciário eIF2-GTP-Met-tRNA<sub>i</sub>, formando-se o complexo de pré-iniciação 43S (Fig.1C). O complexo de pré-iniciação liga-se ao mRNA junto à estrutura *cap* e move-se no sentido 5' para 3' até reconhecer um codão de iniciação AUG. Este reconhecimento é mediado por eIF1, eIF2 e pela complementaridade entre as bases do tRNA e do mRNA <sup>(7)</sup>. Caso se dê o reconhecimento do codão de iniciação, o factor eIF5B estimula a hidrólise do GTP, provocando a dissociação de eIF2-GDP e provavelmente de outros factores de iniciação presentes neste processo. O factor eIF5, por sua vez, catalisa a associação da subunidade 60S formando o ribossoma 80S e consequentemente dá-se o início do alongamento (Fig.1D). Neste processo, é produzida uma cadeia de resíduos de aminoácidos que se separa aquando do reconhecimento de um codão de terminação (UAG, UAA, UGA). No processo de terminação o ribossoma dissocia-se nas suas subunidades libertando-se do mRNA <sup>(7)</sup>.

O processo de iniciação da tradução anteriormente referido é dependente da estrutura *cap*, no entanto, têm sido descritos vários mecanismos e mRNAs atípicos. Um exemplo, foi a descrição de um mecanismo independente desta estrutura, em que o ribossoma é recrutado para a proximidade do codão de iniciação de uma forma análoga às sequências *Shine-Dalgarno* em bactérias, mas mais complexa, por reconhecimento das estruturas IRES (*internal ribosome entry site*) <sup>(3)</sup>.

Os mRNAs contendo estruturas IRES foram descobertos em picornavírus e posteriormente foram encontrados em mRNAs de organismos superiores, incluindo humanos.

Estes mRNAs são principalmente codificantes para proteínas reguladoras como proto-oncogenes, factores de crescimento e seus receptores <sup>(2)</sup>. Não obstante deste processo alternativo de iniciação de tradução, muito dos transcritos que contêm IRES podem também ser traduzidos pelo mecanismo convencional descrito anteriormente, o que revela a possibilidade das estruturas IRES permitirem mudanças ao nível da tradução de forma a expressar, em determinadas circunstância, proteínas essenciais para manter a sobrevivência ou morte da célula <sup>(8)</sup>. Todavia, o mecanismo pelo qual actuam não está completamente esclarecido e nenhuma sequência de consenso foi identificada, impossibilitando o reconhecimento dos transcritos que poderão conter estas estruturas <sup>(9)</sup>.



**Figura 1- Mecanismo de iniciação da tradução.** (A) Primeiro passo da tradução do mRNA consiste na ligação de eIF4E à estrutura *cap*; (B) ligação do eIF4G que interage com o eIF4E e PABP provocando a circularização do mRNA; (C) ligação de eIF4A que permite a resolução de estruturas secundárias e a associação do complexo de pré-iniciação 43S ao mRNA; (D) reconhecimento do AUG com subsequente dissociação de factores de iniciação e formação da subunidade 80S. Adaptado de Mangus *et. al* (2003) e Jackson (2005) <sup>(13, 14)</sup>.

Existem outros elementos que actuam *em cis* que têm a função de modular a expressão génica ao nível da tradução, que, como já referido, estão localizados na região 5' líder ou 3' UTR do transcrito. Exemplos destes elementos são: (1) estruturas secundárias estáveis que bloqueiam a passagem da subunidade 43S do ribossoma quando esta percorre o mRNA; (2) codões de iniciação a montante da grelha de leitura principal (uAUGs) e, por vezes, grelhas de leitura a estes associadas (uORFs), que podem resultar numa diminuição da tradução da ORF

principal; (3) proteínas de ligação ao mRNA num local próximo da estrutura *cap* que inibem o seu reconhecimento por um impedimento estérico; (4) locais de reconhecimento de proteínas de ligação que estimulam ou inibem a tradução como resposta a determinados estímulos; (5) RNAs não codificantes (microRNAs) que reconhecem principalmente sequências na 3' UTR e que inibem a tradução<sup>(3, 10)</sup>.

Outras formas de regulação pós-transcricional afectam a estabilidade do mRNA. Um exemplo é a activação da via de decaimento do mRNA mediado por mutações *nonsense* (NMD – *nonsense mediated mRNA decay*). Este mecanismo regula cerca de 5-10% dos transcritos em levedura, *Drosophila* e humanos<sup>(11)</sup> e ocorre não só para proteger as células de mutações *nonsense*, que podem resultar em proteínas truncadas deletérias<sup>(1)</sup>, mas também como forma de regulação de transcritos com uORF naturais, cujo codão de terminação é reconhecido como prematuro<sup>(2)</sup>.

Estima-se que 12-22% dos genes contenham regiões 5' líder alternativas resultantes de eventos de *splicing* e de local de início de transcrição alternativos. Estes processos permitirem a remodelação da região 5' líder do transcrito tanto no número como no tipo de elementos reguladores da tradução presentes, principalmente uAUGs ou uORFs<sup>(12)</sup>. O presente trabalho tem por objectivo o estudo regulação da expressão génica mediada pela presença de uORFs.

## **I.2. Grelhas de leitura a montante da grelha de leitura principal (uORFs)**

Uma região 5' líder possui, em média, 100 a 200 nucleótidos, número aproximadamente constante em diversas classes taxonómicas. Porém, as 3'UTRs são mais variáveis, tanto entre espécies, podendo estender-se entre os 200 nucleótidos em plantas e fungos e 800 nucleótidos em humanos e outros vertebrados, como dentro da mesma espécie, variando entre dezenas a centenas de nucleótidos. Geralmente, as regiões 5' líder apresentam maior conteúdo em GCs do que as 3'UTRs. Por exemplo, no caso de vertebrados de sangue quente é estimado que este conteúdo nas primeiras seja de cerca de 60% e nas segundas aproximadamente 45%<sup>(2)</sup>.

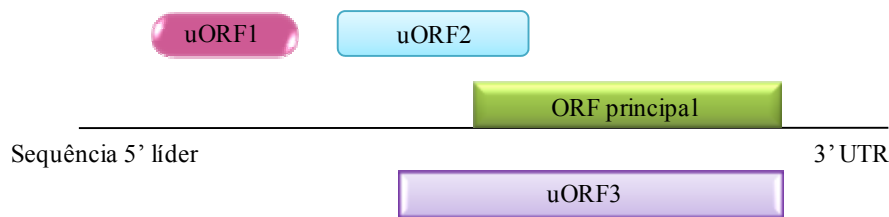
Como referido anteriormente, as características estruturais das regiões 5' líder são relevantes para o controlo da tradução. Fundamentando este facto está a observação de que estas regiões se apresentam, frequentemente, maiores e mais complexas, do que é referido como normal, em genes essenciais para as células, que necessitam de ser estritamente regulados. Na sua maioria, as regiões 5' líder dos correspondentes transcritos apresentam uAUGs, uORFs ou estruturas secundárias estáveis<sup>(2)</sup>.

Diversos estudos pretenderam quantificar uAUGs e uORFs presentes nas diferentes classes taxonómicas. No entanto, esta tarefa mostrou-se extremamente difícil, principalmente por não terem sido determinadas todas as sequências 5' dos genes nos organismos em geral, desde humanos até organismos mais simples como leveduras <sup>(15)</sup>. Logo, nem todos os locais de início da transcrição estão correctamente determinados e uma fracção dos mRNAs pode estar representada no seu estado imaturo, ainda com intrões, e assim não expor regiões 5' líder funcionais <sup>(7)</sup>. Adicionalmente, locais de início da transcrição, processamento de mRNA e codões de iniciação alternativos complicam a determinação da constituição exacta da região 5' líder do transcrito. Não obstante, é notória a sua presença em genes essenciais para a célula, tais como, oncogenes, supressores tumorais, e genes envolvidos no crescimento e diferenciação celular <sup>(16, 17)</sup>. Apesar destes constrangimentos, as estimativas mais recentes apontam para que os uAUGs estejam presentes em cerca de 16-27% dos mRNAs de plantas, 48-52% em invertebrados, 48% em aves, 37-49% em mamíferos e 44% em humanos. Da percentagem obtida em humanos cerca de 18% representam transcritos com três ou mais uAUGs <sup>(18, 19)</sup>, e cerca de 15% poderão corresponder a uORFs <sup>(7)</sup>.

As uORFs têm sido envolvidas na redução dos níveis de proteína em aproximadamente 100 genes eucariotas. Adicionalmente, foram também descritas mutações que introduzem ou interrompem uma uORF como causa de 3 doenças humanas: trombocitopenia hereditária <sup>(20)</sup>, melanoma (aumento da predisposição) <sup>(21)</sup> e hipotricose hereditária de Marie Unna <sup>(22)</sup>. Porém, nenhum estudo comprovou se estes elementos têm um efeito generalizado nos níveis de proteína celular e se existe algum polimorfismo na presença ou tamanho das uORFs na população humana <sup>(23)</sup>. Todos estes factos revelam a necessidade de mais estudos e de uma melhor compreensão dos mecanismos de actuação e relevância biológica das uORFs.

As uORFs são então definidas pela ocorrência de uma grelha de leitura, com um codão de iniciação e um codão de terminação em fase, a montante de uma grelha de leitura principal. No entanto, apesar dos codões de iniciação estarem a montante da ORF principal, o codão de terminação pode situar-se em diferentes posições em relação a esta. O codão de terminação pode estar a montante, ficando toda a uORF localizada na região 5' líder (Fig.2 – uORF1), dentro da região codificante, sobrepondo-se com a ORF principal (Fig.2 – uORF2), ou pode ser o mesmo da ORF principal, formando uma versão extensa da mesma (Fig.2 uORF3) <sup>(18)</sup>. Neste último tipo de uORFs são produzidas isoformas da mesma proteína que diferem na região N-terminal. Visto que esta região pode ser responsável pela localização celular, é possível que seja uma forma mais económica para a célula de produzir a mesma proteína com

destino a compartimentos celulares diferentes. Estima-se que este processo ocorra em cerca de 3-5% dos mRNAs de plantas e mamíferos <sup>(18)</sup>.



**Figura 2- Diferentes tipos de *upstream open reading frames* (ORFs).** A uORF1 está totalmente localizada na sequência 5' líder, a uORF2 sobrepõe-se com a ORF principal e a uORF3 é uma forma mais extensa da ORF principal. Adaptado de Kochetov *et. al* (2008).

Relembrando o mecanismo de iniciação da tradução anteriormente descrito, segundo o qual, o ribossoma reconhece, geralmente, o primeiro AUG, é evidente o constrangimento que pode ser causado pela existência de uORFs. Sendo assim, serão necessários processos alternativos de tradução para que a proteína principal seja produzida. Um desses processos é o *leaky scanning*, em que a subunidade 43S do ribossoma que percorre o mRNA pode reconhecer ou não o primeiro AUG que encontra, e no caso de existir uma uORF, dependendo das condições, pode ou não traduzi-la. O reconhecimento do AUG pode ser afetado pelos nucleótidos que o flanqueiam, ou seja, o contexto em que está inserido, pela sua proximidade à estrutura *cap*, pelo comprimento da região líder, e pela presença de estruturas secundárias tanto a montante como a jusante do AUG <sup>(16,24)</sup>.

O contexto considerado ótimo em que o AUG pode estar inserido é GCC(A/G)CCAUGG, no qual as posições mais importantes são a -3 e +4 (a negrito). Um contexto considerado adequado é AnnAUGn ou GnnAUGGG e um contexto mais fraco é (C/T)nnAUG(A/C/T) <sup>(18)</sup>. As ORFs principais têm 95-97% de AUGs em contexto ótimo ou adequado, no entanto, este valor é menor no caso das uORFs, sendo de 43-63% <sup>(7)</sup>. No caso da existência de um uAUG, o contexto em que o AUG da ORF principal está inserido não interfere no mecanismo de tradução da uORF. É possível ainda que codões AUG em contexto fraco ou codões não AUG (CUG, ACG e GUG) sejam reconhecidos <sup>(25)</sup>. Porém, é comum serem necessários sinais adicionais como estruturas secundárias especificamente localizadas a 17-30 nucleótidos a jusante do codão de iniciação ou um contexto perfeito em codões não AUG <sup>(18)</sup>. Especificamente no caso de existirem estruturas secundárias, pensa-se que estas provoquem uma paragem do ribossoma que desta forma tem mais tempo para o

reconhecimento do codão de iniciação. Se a pausa for longa o suficiente, o eIF5, que funciona como um dispositivo temporal, recruta o complexo eIF2-GTP-Met-tRNA<sub>i</sub> e começa a fase de alongamento <sup>(26)</sup>.

Outro processo que permite a tradução da ORF principal é a reiniciação. Neste caso, a uORF é traduzida, no entanto, em vez do ribossoma se separar nas suas subunidades e estas se dissociarem do mRNA após a terminação da tradução da uORF, a subunidade 40S permanece associada, ao contrário da 60S, podendo retomar o seu movimento de *scanning* do mRNA e reconhecer outro AUG mais a jusante <sup>(7, 9)</sup>. O mecanismo de reiniciação pode ocorrer após a tradução de uma ORF pequena mas não após uma ORF maior, contudo, o factor crítico não parece ser o tamanho da uORF mas o tempo que demora a tradução. Esta dependência do tempo relaciona-se com a necessidade do ribossoma readquirir eIF2-GTP-Met-tRNA<sub>i</sub> e com a necessidade de que determinados factores de iniciação permanecerem ligados durante a fase de alongamento <sup>(27, 28)</sup>. É extremamente difícil definir um comprimento máximo da uORF que permita a reiniciação, no entanto, verifica-se que, na maior parte dos casos, uma uORF com mais do que 30 codões não permite a reiniciação <sup>(2)</sup>. O tamanho da região intercistronica parece, por vezes, estar envolvido neste processo. Por exemplo, foi descrito que o aumento da sequência desta região melhorou a eficiência da reiniciação e remete-se para que se relacione com o aumento do tempo necessário para readquirir eIF2-GTP-Met-tRNA<sub>i</sub> <sup>(9)</sup>. As estruturas secundárias fortes, para além de outros elementos, quando próximas do codão de terminação da uORF, podem impedir que ocorra reiniciação por prevenirem o recomeço do movimento do ribossoma <sup>(9)</sup>. Desta forma uma reiniciação com baixa eficiência reduz a produção de proteína codificada pela ORF principal <sup>(26)</sup>.

Foram descritos alguns casos em que o próprio péptido codificado pela uORF pode inibir a capacidade de reiniciação do ribossoma <sup>(29)</sup>. Neste caso, considera-se que o péptido poderá interagir especificamente com algum componente da maquinaria traducional, como um factor de terminação, e provocar uma paragem do ribossoma nas proximidades do codão de terminação, bloqueando os restantes ribossomas que iniciam a tradução desse transcrito <sup>(17, 30)</sup>. Poucos exemplos da dependência do péptido têm sido identificados e caracterizados e mesmo entre estes não foi encontrada qualquer sequência conservada ou motivos de consenso. Adicionalmente, os casos conhecidos em que este mecanismo ocorre parecem ter formas de actuar diferentes podendo indicar interacções com diferentes elementos da maquinaria traducional. Estudos em *Arabidopsis* e *Oryza sativa* (arroz) apontam para a conservação de péptidos de uORFs desde o seu ancestral comum, à 140-200 milhões de anos

atrás, tornando interessante considerar a hipótese do uso deste mecanismo pela célula para uma resposta eficiente a mudanças fisiológicas ou condições ambientais <sup>(31)</sup>.

De forma a introduzir ainda maior complexidade e variabilidade, e consistente com o facto da existência da uORF não ter um efeito generalizado na redução dos níveis de proteína principal produzida, estudos recentes têm estabelecido uma interacção entre a região 3' UTR e a uORF do mesmo transcrito. Como resultado, ocorre uma diminuição do efeito inibitório exercido pela uORF. Pensa-se que a 3'UTR promova a alteração de eventos específicos, durante ou após a tradução da uORF, que resultam num aumento da capacidade do ribossoma em reiniciar ou num melhor reconhecimento do AUG da ORF principal <sup>(32)</sup>.

### **I.3. Eritropoietina**

A eritropoietina (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 30,4kDa <sup>(33)</sup>, codificados por um transcrito de 1340 nucleótidos (NM\_000799), cujo gene está localizado no braço longo do cromossoma 7 (q11-q22). Em 1986, Shoemaker *et.al* fizeram uma comparação entre o gene humano e o gene murino codificantes para a EPO e verificaram que ambos têm uma região codificante bastante conservada, sendo constituída por 5 exões. A homologia destes genes estende-se às regiões não codificantes, principalmente as regiões 5' líder, 3' UTR e uma sequência no primeiro intrão. Apesar desta observação expor a possibilidade de alguma função específica destas sequências, tal não foi analisada no referido estudo. Estes autores verificaram ainda que a região 5' líder do transcrito é extremamente rica no seu conteúdo em GC em ambos os genes, tendo o gene humano um teor significativamente superior. Adicionalmente, esta região apresenta 181 nucleótidos no gene humano e uma uORF natural de 14 codões, cuja comparação com o gene murino permitiu identificar a mesma uORF apesar de só conservar 7 dos aminoácidos <sup>(34)</sup>. Até ao presente, nenhum estudo foi efectuado de forma a elucidar o papel biológico desta uORF.

O principal órgão de produção de EPO é o rim em adultos, mais precisamente as células intersticiais do tipo I semelhantes a fibroblastos do córtex renal e da medula exterior, e o fígado na vida fetal <sup>(33)</sup>. No entanto, têm sido identificados outros órgãos de produção e secreção secundária de EPO, incluindo o fígado, cérebro, útero e testículos <sup>(35, 36)</sup>. De facto, estudos preliminares indicam que cerca de 10% da EPO em circulação corresponde à sua secreção por outros órgãos para além do rim <sup>(36)</sup>.



Apesar da EPO ser predominantemente produzida nos rins, os seus receptores são expressos em diferentes tipos celulares. Existem dois tipos de receptores, homodímeros ou heterodímeros. Os homodímeros medeiam respostas hematopoiéticas e os heterodímeros medeiam a protecção de tecidos, principalmente quando o receptor da EPO está associado a uma subunidade comum do receptor  $\beta$ , também conhecido como CD131 <sup>(33)</sup>. Desta forma, várias funções não hematopoiéticas da EPO têm sido identificadas, principalmente a protecção de tecidos e prevenção da apoptose, tornando as suas actividades cardioprotectoras e neuroprotectoras o foco de intensa pesquisa para os seus potenciais terapêuticos <sup>(33)</sup>.

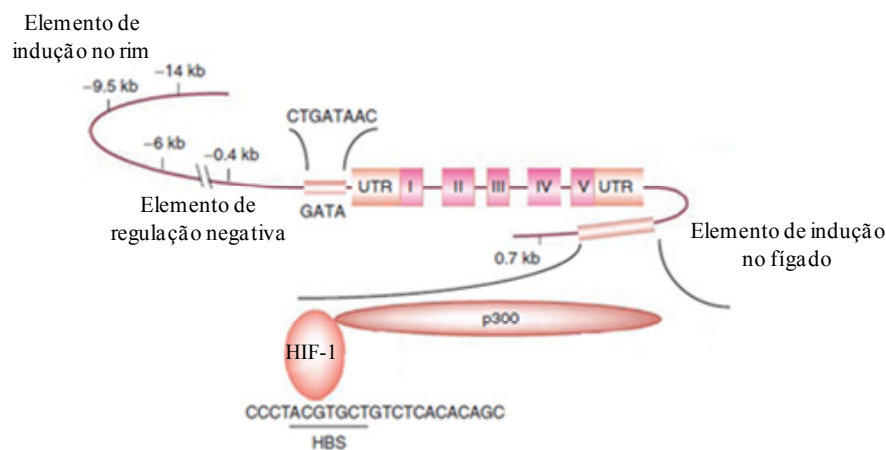
A forma recombinante da EPO tem sido amplamente utilizada para o tratamento de vários tipos de anemias, na sua maioria causadas por hemoglobinopatias hereditárias, falha renal crónica ou quimioterapia <sup>(37)</sup>. No entanto, provou-se que a sua administração pode ser tanto inconveniente como cara <sup>(37)</sup>, principalmente devido a efeitos secundários relacionados com as suas actividades não hematopoiéticas. De facto, as suas capacidades pró-coagulantes, pró-trombóticas, vasoactivas <sup>(33)</sup> e pró-angiogénicas, podem resultar na progressão de tumores e em problemas cardiovasculares e cerebrovasculares, devido à possibilidade da estimulação da actividade da medula óssea <sup>(38,39)</sup>.

O principal estímulo fisiológico para a produção de EPO é a hipoxia, podendo aumentar até 1000 vezes a sua produção <sup>(40)</sup>. A investigação da regulação da expressão do gene da EPO desvendou um controlo a nível transcricional mediado pelo factor HIF-1 (*hipoxia-inducible factor 1*), que contém um local de ligação na região a 3' do gene da EPO <sup>(35)</sup>. Este factor de transcrição em situações de normoxia é degradado por proteossomas visto que se encontra hidroxilado. Todavia o mesmo não ocorre em condições de hipoxia o que vai permitir que este se ligue ao seu local de ligação HBS (*HIF- $\alpha$ -binding site*) (Fig. 3), e que o co-activador transcricional p300 seja recrutado e assim, os níveis de transcrito aumentem e consequentemente os níveis de proteína <sup>(36)</sup>.

Diferentes estudos foram realizados com o objectivo de identificar quais os componentes que controlam a expressão da EPO específica de tecido e quais os mecanismos que medeiam a transição da sua expressão principalmente no fígado, na vida fetal, para o rim, na vida adulta. Estes verificaram que existem diversos elementos de indução de expressão em diferentes tecidos da EPO ao longo do seu gene. Especificamente, existe um local a jusante do gene da EPO que induz a expressão no fígado, onde está localizado o HBS, um outro, de indução da expressão no rim a montante do gene, cerca de -9,5 kb a -14 kb, e um de regulação negativa em locais ectópicos entre -0,4 kb e -6 kb (Fig.3). São ainda desconhecidos os

factores que interactuam com estas regiões, tanto em *cis* como em *trans*, porém tal informação pode ter implicações terapêuticas importantes<sup>(36)</sup>.

Adicionalmente a estes estudos que se baseiam numa regulação ao nível transcricional, foram realizados estudos sobre o comportamento dos níveis de mRNA e de proteína na linha celular Hep3B após a indução de hipoxia, que revelaram a necessidade de uma regulação ao nível pós-transcricional do gene da EPO. Nesses estudos esta regulação foi relacionada com a proteína ERBP (*EPO mRNA-binding protein*) que se liga especificamente na região 3' UTR do transcrito da EPO numa zona de 120 nucleótidos rica em pirimidinas. É sugerido que a ERBP só seja produzida em condições de hipoxia, podendo ligar-se ao mRNA da EPO e prolongar a sua estabilidade<sup>(41)</sup>.



**Figura 3 – Representação esquemática do gene da EPO.** Encontram-se realçados diversos locais regulatórios, incluindo elementos de indução no rim e fígado, um elemento de regulação negativa e a sequência reconhecida pelo HIF-1 (HBS). A p300 é um activador co-transcricional recrutado aquando da ligação do HIF-1 permitindo o aumento da transcrição do gene da EPO. Adaptado de Weidemann *et al.* (2009)<sup>(36)</sup>.

A hipoxia é o estímulo externo mais conhecido e estudado que provoca um aumento na produção dos níveis de EPO. No entanto, outros estímulos podem produzir o mesmo efeito, tais como a glicopenia (baixo teor de açúcar no sangue), infecção, trauma e intenso metabolismo celular<sup>(42)</sup>, apesar de não estar identificado o mecanismo envolvido.

Todos estes estudos sugerem que a expressão da EPO é finamente regulada por diversos factores e a diferentes níveis. Consequentemente, devido à elevada importância a nível biológico e a nível terapêutico o estudo da sua regulação génica revela-se necessário e promissor. Assim, o objectivo do presente trabalho é investigar se a expressão de EPO é também regulada pela ocorrência natural da sua uORF em condições fisiológicas e sob determinados estímulos externos, nomeadamente a falta de nutrientes.

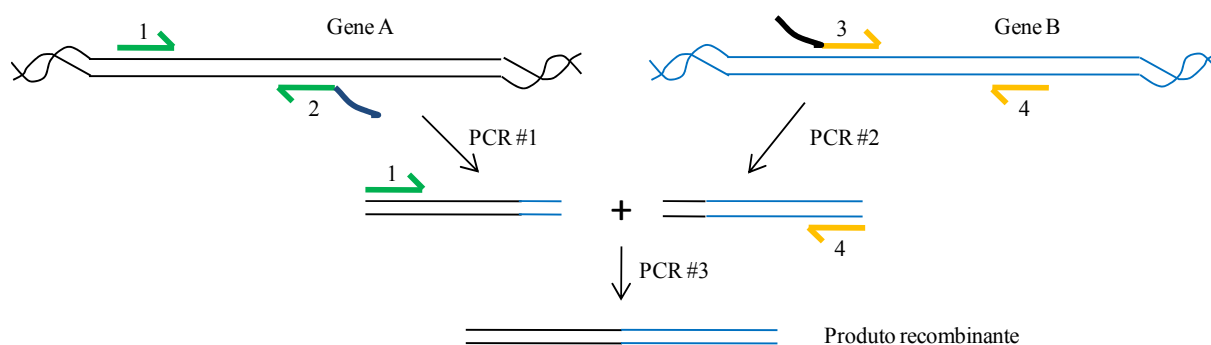
## II. Metodologia experimental

### II.1 Construção de vectores de expressão

Dois vectores de expressão foram utilizados de forma a proceder-se aos ensaios de luminescência. Os vectores escolhidos foram o pGL2-*enhancer* (Promega), que codifica para a luciferase do pirilampo (*Firefly*), e o pGL4.70 (Promega), que codifica para a luciferase da *Renilla*. De forma a ocorrer a sua expressão em células animais, o promotor do citomegalovírus humano (hCMV), presente no vector pcDNA3/hygro+ (*Invitrogen*) (Anexo A – Fig.A1), foi clonado a montante dos genes de luciferase. Os vectores referidos foram digeridos com as enzimas de restrição *Bgl* II (*Amersham*) e *Hind* III (*Promega*) de forma a isolar o segmento de DNA com o promotor do hCMV. Este fragmento e os vectores pGL2-*enhancer* e pGL4.70 foram purificados e ligados utilizando a enzima ligase de DNA do fago T4 (*Fermentas*), durante aproximadamente 16h a 22°C. Parte do produto da ligação foi utilizado para transformar bactérias quimicamente competentes DH5a (*Invitrogen*), que seguidamente foram plaqueadas em meio LB-agar suplementado com o antibiótico de selecção positiva ampicilina (Amp), a uma concentração final de 50µg/mL. As placas foram colocadas numa estufa a 37°C e cerca de 16h depois foi removida uma colónia, inoculada em 4 mL de meio LB-líquido suplementado com Amp, na mesma concentração, e incubada a 37°C, com uma agitação de 220rpm. Dezasseis horas depois, a cultura foi utilizada para extracção de DNA plasmídico, utilizando um *kit* de extracção de DNA plasmídico (*JETquick Plasmid MiniPrep SpinII. Kit/250; Genomed*). As construções foram posteriormente confirmadas por sequenciação automática e foram designadas por pGL2+hCMV (Anexo A – Fig. A2) e pGL4.70+hCMV (Anexo A – Fig. A3).

De modo a proceder aos ensaios para a medição da actividade da luciferase, sob a influência da uORF da EPO, foi necessário clonar toda a região 5' líder do gene da EPO a montante da sequência de cDNA da luciferase do pirilampo (Anexo A – Fig.A2). Para isso, criou-se uma fusão em que o AUG da ORF da luciferase se situa exactamente no mesmo local do AUG da ORF da EPO. Para a formação deste produto recombinante recorreu-se ao processo de *SOEing*<sup>(43)</sup>. Este baseia-se na realização de duas PCR's independentes realizadas com oligonucleótidos específicos e DNAs alvo diferentes (Fig. 4 – PCR #1 e PCR #2), cujos produtos são amplificados numa terceira PCR apenas com os oligonucleótidos iniciadores externos utilizados nas duas primeiras amplificações (Fig. 4 – PCR #3). Sendo assim, de forma a clonar a região 5' líder da EPO a montante da ORF da luciferase, o gene da EPO corresponde ao gene A e o vector contendo o cDNA da luciferase corresponde ao gene B. Um

dos oligonucleótidos iniciadores externos utilizado (Fig. 4 – oligonucleótido 1) é complementar do início da sequência da região 5' do transcrito da EPO, contendo um *linker* para o local de reconhecimento da enzima de restrição *Hind* III, e o outro (Fig.4 – oligonucleótido 4) é complementar a uma região do gene da luciferase a jusante do local de reconhecimento da enzima *Xba* I (Anexo A – Tabela I). Os oligonucleótidos internos contêm *linkers* correspondentes à região que se pretende fundir, ou seja, o correspondente ao oligonucleótido 3, (Fig.4) contém uma sequência complementar ao fim da região 5' líder do gene da EPO, imediatamente antes do AUG, e o correspondente ao 2, (Fig.4) tem a sequência do início do gene da luciferase mesmo após o AUG (Anexo A – Tabela I).

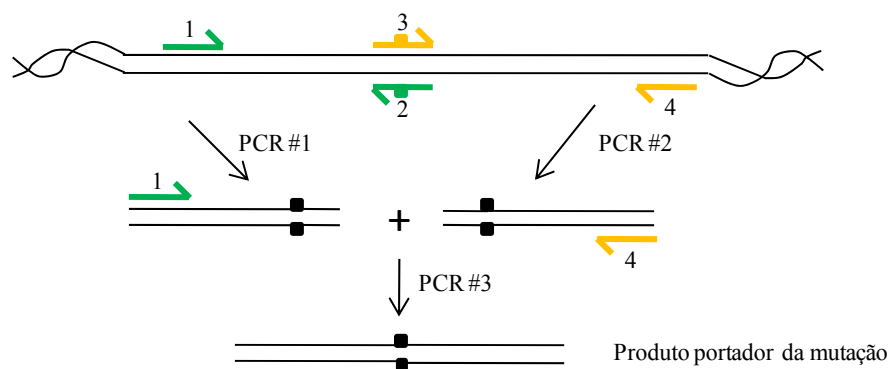


**Figura 4 – Processo de SOEing de forma a obter um produto recombinante.** O DNA alvo que se pretende recombinar está representado com cores diferentes. Os oligonucleótidos utilizados para a PCR #1 são os representados a verde, denominados 1 e 2, por sua vez, para a PCR #2 são utilizados os de cor amarela, denominados por 3 e 4. Os oligonucleótidos 2 e 3 apresentam *linkers* correspondentes a sequências do gene ao qual pretende realizar a fusão, o que os torna complementares. A PCR #3 é realizada utilizando os produtos das PCR referidos anteriormente e os oligonucleótidos 1 e 4. O produto obtido será a fusão dos dois genes alvo. Para realizar a fusão da região 5' líder da EPO com a ORF da luciferase, usou-se como o gene A o DNA da EPO e como gene B o vector com o cDNA da luciferase. Desta forma, o *linker* do oligonucleótido 2 contém a região complementar ao gene da luciferase imediatamente a jusante do ATG, e o *linker* do oligonucleótido 3 é complementar à sequência final da região 5' líder da EPO, imediatamente a montante do ATG, com este incluído. Adaptado de Horton, R (1995) <sup>(43)</sup>.

A amplificação por PCR foi realizada utilizando a enzima *Pfu* recombinante (*Fermentas*) e o programa consistiu numa desnaturação inicial a 95°C durante 2min, seguida de 30 ciclos de: desnaturação a 95°C durante 30seg, hibridação a 57°C durante 30seg e extensão a 72°C durante 30seg; e uma extensão adicional a 72°C durante 2min. O DNA alvo utilizado foi para a primeira reacção DNA genómico humano, e para a segunda o vector pGL2-*enhancer*. Após a terceira reacção o produto de PCR foi purificado utilizando um *kit* de purificação de produtos de PCR (*JETquick PCR product Purification Spin Kit/250; Genomed*)

e, procedeu-se à digestão do produto de PCR purificado e do vector pGL2+hCMV com as enzimas *Hind* III e *Xba* I (*Promega*). Os produtos da digestão foram ligados utilizando a ligase de DNA do fago T4 (*Fermentas*) e usados para transformar bactérias quimicamente competentes DH5 $\alpha$ , do mesmo modo já referido. Realizou-se a extracção de DNA plasmídico e as sequências foram confirmadas por sequenciação automática. A construção foi identificada como pGL2+uORF.

O passo seguinte foi utilizar esta construção como DNA molde para fazer diversas mutações tanto na uORF como na ORF da luciferase. Para este fim, foi utilizado o processo de *SOEing*, como anteriormente. No entanto, apesar dos oligonucleótidos externos (Fig.5 – oligonucleótidos 1 e 4) serem os mesmos, os oligonucleótidos internos apresentam agora a mutação que se deseja introduzir na construção (Fig. 5 – oligonucleótidos 2 e 3). Seguidamente, o produto do PCR#3 e o plasmídeo pGL2+hCMV foram digeridos com as enzimas *Hind* III e *Xba* I e o restante procedimento ocorreu como anteriormente descrito. Todas as construções foram confirmadas por sequenciação automática.



**Figura 5 – Processo de *SOEing* com o objectivo de obter um produto portador de uma mutação.** Na PCR #1 o DNA alvo é amplificado pelos oligonucleótidos representados a verde, denominados 1 e 2, na PCR #2 são utilizados os de cor amarela, 3 e 4. Os oligonucleótidos 2 e 3 são complementares e apresentam a mutação que se pretende clonar. A PCR #3 é realizada utilizando os produtos dos PCRs referidos anteriormente e os oligonucleótidos 1 e 4. O produto obtido será portador da mutação. Adaptado de Horton, R (1995) <sup>(43)</sup>. Especificamente, o DNA alvo foi a construção pGL2+uORF e os oligonucleótidos iniciadores 1 e 4 foram os mesmos do processo de *SOEing* já referido.

As construções realizadas foram (Fig.6):

- pGL2+mut uAUG: portadora de uma mutação no uATG (ATG  $\rightarrow$  TTG) (Anexo A – Tabela I).

- pGL2+mut uTer: portadora da deleção de um nucleótido no codão de terminação da uORF (TGA → -GA) de modo a que as duas grelhas de leitura (da uORF da EPO e da ORF da luciferase) se encontrem em fase (Anexo A – Tabela I).
- pGL2+mut uTer+mAUG: portadora de uma mutação no ATG da luciferase (ATG → TTG), em *cis* à deleção referida (TGA → -GA) no codão de terminação da uORF. Para a obtenção desta construção usou-se a construção pGL2+mut uTer como DNA alvo (Anexo A – Tabela I).

## II.2 Culturas celulares

A linha celular HEK293, cedida pelo grupo de Oncobiologia do Departamento de Genética do Instituto Nacional de Saúde Dr. Ricardo Jorge, foi mantida em frascos T75 com 20mL de meio DMEM + Glutamax-I (*Gibco*), suplementado com 10% de soro fetal bovino (FBS) e numa estufa a 37°C com uma atmosfera de 4% de CO<sub>2</sub>. A passagem das células foi efectuada a cada 3/4 dias para diluições que variaram desde 1:4 até 1:8. No entanto, no dia anterior à passagem foi necessária a mudança de todo o meio de cultura, devido ao esgotamento dos nutrientes, pela mesma quantidade de meio DMEM suplementado com 10% de FBS. Visto que são células aderentes, a passagem destas foi efectuada através da remoção do meio de cultura e adição de 1mL de tripsina durante aproximadamente 2min a 37°C, de modo a quebrar as ligações destas com a superfície do frasco. As células foram ressuspensas em meio DMEM e a fracção desejada mudada para outro frasco T75, previamente preparado com 20mL de meio suplementado com 10% de FBS.

A linha celular HepG2, cedida pelo grupo de Hemoglobinopatias do Departamento de Genética do Instituto Nacional de Saúde Dr. Ricardo Jorge, foi mantida em frascos de cultura celular T75 (*Nunc*) com 20mL de meio RPMI 1640 + Glutamax-I (*Gibco*), suplementado com 10% de FBS e numa estufa a 37°C com uma atmosfera de 4% de CO<sub>2</sub>. A passagem das células foi efectuada a cada 3/4 dias para diluições desde 1:3 a 1:5. Neste caso, foram utilizados 0,5mL de tripsina a 37°C, durante aproximadamente 5min.

Com vista a manter um *stock* destas linhas celulares procedeu-se ao seu congelamento. Para tal, as células foram incubadas em frascos T75 até atingirem uma confluência de aproximadamente 90%. Depois de tripsinizadas, as células foram ressuspensas em 5mL do respectivo meio de cultura e centrifugadas durante 5min a 1200rpm. O meio de cultura foi removido e adicionou-se ao sedimento celular 2mL de solução de congelamento (90% de FBS; 10% de dimetilsulfóxido). As células foram suavemente ressuspensas nesta solução e divididas por dois tubos de congelamento. Procedeu-se por fim a um congelamento gradual

das células e ao seu armazenamento em azoto líquido. Sempre que necessário, procedeu-se ao descongelamento rápido destas alíquotas na estufa a 37°C, e à preparação cuidadosa de culturas em frascos contendo meio suplementado com 20% de FBS.

Neste trabalho, a linha celular HepG2 foi utilizada entre as passagens número 4 e 20, enquanto a linha HEK293 entre a número 3 e 25.

### **II.3 Transfecção de células em cultura**

As linhas celulares anteriormente referidas foram semeadas em placas de seis poços (*Nunc*) usando uma quantidade de células de modo a que no dia seguinte se encontrassem a uma confluência de cerca de 70%. Por poço, foram usados 2mL do respectivo meio suplementado com 10% de FBS. As células foram incubadas durante 18h numa estufa a 37°C com atmosfera de CO<sub>2</sub> a 4%, após a qual se procedeu à sua transfecção. Para a linha celular HEK293, realizou-se uma co-transfecção com 20ng do vector de expressão pGL4.70+hCMV e 2µg dos restantes vectores de expressão (pGL2+hCMV, pGL2+uORF, pGL2+mut uAUG, pGL2+mut uTer ou pGL2+mut uTer+mAUG), de forma a ter uma razão entre os dois vectores de 1:100. Para a co-transfecção da linha celular HepG2, foi usada uma razão de 1:200, tendo-se adicionado 8ng do vector pGL4.70+hCMV e 2µg de cada vector de expressão. Para tal, diluíram-se os plasmídeos em 125µL de meio Opti-MEM (*Gibco*), misturando suavemente. Em paralelo, procedeu-se à diluição de 5µL de lipofectamina<sup>TM</sup> 2000 (*Invitrogen*) em 125µL de Opti-MEM, seguida de incubação durante 5min à temperatura ambiente. Após este período, a cada preparação dos vectores de expressão foi adicionada a diluição de lipofectamina<sup>TM</sup> 2000, tendo o cuidado de misturar suavemente e posteriormente procedeu-se à sua incubação, durante 20min à temperatura ambiente. Por fim, cada uma destas misturas foi distribuída por cada poço da placa de cultura, cujo meio tinha sido previamente substituído. A placa foi de novo colocada na estufa a 37°C.

### **II.4 Transfecção de células em cultura com privação de soro**

As linhas celulares HepG2 e HEK293 foram semeadas com o respectivo meio suplementado com 10% de FBS, e incubadas 18h a 37°C, como referido anteriormente. No entanto, imediatamente antes do momento da transfecção, este meio foi retirado, as células foram lavadas, duas vezes, com 1mL do respectivo meio sem soro e foi-lhes adicionado 2mL de meio, sem qualquer suplemento de FBS. Procedeu-se então à transfecção como referido anteriormente e incubação na estufa a 37°C.

No caso da linha celular HEK293, o crescimento celular sem FBS ocorreu por 24h, sendo notório o efeito da falta do soro nas células. No caso da linha celular HepG2, foram feitas incubações de 24, 48 e 72h sem alteração morfológica das células, tendo-se escolhido as 72h como o tempo indicado para os estudos posteriores.

### II.5 Ensaio de luminescência

A actividade da luciferase foi determinada usando o *Dual Luciferase Reporter Assay System (Promega)* e o luminómetro *Lucy 2 (Anthos)*. Após dezasseis horas de transfecção para a linha celular HEK293, e vinte e quatro horas para a linha celular HepG2, procedeu-se à obtenção de extractos proteicos. Para isso, realizou-se uma lise activa das células, usando, por poço, 250µL de *passive lysis buffer* 1x concentrado (PLB, *Promega*), no caso da linha celular HEK293 e 100µL no caso das células de HepG2. De seguida, os lisados foram rapidamente congelados a -80°C, descongelados a 37°C, e centrifugados a 4°C durante 30seg. A posterior análise da actividade da luciferase foi efectuada no luminómetro de acordo com as instruções do manual da *Promega*.

O ensaio de luminescência consiste na medição sequencial da actividade da luciferase do pirilampo e da *Renilla* na mesma reacção. Tal é possível pois as duas enzimas têm estruturas e requisição de substratos diferentes. Assim, depois da adição do substrato e medição da actividade da luciferase do pirilampo, esta reacção é interrompida ao mesmo tempo que é estimulada a actividade da luciferase da *Renilla*, permitindo, passado alguns segundos, a sua medição. A luciferase da *Renilla* funciona como o controlo interno da reacção de forma a normalizar a quantidade de luciferase de pirilampo medida.

A proteína produzida pelas células que cresceram na ausência de soro, foi quantificada pelo mesmo procedimento, com a diferença de que a lise ocorreu após 24h de incubação (após transfecção) para a linha celular HEK293 e 72h para HepG2.

### II.6 RT-PCR quantitativo

As linhas celulares HEK293 e HepG2, transfectadas como anteriormente referido, foram utilizadas para a extracção do correspondente RNA total. Para tal procedeu-se à sua lise com o tampão RA1 fornecido pelo *kit* de extracção de RNA *NucleoSpin RNA II (Macherey-Nagel)*. Este *kit* foi de seguida utilizado na extracção do RNA das diferentes amostras, de acordo com o manual fornecido pelo fabricante. Em seguida, as cadeias de cDNA foram sintetizadas a partir do RNA isolado. Com este objectivo, foram utilizadas esferas comerciais (*Ready-To-Go You-Prime First-Strand Beads, GE Healthcare*) contendo transcriptase reversa,



inibidores de RNases e nucleótidos. O procedimento foi efectuado de acordo com as especificações do manual, juntando-se à mistura referida, 2ng de RNA extraído e um oligonucleótido iniciador aleatório (*random primer*, *Invitrogen*).

A reacção de PCR quantitativo foi realizada utilizando oligonucleótidos iniciadores desenhados especificamente para o gene da luciferase do pirilampo e para o gene da luciferase da *Renilla* (Anexo A – Tabela II) obtendo-se, em ambos os casos, fragmentos de 51bp. A estes adicionou-se uma mistura de reacção (*SYBR green PCR master mix*, *Applied Biosystems*), que contém a polimerase *AmpliTaq Gold*, o intercalante de DNA em dupla cadeia *SYBR green*, e nucleótidos, incluindo dUTP. O programa consistiu num passo inicial de desnaturação a 95°C durante 10min, seguido de quarenta ciclos que incluem um passo de desnaturação a 95°C por 15seg e um passo de extensão a 60°C por 1min. A reacção foi quantificada utilizando o aparelho de PCR em tempo real *ABI PRISM 7000* (*Applied Biosystems*). Como DNA alvo foi utilizado o cDNA de cada amostra obtido como anteriormente descrito. Para que a sua quantidade se encontrasse em quantidades detectáveis, várias diluições tiveram que ser realizadas. No caso do cDNA correspondente à lise das células HepG2 foi necessária uma diluição inicial de todas as amostras de 1:10. Adicionalmente, a amostra correspondente à transfecção com pGL2+hCMV foi utilizada para fazer a recta de calibração e foi diluída de uma forma seriada de 1:5 até se obterem 4 pontos da recta. O mesmo procedimento foi efectuado com os lisados obtidos pela transfecção das células HEK293, no entanto, estas requereram uma diluição inicial de 1:100 em todas as amostras.

### **II.7 Imuno-deteccção por *Western Blot***

As células HEK293 foram semeadas numa placa de seis poços e transfectadas de modo idêntico ao anteriormente descrito. Vinte e quatro horas após a transfecção, as células foram lisadas como referido anteriormente e a 60µL do lisado celular foi adicionado 20µL de SDS 5x. De seguida, procedeu-se à sua desnaturação, a 95°C, durante 10min, e armazenamento a -20°C. Posteriormente, procedeu-se à preparação de um gel de SDS-PAGE (gel de resolução a 10% e gel de concentração a 4%; Anexo B) onde foram aplicadas as amostras dos lisados. A corrida do gel fez-se a uma amperagem fixa de 20mA durante aproximadamente 3h. Durante a electroforese, uma membrana de nitrocelulose foi hidratada em metanol e posteriormente lavada em H<sub>2</sub>O bidestilada. A transferência das proteínas para a membrana foi efectuada a uma voltagem fixa de 100V durante 1h. Lavou-se a membrana com TBST 1x e bloqueou-se com uma solução de leite em pó em TBST durante 1h. Removeu-se o

excesso de leite com TBST e efectuou-se a imunomarcação. Visto que a proteína da luciferase tem aproximadamente 61 kDa e a que  $\beta$ -catenina, utilizada como controlo interno da reacção, tem cerca de 92 kDa, a membrana foi cortada na região dos 75 kDa. Metade desta foi incubada com o anticorpo primário de rato anti-Luciferase (*Invitrogen*), a uma diluição de 1:200, e a outra metade com o anticorpo primário de rato anti- $\beta$  catenina numa diluição de 1:3000. Após 1h, lavou-se várias vezes as membranas em TBST, num total de 20min de lavagem e procedeu-se à incubação com o anticorpo secundário de cabra anti-IgG de rato (*Abcam*) a uma diluição de 1:4000, durante 1h. Este último anticorpo está conjugado com HPR (*horseradish-peroxidase*). As membranas foram novamente lavadas com TBST durante 30min. Finalmente, foi realizada uma reacção do tipo ECL (*Enhanced chemiluminescence*) durante 1min. Esta reacção consiste na exposição das membranas a uma mistura de duas soluções (solução 1 e 2, Anexo B) constituídas por luminol,  $H_2O_2$  e ácido cumárico. Finalmente, foram feitas exposições em chapas radiográficas a vários tempos. A revelação foi efectuada no processador Kodak X-OMAT 1000.

### III. Resultados

#### III.1 A uORF da EPO inibe a eficiência de tradução da ORF principal

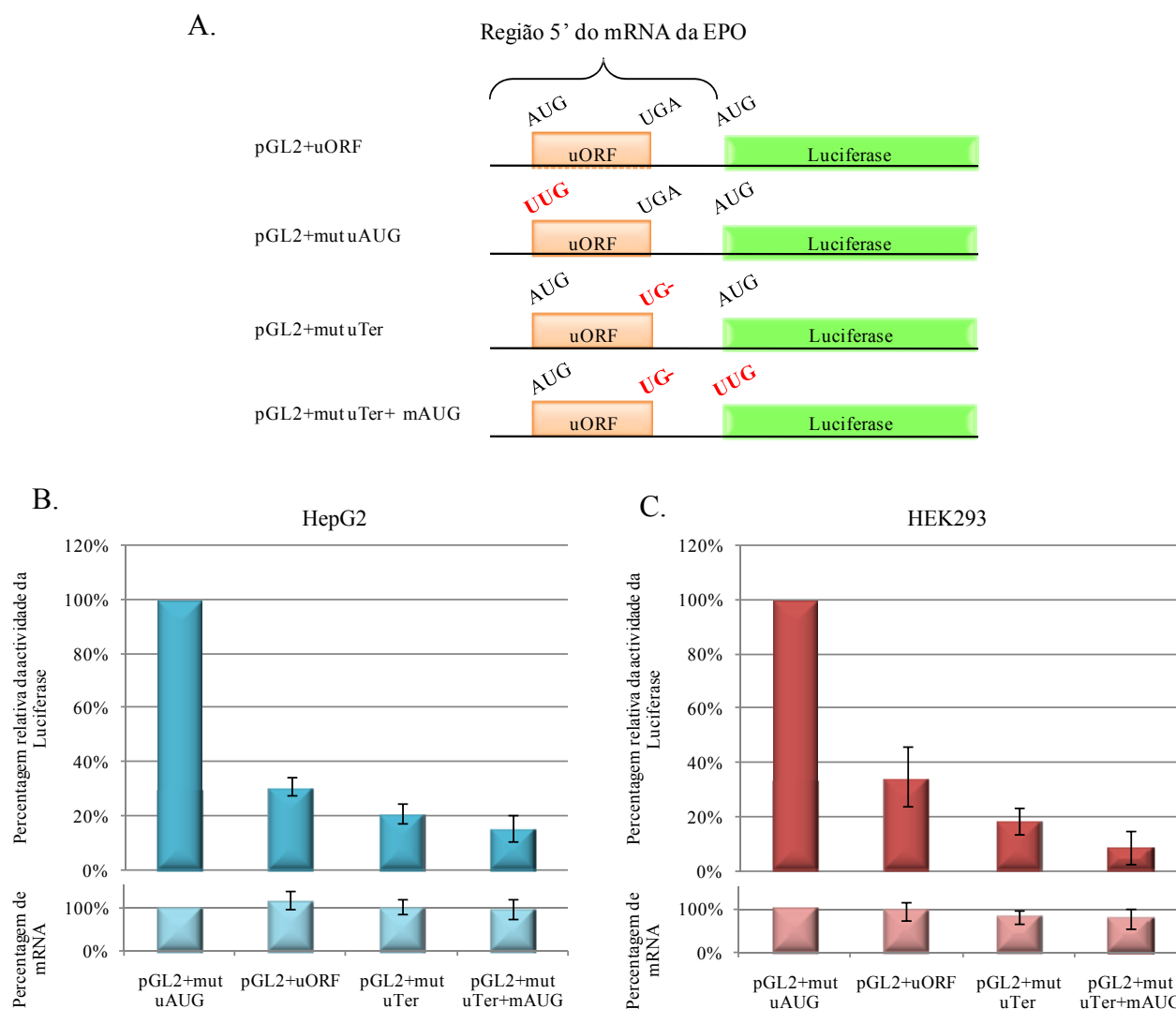
O transcrito da EPO contém um codão de iniciação, na sua região 5' líder, que precede o codão de iniciação da ORF principal. Este codão AUG é seguido por uma pequena ORF de 14 codões e de uma região intercistronica de 22 nucleótidos, no fim da qual está presente o codão AUG da ORF da EPO. De modo a avaliar a função desta uORF na expressão da EPO, foram realizadas construções repórter, clonando a sequência 5' líder deste transcrito a montante do gene da luciferase (pGL2+uORF). Adicionalmente foram realizadas mutações no uAUG (pGL2+mut uAUG), no codão de terminação da uORF (pGL2+mut uTer) e neste em simultâneo com uma mutação no AUG da ORF da luciferase (pGL2+mut uTer+mAUG) (Fig.6A). Estas construções foram utilizadas para transfectar duas linhas celulares diferentes, HepG2 e HEK293, de modo a permitir a distinção de mecanismos específicos de tecido.

Visto que o objectivo é estudar os efeitos ocorridos ao nível da tradução, a actividade da luciferase de cada construção foi normalizada aos correspondentes níveis de mRNA, determinados por RT-PCR em tempo real.

A construção em que o uAUG está mutado corresponde à ausência da possível inibição exercida pela uORF, como tal, esta foi utilizada para a normalização dos valores da quantidade de actividade enzimática obtidos pelas restantes construções. Não obstante, a situação considerada fisiológica é a que apresenta a uORF intacta visto ser a que corresponde ao que ocorre naturalmente. Os resultados revelam que a quantidade de proteína detectada quando a uORF se encontra intacta é cerca de 30% em ambas as linhas celulares estudadas ( $p=0,0001$  em HepG2 e  $p=0,0004$  em HEK293). Nas mesmas condições os níveis de mRNA com e sem uORF funcional mostraram-se equivalentes em ambas as linhas celulares ( $p=0,4422$  em HepG2 e  $p=0,3400$  em HEK293) (Fig.6B e C – construção pGL2+uORF). Consequentemente, é possível concluir que a uORF inibe, em condições fisiológicas, cerca de 70% da tradução da ORF a jusante.

Quando o codão de terminação da uORF está alterado, de modo a que o uAUG fica em fase com o AUG da ORF principal, espera-se a obtenção de uma proteína de fusão e, caso ocorra reconhecimento do AUG da luciferase, que ocorra a expressão da proteína nativa. Os resultados obtidos revelam uma diminuição significativa da quantidade de proteína detectada em comparação com a situação fisiológica (pGL2+uORF), em ambas as linhas celulares (em HepG2  $p=0,0052$ , e em HEK293  $p=0,0000$ ). No entanto, os níveis de mRNA mantêm-se constantes (em HepG2  $p=0,3862$ , e em HEK293  $p=0,3259$ ) (Fig.6B e C – construção

pGL2+mut uTer), o que permite concluir que esta diferença poderá estar relacionada com um outro mecanismo de regulação ao nível traducional ou com uma diminuição da actividade enzimática da proteína de fusão.



**Figura 6 – A uORF da EPO afecta a eficiência de expressão da luciferase, mas não do correspondente nível de mRNA.** (A) Representação esquemática das construções realizadas, todas incluindo a região 5' líder do gene da EPO em fusão com o gene repórter da luciferase. As mutações realizadas estão representadas a vermelho. (B) Actividade de luciferase (em cima) e de mRNA da mesma (em baixo) em células HepG2 para cada uma das construções esquematizadas em (A). Os dados representam a razão da actividade enzimática da luciferase do pirilampo em relação à da luciferase da *Renilla* e foram normalizadas relativamente à quantidade de proteína codificada pela construção pGL2+mut uAUG. (C) A mesma análise que em (B) mas obtida na linha celular HEK293.

De modo a distinguir a quantidade de proteína de fusão da quantidade de proteína nativa de luciferase produzidas pela construção pGL2+mut uTer, o codão de terminação da uORF foi alterado em conjugação com uma mutação no codão AUG da ORF da luciferase,

construção pGL2+mut uTer+mAUG. Desta forma só ocorre reconhecimento do uAUG e só é produzida a proteína de fusão. Os resultados obtidos revelam, de novo, uma diminuição da actividade enzimática quantificada em ambas as linhas celulares (em HepG2,  $p=0,0114$ , e em HEK293  $p=0,0024$ ), enquanto os níveis de mRNA se mantêm equivalentes (em HepG2  $p=0,3961$ , e em HEK293  $p=0,1703$ ) (Fig.6 B e C – construção pGL2+uTer+mAUG). Tal como anteriormente, o mecanismo que explica estes resultados não está ainda identificado, podendo remeter para outros processos de síntese da ORF principal ou para a diminuição da actividade do péptido de fusão sintetizado. No entanto, como esta diminuição foi ligeiramente superior à anterior, indicia que na construção pGL2+mut uTer existe a possibilidade de reconhecimento do AUG da ORF principal.

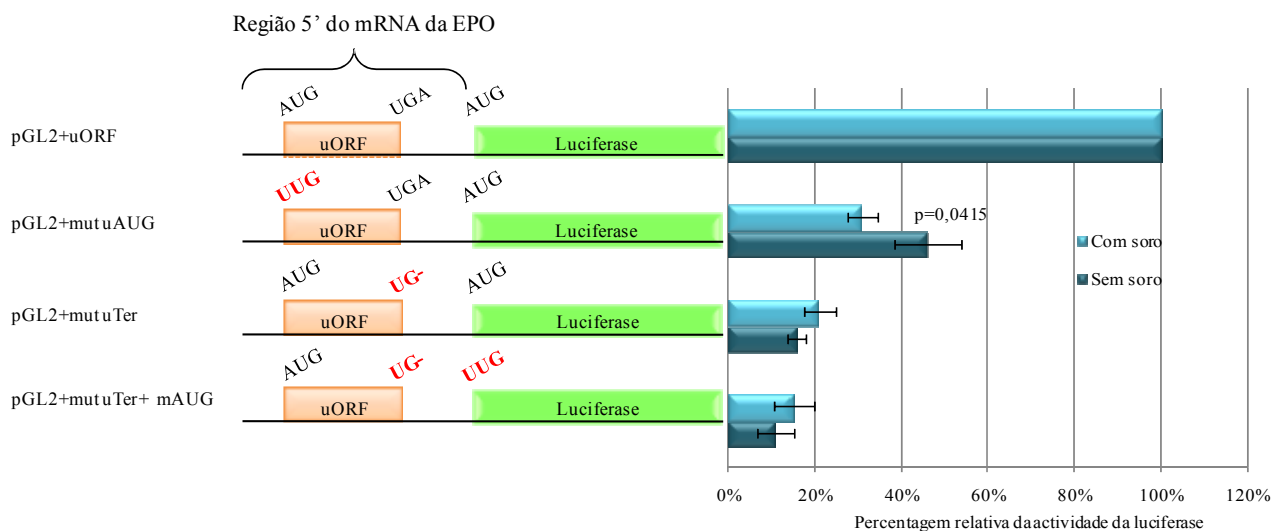
### **III.2 A inibição da expressão da EPO, exercida pela uORF, responde a estímulos externos**

O principal estímulo externo que modula a expressão da EPO é a hipoxia. Esta provoca um aumento da quantidade de EPO produzida, que é explicado por um mecanismo de indução da transcrição mediado pelo HIF<sup>(35)</sup>. No entanto, outros estímulos provocam o mesmo efeito, tais como, glicopenia, infecção, trauma e intenso metabolismo celular<sup>(42)</sup>. De modo a verificar se a uORF tem algum papel neste tipo de resposta, utilizaram-se as mesmas construções e as linhas celulares anteriormente referidas, em situações que mimetizam a falta de nutrientes que pode ocorrer ao nível fisiológico. Ou seja, realizaram-se transfecções das células em cultura com privação de soro e foram posteriormente medidos os níveis de proteína por ensaios de luminescência.

Os resultados obtidos em HepG2 indicam que ocorre uma diminuição da quantidade de proteína produzida quando a uORF se encontra funcional ( $p=0,0007$ ). No entanto, ao comparar este valor em ambas as condições em estudo, observa-se que aquando da privação de soro, a diminuição é significativamente menor ( $p=0,0415$ ). Mais especificamente, observa-se que na presença de soro a quantidade de luciferase detectada é de 31% e na ausência de soro passa a ser de 46% (Fig.7 – construção pGL2+uORF), o que significa que a inibição pela uORF inicialmente é de 69% e passa para 54%. Desta forma, é possível concluir que a inibição exercida pela uORF, sobre a ORF principal, diminui como resposta a um estímulo anómalo.

A transfecção na presença de soro revelou uma diminuição da quantidade de actividade luciferásica detectada na construção em que o codão de terminação da uORF está mutado, tal como na construção portadora desta mutação em conjunto com uma outra no

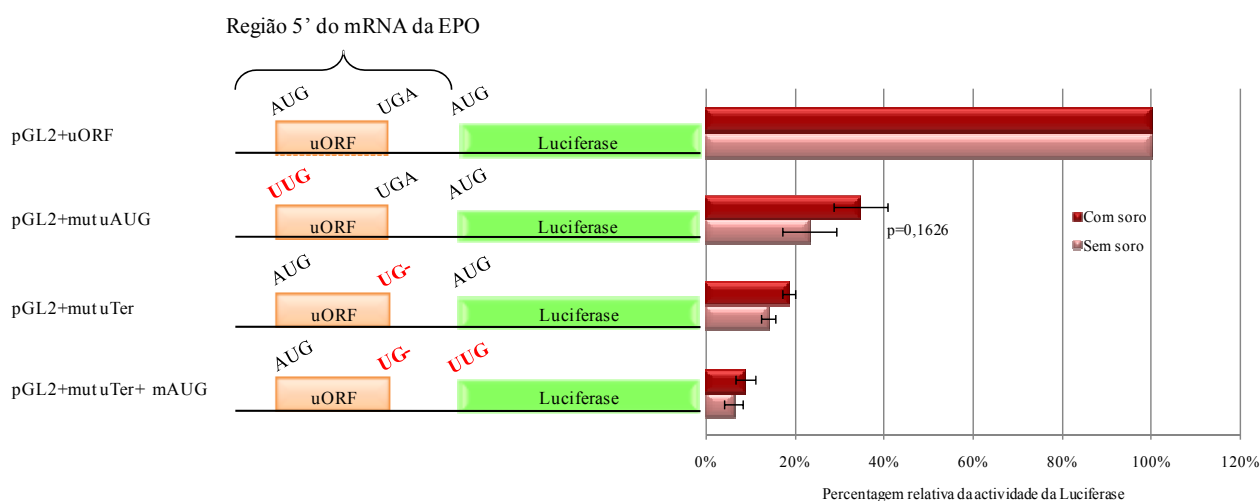
codão AUG da ORF da luciferase. O mesmo foi observado em condições de crescimento celular sem soro ( $p=0,0007$  para pGL2+mut uTer, e  $p=0,0129$  para pGL2+mut uTer+mAUG) (Fig.7 – construções pGL2+mut uTer e pGL2+mut uTer+mAUG). Apesar, de tal como anteriormente, a justificação deste valor ser desconhecida, estes dados parecem indicar uma menor eficiência no reconhecimento do uAUG.



**Figura 7 – A inibição exercida pela uORF diminui como resposta à privação de soro na linha celular HepG2.** À esquerda estão esquematizadas as construções realizadas. À direita estão representados os resultados obtidos em condições de expressão das construções na linha celular HepG2, com e sem soro. É possível observar uma descida significativa da actividade enzimática de luciferase quando a uORF está intacta que corresponde, contudo, a uma diminuição do efeito exercido pela uORF para 54%. A mesma diminuição é observada nas restantes construções, o que parece indicar uma menor eficiência de reconhecimento do uAUG.

O mesmo tipo de análise foi realizado na linha celular HEK293, de forma a verificar se a diminuição da inibição pela uORF sob este tipo de *stress*, é uma resposta específica de tecido. Os resultados obtidos não revelam, ao contrário do que acontece em HepG2, uma diferença estatisticamente significativa quando se compara os resultados da transfecção na presença e ausência de soro ( $p=0,1626$  para pGL2+uORF,  $p=0,3050$  para pGL2+mut uTer e  $p=0,8802$  para pGL2+mut uTer+mAUG) (Fig.8). Ou seja, o padrão da actividade enzimática detectada para cada uma das construções revela-se semelhante, e na mesma ordem de grandeza, ao obtido em condições com soro. Especificamente, observa-se a mesma diminuição significativa dos níveis de proteína como resultado do efeito da uORF e a diminuição da actividade enzimática da proteína codificada pelas restantes construções ( $p=0,0413$  para pGL2+uORF,  $p=0,0424$  para pGL2+mut uTer e  $p=0,0225$  para pGL2+mut uTer+mAUG). Apesar da diferença de proteína detectada quando a uORF está intacta, com e sem soro, não ser estatisticamente significativa, esta indicia a possibilidade da síntese proteica

ser menor. Este dado está de acordo com o que é esperado como resposta a uma situação de carência de nutrientes, i.e., como a disponibilidade de aminoácidos é menor, nestas condições, ocorre uma diminuição generalizada da tradução.



**Figura 8 – A uORF não responde à privação de soro nas células HEK293.** À esquerda estão esquematizadas as construções utilizadas para este estudo. À direita estão representados os resultados obtidos em condições de expressão de construções na linha celular HEK293 com e sem soro. É possível observar que, comparando ambas as condições, não existe uma diferença significativa dos níveis de proteína obtidos.

A nível técnico verificou-se não ser possível efectuar um maior tempo de exposição das células à privação do soro, pelo que os valores apresentados revelam apenas uma tendência da resposta da uORF a estímulos externos anómalos. Espera-se assim, que o estudo de outros tipos de *stress*, como estimulação por hipoxia ou TNF- $\alpha$ , venham corroborar a capacidade de resposta da uORF.

Concluindo, a expressão da EPO responde a estímulos externos, nomeadamente, à falta de nutrientes, por uma diminuição da inibição exercida pela sua uORF. Esta resposta resulta num aumento de expressão da EPO e é específica de tecido, pois as células hepáticas parecem ter capacidade de resposta enquanto as células de rim embrionário não apresentam qualquer variação na eficiência de tradução da ORF principal.

### III.3 A iniciação da tradução no AUG principal da EPO ocorre pelo mecanismo de *leaky scanning*

Segundo o modelo convencional de iniciação da tradução, o ribossoma começa a síntese proteica após o reconhecimento do primeiro AUG que encontra. No entanto, no caso de um transcrito conter uma uORF, este modelo nem sempre se verifica. Assim, existem dois

mecanismos principais que permitem o reconhecimento do codão AUG da ORF principal: o *leaky scanning* e a reiniciação. De modo a avaliar a ocorrência de *leaky scanning*, para que se dê o reconhecimento do AUG da EPO, foram utilizadas as construções, referidas anteriormente, transfectadas na linha celular HEK293 e de seguida detectaram-se os níveis de luciferase por imuno-deteccção por *Western Blot*. A mesma análise da quantidade de proteína de luciferase produzida nas células HepG2 transfectadas não foi possível visto que estas apresentam uma baixa eficiência de transfecção, comparativamente às células HEK293. Assim, a quantidade de proteína não foi o suficiente para ser detectável por esta técnica.

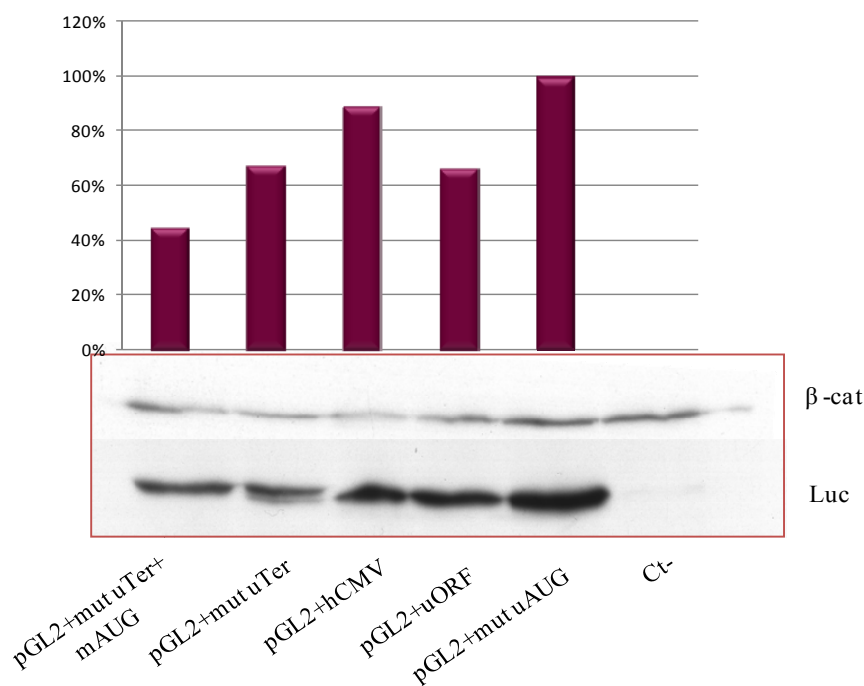
Os resultados obtidos em células HEK293 transfectadas com a construção pGL2+mut uTer, revelam a deteção de duas bandas (Fig.9 – construção pGL2+mut uTer), i.e., parece estar a ocorrer o reconhecimento dos dois codões AUGs presentes no correspondente transcrito, produzindo-se, desta forma, uma proteína de fusão e a proteína normal da luciferase, que parece estar em menor quantidade. A quantificação, em conjunto, destas duas bandas, revelou níveis equitativos aos detectados em células transfectadas com a construção pGL2+uORF (Fig.9 – construção pGL2+mut uTer). Este valor contradiz o decréscimo dos níveis de actividade enzimática obtido pelo ensaio de luminescência (Fig.6C – construção pGL2+mut uTer) em extractos provenientes de células HEK293 transfectadas com a referida construção. Sendo assim, é possível concluir que esta diminuição ocorre devido à menor actividade enzimática de proteína de fusão sintetizada nestas condições, e não como resultado de um processo relacionado com a eficiência de tradução da ORF principal.

A quantidade de proteína obtida e as duas bandas observadas por imuno-deteccção, quando o codão de terminação está mutado, não podem ser explicadas pela ocorrência de reiniciação, já que não ocorre a terminação da tradução da uORF, portanto, conclui-se que o mecanismo que permite a tradução destas duas formas é o *leaky scanning*. Ou seja, a subunidade 43S do ribossoma, pode reconhecer ou não o uAUG. No caso de não o reconhecer, esta subunidade ribossomal continua o *scanning* do transcrito, e a tradução pode ser iniciada no AUG da ORF principal.

Como resultado da mutação do codão de terminação da uORF juntamente com a mutação do codão AUG da luciferase observa-se, como esperado, apenas a produção da proteína de fusão. A quantificação desta banda revelou que ocorre uma diminuição da quantidade de proteína sintetizada nestas condições (Fig.9 – construção pGL2+mut uTer+mAUG). Esta diminuição foi também observada em ensaios de luciferase, porém, numa ordem de grandeza bastante diferente devido à menor actividade enzimática da proteína de fusão (Fig.6C – construção pGL2+mut uTer+mAUG). Este resultado indica que, cerca de



70% das vezes que um ribossoma percorre o transcrito, ocorre o reconhecimento do uAUG da uORF. Em consequência, a subunidade ribossomal 43S que não o reconhece pode continuar a realizar o *scanning* do transcrito até encontrar um AUG mais a jusante, permitindo a tradução da EPO.



**Figura 9 – O processo de *leaky scanning* permite a tradução da ORF principal.** Em cima estão representados os valores obtidos pela quantificação da imuno-deteção. Em baixo está representada a imuno-deteção por *Western Blot* da proteína da luciferase em extractos de HEK293. Para além das construções já referidas estão representados um controlo positivo (pGL2+hCMV), que apresenta a ORF de luciferase intacta com a sequência do promotor do hCMV a montante, e um controlo negativo da experiência (Ct-), correspondente a extractos proteicos de células não transfectadas.

Quando a uORF se encontra funcional, revela-se uma diminuição da quantidade de proteína expressa (Fig.9 – construção pGL2+uORF). No entanto, este valor é maior do que o detectado com o ensaio de luminescência (Fig.6C – construção pGL2+uORF). Pensa-se que tal se deve, primeiro a uma limitação da técnica, visto que a quantificação do sinal saturado, observado na detecção da proteína produzida quando a uORF não está funcional, não é fidedigna. Em segundo lugar, os lisados utilizados não são os mais representativos, visto que, no ensaio de luciferase, a diminuição de proteína detectada como resultado da transfecção com pGL2+uORF foi cerca de 50%.

## Discussão

Os estudos da regulação da expressão génica têm revelado elevada complexidade e diversidade de processos responsáveis por uma correcta 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 volúvel <sup>(1,3)</sup>.

Pequenas uORFs são um exemplo de elementos, que actuam 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 principal, e estão presentes principalmente em proto-oncogenes, e em genes envolvidos no crescimento e diferenciação celular <sup>(16, 17)</sup>. Estes elementos contêm o primeiro codão AUG do transcrito, sendo reconhecidos e traduzidos pelo modelo convencional de iniciação da tradução. Nos transcritos com uORF, a tradução inicia-se no codão AUG da ORF principal tanto por reiniciação como por *leaky scanning*.

O objectivo do presente trabalho foi o estudo de uma uORF natural de 14 codões do mRNA da EPO, descrita em 1986 por Shoemaker *et.al*, e o modo como este elemento afecta a expressão da ORF principal. Para tal, tirou-se partido do sistema de fusão com o gene repórter da luciferase e de diversas construções com mutações pontuais de forma a serem quantificados, por luminescência, os níveis de proteína produzida. Estes resultados foram normalizados aos correspondentes níveis de mRNA, para confirmar que os efeitos são consequência de uma regulação traducional. Assim recorreu-se à técnica de RT-PCR em tempo real. Com o objectivo de identificar a relevância biológica deste elemento foi utilizada a linha celular derivada de um carcinoma hepatocelular (HepG2), visto ser a utilizada nos estudos de regulação da expressão do gene da EPO <sup>(36)</sup>. Como controlo foi utilizada a linha celular derivada de rim embrionário (HEK293). Embora se pense que uma linha celular derivada de tecido de rim fosse a mais indicada para os estudos de expressão génica da EPO, as tentativas para o seu estabelecimento têm-se revelado infrutíferas <sup>(36)</sup>. Os resultados obtidos permitiram concluir que, efectivamente, a uORF da EPO é capaz de inibir a eficiência traducional da ORF principal. Esta inibição revelou ser cerca de 70% e apresentou níveis semelhantes nas linhas celulares estudadas, o que demonstra ser um mecanismo generalizado.

Foi também comprovado que a uORF exerce a sua função apenas se for reconhecida e traduzida.

O principal estímulo para o aumento da produção de EPO é a hipoxia, no entanto, pensa-se que outras situações de *stress*, como a falta de nutrientes e a estimulação por citocinas pro-inflamatórias, possam resultar no mesmo efeito <sup>(43)</sup>. Nestas condições, os níveis de EPO aumentam para que ocorra uma maior produção de eritrócitos que, por sua vez, vão colmatar essas situações de *stress*. Sendo assim, é importante verificar se a presença da uORF pode modular a expressão da ORF principal nestas situações. Consequentemente, foi aplicado como estímulo externo, a privação de soro nas linhas celulares utilizadas neste estudo, mimetizando a falta de nutrientes na situação fisiológica. Foi observado que, nas células HepG2, ocorreu uma expressão diferencial nas condições com e sem soro em todas as construções efectuadas. Mais especificamente, observou-se que na situação sem soro a quantidade de proteína detectada quando a uORF se encontra intacta foi significativamente superior, cerca de 46% (*versus* 31% em células com soro). Tal revela uma repressão da inibição exercida pela uORF que passou de 69% em condições normais para 54% em condições de *stress*. Portanto, é possível concluir que o efeito inibitório da uORF é reprimido como resposta a condições anómalas, provocando um aumento da quantidade de EPO produzida. Adicionalmente, na ausência de soro, verificou-se uma diminuição da quantidade de actividade enzimática detectada quando o codão de terminação da uORF está mutado e quando este ocorre em conjunto com a mutação no AUG da ORF da luciferase, relativamente à transfecção na presença de soro. Apesar de se ter comprovado, por *Western Blot*, que tal diminuição ocorre devido à proteína de fusão sintetizada não ser totalmente funcional, estes dados indicam que poderá estar a ocorrer uma diminuição da eficiência de tradução iniciada no uAUG. Esta diminuição pode estar relacionada com a escassez de nutrientes, que provoca um decréscimo do nível de aminoácidos disponíveis para a tradução, ou com um mecanismo adicional, ainda por elucidar, que permita um aumento da eficiência de tradução da ORF principal.

Quando os mesmos estudos foram efectuados na linha celular HEK293 não se observaram os mesmos resultados, visto que a resposta dada pelas células na ausência de soro foi em tudo semelhante à que ocorre na situação com soro. Tal permite concluir que a regulação traducional da EPO é específica de tecido. No entanto, a base molecular desta regulação está ainda por identificar.

Um dos mecanismos investigados que permite explicar a iniciação da tradução no AUG da ORF da EPO, foi o *leaky scanning*. Este postula que, por vezes, o ribossoma não

reconhece o primeiro AUG que encontra e continua a percorrer o transcrito até encontrar o codão AUG seguinte e assim iniciar a tradução. A sua existência foi comprovada por imunodeteção por *Western Blot* que revelou a presença de duas proteínas quando o codão de terminação da uORF está mutado. Só o reconhecimento dos dois AUGs pode explicar a presença dos dois péptidos. Salienta-se ainda que o facto de não existir terminação da tradução da uORF impossibilita a ocorrência de reiniciação nesta construção.

A eficiência do reconhecimento do codão AUG depende do contexto em que este está inserido. Neste caso o contexto do AUG da uORF é bom mas não óptimo (CCCGGGAUGAG), o que está de acordo com a possibilidade da ocorrência de *leaky scanning*.

A quantidade de proteína detectada, por *Western Blot*, quando o codão de terminação da uORF está mutado, e quando, para além deste, o AUG da luciferase também se encontra mutado, permitiram confirmar que a proteína de fusão obtida apresenta uma diminuição da capacidade de degradação do substrato. Adicionalmente, na situação em que só ocorre produção da proteína de fusão, os resultados revelam que o uAUG é reconhecido cerca de 70% das vezes (em relação à condição fisiológica – pGL2+uORF), o que está de acordo com o nível de expressão de proteína quando a uORF está intacta (30%).

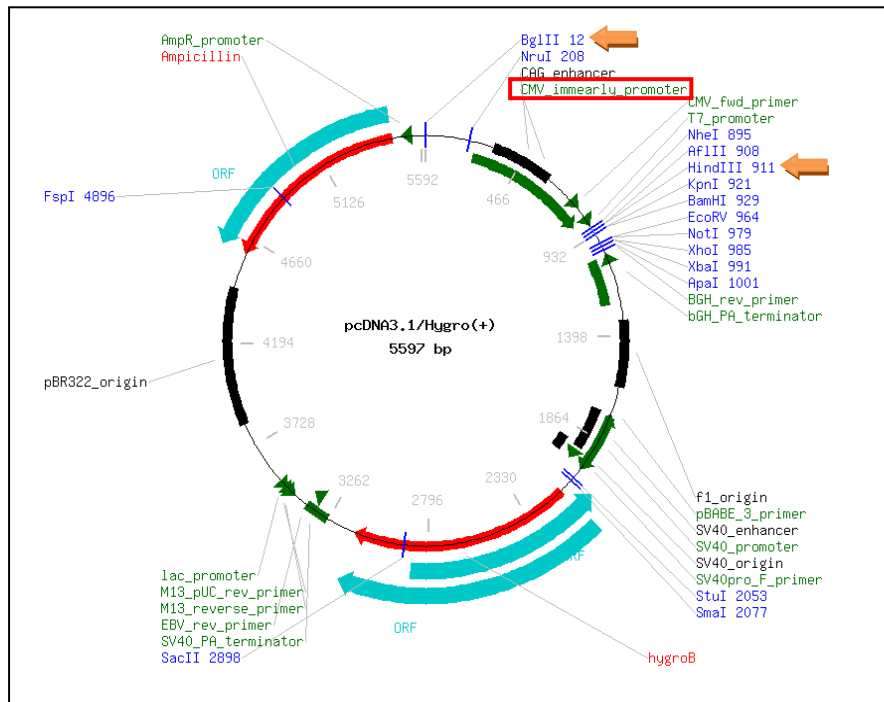
Concluindo, o presente estudo comprova que a uORF natural do mRNA da EPO regula negativamente a tradução da ORF principal, de uma forma generalizada nos diferentes tecidos, e responde a condições de *stress* de uma forma específica de tecido, aumentando a quantidade de proteína produzida. Adicionalmente, verificou-se que o mecanismo de *leaky scanning* está envolvido no reconhecimento do AUG da ORF principal. No entanto, outros pontos permanecem por estudar, nomeadamente a ocorrência de reiniciação e o mecanismo pelo qual o efeito da uORF é inibido em células onde a EPO aumenta de expressão em resposta a um estímulo externo.

## Bibliografia

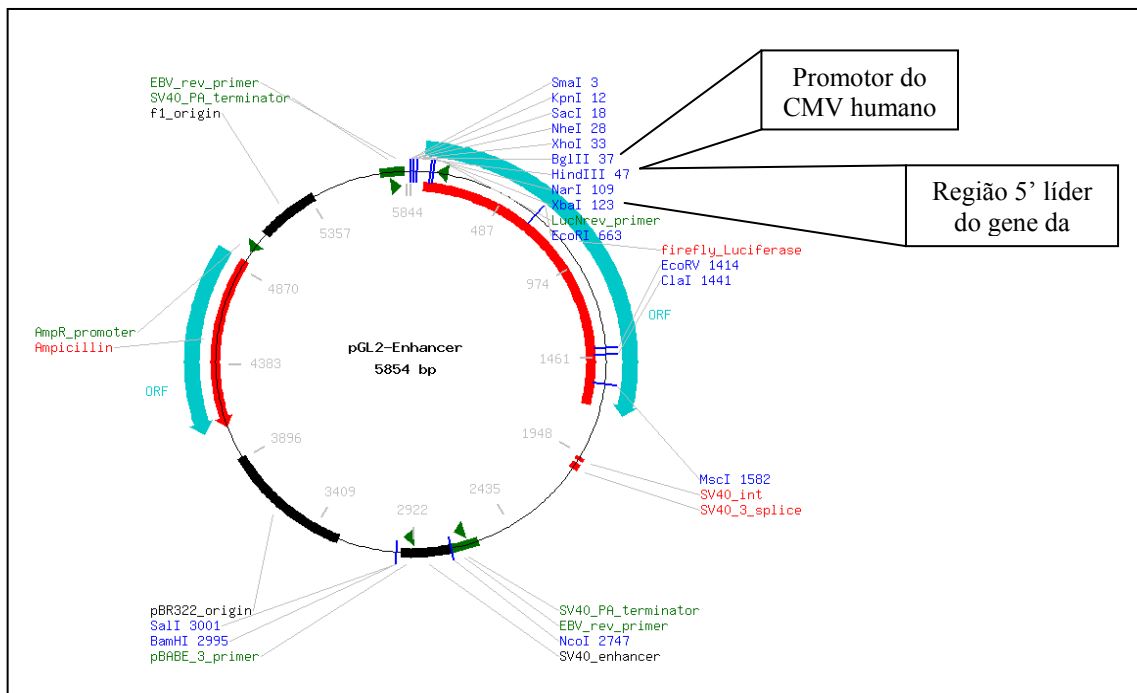
1. Mata, J., Marguerat, S., Bahler, J. (2005) Post-transcriptional control of gene expression: a genome-wide perspective. *Trends Biochem Sci.* **30** (9): 506 – 514
2. Mignone, F., Gissi, C., Liuni, S., Pesole, G. (2002) Untranslated regions of mRNAs. *Genome Biol.* **3**: 1-10
3. Sonenberg, N., Hinnebusch, A. (2009) Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell* **136**: 731-745
4. Cazzola, M., Skoda, R. (2000) Translational pathophysiology: a novel molecular mechanism of human disease. *Blood* **95**: 3280-3288
5. Law, G., Raney, A., Heusner, C., Morris, D. (2001) Polyamine Regulation Pausing at the Upstream Open Reading Frame of S-Adenosylmethionine Decarboxylase. *J. Biol. Chem.* **276**: 38036-38043
6. Anthony, D., Merrick, W. (1991) Eukaryotic Initiation Factor (eIF)-4F. Implications for a role in internal initiation of translation. *J. Biol. Chem.* **266**: 10218-10226
7. Meijer, H., Thomas, A. (2002) Control of eukaryotic protein synthesis by upstream open reading frames in the 5'-untranslated region of an mRNA. *Biochem. J.* **367**: 1-11
8. Pickering, B., Willis, A. (2005) The implications of structured 5' untranslated regions on translation and disease. *Semin. Cell. Dev. Biol.* **16**: 39-47
9. Kozak, M. (2002) Pushing the limits of the scanning mechanism for initiation of translation. *Gene* **299**: 1-34
10. Hemmings-Mieszczak, M., Hohn, T. (2006) A stable hairpin preceded by a short open reading frame promotes nonlinear ribosome migration on a synthetic mRNA leader. *RNA* **5**: 1149-1157
11. Silva, A., Romão, L. (2009) The mammalian nonsense-mediated decay pathway: to decay or not to decay! Which players make the decision? *FEBS Lett.* **583**: 499-505
12. Resch, A., Ogurtsov, A., Rogozin, I., Shabalina, S., Koonin, E. (2009) Evolution of alternative and constitutive regions of mammalian 5'UTRs. *BMG Genomics* **10**: 162
13. Mangus, D., Evans, M., Jacobson, A. (2003) Poly(A)-binding proteins: multifunctional scaffolds for the post-transcriptional control of gene expression. *Genome Biol.* **4**: 223
14. Jackson, R. (2005) Alternative mechanisms of initiating translation of mammalian mRNAs. *Biochem. Soc. Trans.* **33**: 1231-1241
15. Vilela, C., McCarthy, J. (2003) Regulation of fungal gene expression via short open reading frames in the mRNA 5' untranslated region. *Mol. Microbiol.* **49**: 859-867
16. Morris, D., Geballe, A. (2000) Upstream Open Reading Frames as Regulators of mRNA Translation. *Mol. Cell. Biol.* **20**: 8635-8642
17. Mize, G., Ruan, H., Low, J., Morris, D. (1998) The Inhibitory Upstream Open Reading Frame from Mammalian S-Adenosylmethionine Decarboxylase mRNA Has a Strict Sequence Specificity in Critical Positions. *J. Biol. Chem.* **273**: 32500-32505
18. Kochetov, A. (2008) Alternative translation start sites and hidden coding potential of eukaryotic mRNAs. *Bioessays* **30**: 683-691
19. Iacono, M., Mignone, F., Pesole, G. (2005) uAUG and uORFs in human and rodent 5' untranslated mRNAs. *Gene* **349**: 97-105
20. Wiestner A, Schlemper RJ, van der Maas AP, Skoda RC (1998) An activating splice donor mutation in the thrombopoietin gene causes hereditary thrombocythaemia. *Nat. Genet.* **18**:49–52
21. Liu L, et al. (1999) Mutation of the CDKN2A 5' UTR creates an aberrant initiation codon and predisposes to melanoma. *Nat. Genet.* **21**:128–132

22. Wen Y, 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.
23. Calvo, S., Pagliarini, D., Mootha, V. (2009) Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. *Proc. Natl. Acad. Sci. USA* **106**: 7507-7512
24. Kozak, M. (1991) An Analysis of Vertebrate mRNA Sequences: Intimations of Translational Control. *J. Cell Biol.* **115**: 887-903
25. Mathews, M., Sonenberg, N., Hershey, J. (2007) *Translational Control in Biology and Medicine.* Cold Spring Harbor Laboratory Press
26. Kozak, M. (2005) Regulation of translation via mRNA structure in prokaryotes and eukaryotes. *Gene* **361**: 13-37
27. Poyry, T., Kaminski, A., Jackson R. (2003) What determines whether mammalian ribosomes resume scanning after translation of a short upstream open reading frame? *Genes Dev.* **18**: 62-75
28. Kozak, M. (2001) Constraints on reinitiation of translation in mammals. *Nucleic Acids Res.* **29**: 5226-5232
29. Lovett, P., Rogers, E. (1996) Ribosome Regulation by the Nascent Peptide. *Microbiol Rev.* **60**: 366-385
30. Cao, J., Geballe, A. (1996) Coding Sequence-Dependent Ribosomal Arrest at Termination of Translation. *Mol. Cell. Biol.* **16**: 603-608
31. Hayden, C., Jorgensen, R. (2007) Identification of a 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
32. Sachs, M., Geballe, A. (2006) Downstream control of upstream open reading frames. *Genes Dev.* **20**:915-921
33. Casals-Pascual, C., Idro, R., Picot, S., Roberts, D., Newton, C. (2008) Can erythropoietin be used to prevent brain damage in cerebral malaria? *Trends Parasitol.* **25**: 30-36
34. Shoemaker, C., Mitsock, L. (1986) Murine Erythropoietin Gene: Cloning, Expression, and Human Gene Homology. *Mol. Cell. Biol.* **6**: 849-858
35. Maiese, K., Chong, Z., Shang, Y. (2008) Raves and risks for erythropoietin. *Cytokine Growth Factor Rev.* **19**: 145-155
36. Weidemann, A., Johnson, R. (2009) Nonrenal regulation of EPO synthesis. *Kidney Int.* **75**: 682-688
37. Vilalta, A., Wu, D., Margalith, M., Hobart, P. (2001) Rabbit EPO gene and cDNA: Expression of Rabbit EPO after Intramuscular Injection of pDNA. *Biochem. Biophys. Res. Commun.* **284**: 823-827
38. Milano, M., Schneider, M. (2007) EPO in cancer anemia: Benefits and potential risks. *Crit. Rev. Oncol. Hematol.* **62**: 119-125
39. Van der Meer, P., Voors, A., Lipsic, E., van Gilst W., van Veldhuisen, D. (2004) Erythropoietin in cardiovascular diseases. *Eur. Heart J.* **25**: 285-291
40. Ebert, B., Bunn, F. (1999) Regulation of the Erythropoietin Gene. *Blood* **94**:1864-1877
41. McGary, E., Rondon, I., Beckman, B. (1997) Post-transcriptional Regulation of Erythropoietin mRNA Stability by Erythropoietin mRNA-binding Protein. *J. Biol. Chem.* **272**: 8628-8634
42. Brines, M., Cerami, A. (2008) Erythropoietin-mediated tissue protection: reducing collateral damage from the primary injury response. *J. Int. Med.* **264**: 405-432
43. Horton, R. (1995) PCR Mediated Recombination and Mutagenesis: SOEing Together Tailor-Made Genes. *Mol Biotechnol* **3**: 94-99

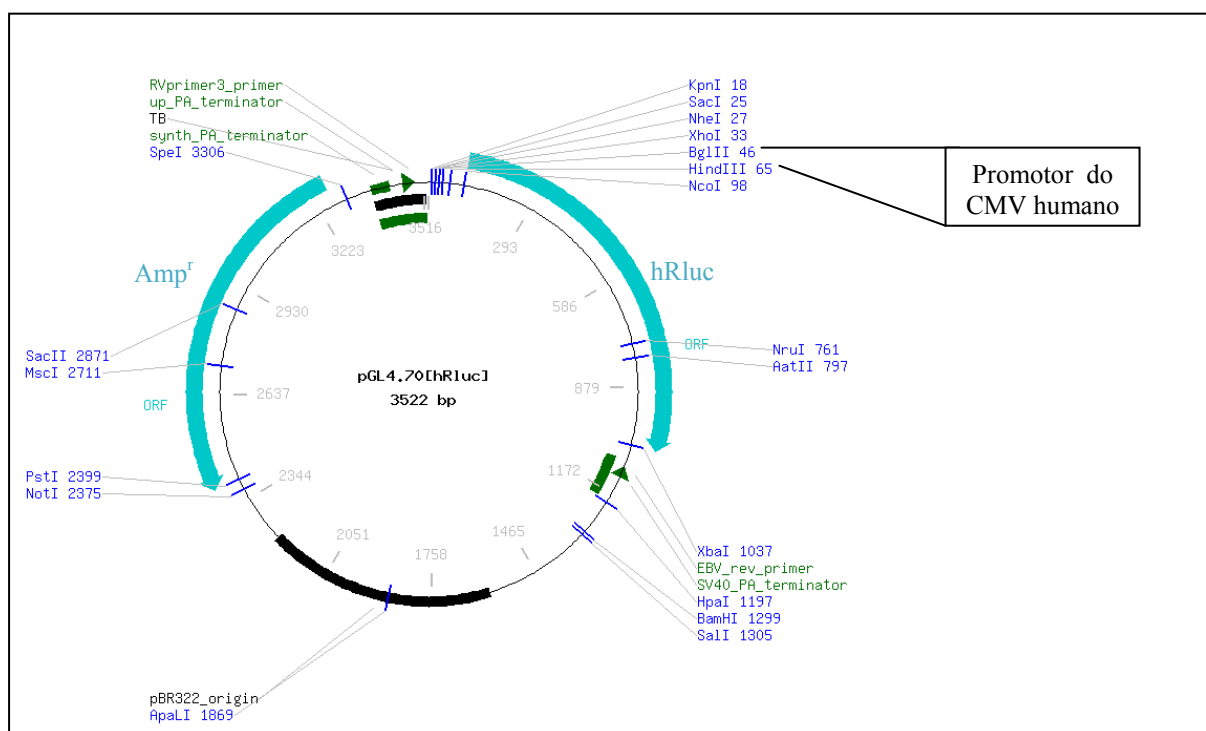
## Anexo A – Oligonucleótidos, Vetores e Sequências



**Figura A1 – Mapa físico do vetor pcDNA3.1/Hygro(+).** Encontram-se destacados por setas os locais de reconhecimento das enzimas de restrição utilizadas para a clonagem da sequência do promotor do CMV humano.



**Figura A2 – Mapa físico do vetor pGL2-enhancer.** Estão identificados os locais de reconhecimento das enzimas de restrição onde foram inseridos o promotor do CMV humano e a região 5' líder do gene da EPO.



**Figura A3 – Mapa físico do vetor pGL4.70.** Encontra-se identificado o local onde foi inserido o promotor do CMV humano.

**Tabela I. Sequência dos oligonucleótidos iniciadores**

Linker HindIII-EPO	5' – CCC <b>AAGCTT</b> CCCGGAGCCGGACCGG – 3'	Oligonucleótidos externos
pGL2 (as)	5' – CGTACGTGATGTTACCTC – 3'	
uORF-linker Luc (s)	5' – GCCAGGCGGGAGATGGAAGACGCCAAAAACATAAAAG – 3'	Oligonucleótidos para clonagem da região 5' líder
Luc-linker uORF (as)	5' – CTTTATGTTTTTGCGTCTTCCATCTCCGCGCCTGGC – 3'	
Mut uAUG (s)	5' – CTTCCCGGG <b>T</b> TGAGGGC – 3'	Oligonucleótidos para mutagêneses
Mut uAUG (as)	5' – GCCCTGA <b>A</b> CCCGGGAAG – 3'	
Mut uTer (s)	5' – AGGTTCG <b>C</b> ^GAGGGACCCC – 3'	
Mut uTer (as)	5' – GGTCCCT <b>C</b> ^GCGACCTGG – 3'	
Mut AUG Luc (s)	5' – GCGGAG <b>T</b> TGGAAGACGCC – 3'	
Mut AUG Luc (as)	5' – GTCTTCCA <b>A</b> CTCCGCGCC – 3'	

A vermelho está assinalada a zona de reconhecimento da enzima de restrição *Hind* III. O sublinhado nos oligonucleótidos iniciadores para a clonagem representam *linkers* para o início do gene da luciferase no primeiro e para o final da região 5' líder da EPO no segundo de forma a ocorrer a fusão destas duas sequências. Os nucleótidos a negrito e sublinhado correspondem aos alterados, o ^ representa a deleção de um nucleótido.



**Tabela II. Sequência dos oligonucleótidos iniciadores usados em PCR em tempo real**

Real-time Firefly (s)	5' – CGCTGGAGAGCAACTGCAT – 3'	Fragmento = 51bp
Real-time Firefly (as)	5' – CCAGGAACCAGGGCGTATCT – 3'	
Real-time Renilla (s)	5' – GAGAACGCCGTGATTTTTCTG – 3'	Fragmento = 51bp
Real-time Renilla (as)	5' – CCACAGGTAGCTGGAGGCAG – 3'	

## Anexo B – Soluções para imuno-deteção por *Western Blot*

Gel SDS de resolução a 10%:

- H<sub>2</sub>O destilada 2,45mL
- Tampão 1,25mL:  
1,5 M Tris/HCl, pH=8,8
- Acrilamida (40%) 1,25mL
- SDS (10%) 50μL
- APS (10%) 50μL
- TEMED 2.5μL

Gel SDS de concentração a 4%:

- H<sub>2</sub>O destilada 1,5mL
- Tampão 0,25mL:  
0,5M Tris/HCl, pH=6,8
- Acrilamida 0,2mL
- SDS (10%) 20μL
- APS (10%) 40μL
- TEMED 2μL

TBST, 10x:

- Tris 0,125M
- NaCl 0,37M
- TritonX-100 0,1%
- pH 7,6

TBST-*milk*:

- 7,5g leite em pó
- TBST 1x 50mL

Solução 1 ECL:

- Tampão ECL 980μL
- Luminol 15μL
- Ácido cumárico 5μL

Solução 2 ECL:

- Tampão ECL 1000μL
- H<sub>2</sub>O<sub>2</sub> 3μL