UNIVERSIDADE DE LISBOA

FACULDADE DE CIÊNCIAS

DEPARTAMENTO DE BIOLOGIA VEGETAL

Regulation of PERK gene expression by its upstream open reading frames

João Gonçalves da Ponte

Mestrado em Biologia Molecular e Genética

Dissertação orientada por:

Doutora Rita Zilhão

Doutora Luísa Romão

2025

Agradecimentos

Agradeço à Faculdade de Ciências da Universidade de Lisboa por me ter aceitado como aluno deste Mestrado. Agradeço também ao Instituto Nacional de Saúde Doutor Ricardo Jorge pelo acolhimento durante o campo de estágio.

À minha orientadora interna, Professora Drª Rita Zilhão, que me acompanhou durante a Licenciatura e Mestrado e que, pelas suas competências pessoais e profissionais, foi exemplo e referência para o meu desenvolvimento pessoal e académico, deixo um especial apreço por toda a atenção e disponibilidade dispensada.

À minha orientadora externa, Drª Luísa Romão, agradeço a transmissão do seu conhecimento e experiência que me permitiram adquirir e desenvolver competências acrescidas laboratoriais e de investigação, bem como a oportunidade e a confiança na realização deste trabalho que integra uma pequena parte de uma grande história do seu laboratório. À Drª Rafaela Lacerda, agradeço a ajuda e paciência durante todo o meu percurso no INSA.

Agradeço à coordenadora do Departamento de Genética Humana, Drª Glória Isidro, que me deu a oportunidade de integrar este departamento e desenvolver o meu trabalho.

Aos meus amigos, colegas de Mestrado e de laboratório, com quem partilhei momentos de alegrias e angústias, um agradecimento especial por termos estado juntos neste percurso.

À minha família, agradeço por me dar força e apoio incondicional em todos os momentos e por me incentivar neste caminho da formação contínua.

Abstract

Upstream open reading frames (uORFs) are *cis*-acting elements located within the 5' untranslated region (5'UTR) of transcripts, which can regulate the translation of the corresponding main open reading frame (mORF). In normal conditions, uORFs are typically repressors of downstream translation, as they can block ribosomal access to the mORF or even induce mRNA degradation through the nonsense-mediated mRNA decay (NMD) pathway. However, in stress conditions, phosphorylation of the eukaryotic initiation factor 2 (eIF2) allows the expression of several stress-responsive proteins via uORF-mediated mechanisms, while global mRNA translation is inhibited. During endoplasmic reticulum (ER) stress, the accumulation of unfolded proteins leads to activation of the ER-resident PKR-like ER kinase (PERK), which will phosphorylate the α-subunit of eIF2 as part of the stress-protective mechanisms of the unfolded protein response (UPR) and integrated stress response (ISR). This results in selective uORF-mediated translation of downstream effectors, which will drive stress resolution or cell death in case of prolonged stress. The dual role of PERK in regulating cell fate is implicated in a growing list of pathophysiological conditions, including neurological and cardiovascular diseases, ophthalmological disorders, viral infections, cancer, and diabetes. Moreover, mutations in the *EIF2AK3* gene encoding PERK are implicated in a rare autosomal recessive disorder, the Wolcott-Rallison Syndrome (WRS). Data from ribosome-profiling (Ribo-seq) studies indicate the existence of uORFs within *PERK* 5'UTR which could be involved in regulating PERK expression.

This work aims to study the translational regulatory role of the uORFs identified in *PERK*'s 5'UTR and estimate its impact on cell homeostasis and human disease. We wish to highlight the importance of including 5'UTRs in the screening of disease-related mutations, as well as the necessity of functional studies to assess their relevance in the pathogenesis of human diseases, as it may provide vital information for developing new therapeutic strategies.

Keywords

ER stress, PERK, uORF, UPR, Wolcott-Rallison Syndrome.

Resumo

A expressão génica eucariótica é um processo complexo que compreende vários passos altamente regulados. Em particular, a tradução do RNA mensageiro (mRNA) representa um passo chave da expressão génica, cuja regulação permite à célula rapidamente alterar a síntese proteica de uma forma espacial e temporal em resposta a diferentes estímulos. O processo de tradução divide-se em quatro etapas: iniciação, alongamento, terminação e reciclagem de ribossomas. A etapa de iniciação representa o passo limitante e, como tal, é o mais regulado. Essa regulação ocorre com o envolvimento de múltiplos fatores e elementos em *cis*. De entre os múltiplos elementos que podem regular esta etapa, encontram-se as pequenas grelhas de leitura a montante da grelha principal (do inglês, *upstream open reading frames*, uORFs), que correspondem a pequenas regiões potencialmente traduzidas definidas por um codão de iniciação na região 5' não traduzida (do inglês, 5' *untranslated region*, 5'UTR) dos mRNAs, em fase com um codão de terminação a montante ou sobreposto com a grelha de leitura principal (do inglês, *main open reading frame*, mORF) que codifica a proteína. As uORFs são tipicamente consideradas repressoras da tradução em condições fisiológicas normais, uma vez que funcionam como "barreiras" aos ribossomas que, ao traduzirem-nas, podem não chegar a traduzir a mORF. Outra característica reguladora importante das uORFs é a possibilidade de induzirem a degradação do mRNA pelo processo de decaimento do RNA mensageiro mediado por mutações *"nonsense"* (do inglês, *nonsense-mediated mRNA decay*, NMD), uma vez que os seus codões de terminação podem ser reconhecidos como prematuros aquando da terminação da tradução. Contudo, em condições de stress, as uORFs são muitas vezes responsáveis por permitir a tradução de determinadas mORFs. Isto acontece porque nestas condições ocorre a fosforilação do fator eucariótico de iniciação da tradução 2α (do inglês, *eukaryotic initiation factor* 2α, eIF2α), o que favorece o não reconhecimento de codões de iniciação de algumas uORFs normalmente com contextos Kozak fracos (do inglês, *leaky scanning* ou *ribossome bypass*).

A proteína cinase tipo PKR residente no retículo endoplasmático (do inglês, PKR-*like* ER *kinase*, PERK) é uma das responsáveis por fosforilar o eIF2α em condições de stress. Estruturalmente, a PERK é uma proteína transmembranar do retículo endoplasmático (ER) que possui três domínios: (i) um domínio lumenal regulatório; (ii) um domínio transmembranar; e (iii) um domínio catalítico citoplasmático. Em condições de homeostasia, a PERK encontra-se num estado monomérico inativo promovido por chaperonas. No entanto, quando ocorre stress do ER, a acumulação de proteínas mal enroladas promove a oligomerização de domínios lumenais da PERK, causando uma aproximação dos respetivos domínios citoplasmáticos de forma a promover a sua auto fosforilação e consequente ativação. Uma vez ativada, a PERK fosforila o eIF2α de forma a inibir globalmente a tradução do mRNA como mecanismo de proteção, enquanto favorece a tradução mediada por uORFs de proteínas responsáveis por aumentar a capacidade processadora do ER como parte dos mecanismos da resposta a proteínas mal enroladas (do inglês, *unfolded protein response*, UPR) e da resposta integrada ao stress (do inglês, *integrated stress response*, ISR). Dentro destas proteínas encontra-se a ATF4 (do inglês, *activating transcription factor* 4), a CHOP (do inglês, CCAAT-*enhancer-binding protein homologous protein*) e a GADD34 (do inglês, *growth arrest and* DNA *damage inducible protein* 34) que, no caso de um estímulo prolongado ou intenso, podem também desencadear vias pró-apoptóticas. O papel da PERK na regulação da homeostasia da célula e da sua sobrevivência tem sido implicado em várias doenças, incluindo doenças neurológicas, cardiovasculares e oftalmológicas, bem como infeções virais, cancro e diabetes *mellitus*. Além disso, mutações no gene que codifica a PERK (o gene *EIF2AK3*) foram associadas ao desenvolvimento de uma doença rara autossómica recessiva, síndrome de Wolcott-Rallison (do inglês, Wolcott Rallison syndrome, WRS), caracterizada por diabetes mellitus neonatal permanente e displasia epifisária múltipla. Isto sugere que a expressão e a atividade da PERK são essenciais para o normal funcionamento da célula, devendo ser corretamente reguladas. Curiosamente, dados de perfil ribossomal

(do inglês, *ribosome profiling*) mostram a existência de vários eventos de iniciação de tradução na 5'UTR do mRNA da PERK, incluindo em codões de iniciação de tradução não canónicos, o que sugere a existência de uORFs com potencial regulatório.

Tendo em conta a escassez de estudos sobre as uORFs do mRNA da PERK e os seus efeitos nesta mesma proteína, o objetivo do presente trabalho focou-se na complementaridade de estudos realizados em laboratório sobre este mesmo tema. Anteriormente, tinha sido observado que o transcrito da PERK possui oito uORFs — cinco iniciadas por AUG e três iniciadas por codões não canónicos — que, em conjunto, inibem fortemente a tradução da mORF, tendo sido registado uma redução de 92% na atividade relativa da luciferase promovida pela 5'UTR da PERK numa linha celular de cancro colorretal. Destas oito, foi verificado também que este efeito repressor é mediado principalmente pela uORF1, iniciada por um codão AUG com contexto Kozak forte, sendo maximizado pelas uORFs 3, 4 e 8, também iniciadas por AUG, mas com contextos intermédios. As restantes uORFs da PERK parecem não contribuir significativamente para este processo regulatório. Em concordância com o seu papel regulatório ao nível da tradução, estudou-se a possibilidade destas primeiras uORFs incluírem-se em transcritos que gerassem pequenos péptidos-sinal independentes do seu gene principal, podendo revelar novas funções. No entanto, nenhuma das uORFS apresentou um resultado preditivo significante para a existência deste tipo de estruturas. De maneira a perceber os mecanismos reguladores da tradução mediados pelas uORFs da PERK que poderiam atuar na tradução da mORF, — fora dos já descritos — foram silenciados vários fatores inerentes de iniciação. Adicionalmente, para verificar se a presença das uORFs é determinante para a expressão da PERK na ausência destes fatores foram feitos ensaios de luminometria com plasmídeos repórter, contendo a 5' UTR da PERK na presença e ausência das suas uORFS. Dos fatores estudados, a ausência de eIF2D e eIF4G2 apresentou sempre resultados estatisticamente significativos, revelando a importância destes na regulação dos níveis proteicos de PERK. Adicionalmente, a análise resultante dos ensaios de luminometria sugere que, apesar da presença de uORFs ser necessária para inibir a tradução geral da mORF, o knockdown das proteínas em estudo não afetou a expressão da mesma, estando esta sobre o efeito das uORFs da PERK ou não. Contudo, são ainda necessários testes de maneira a verificar a contribuição de todos os fatores em estudo na manutenção dos níveis de mRNA da PERK.

Tendo em conta o papel notório das uORFs na regulação dos níveis da PERK e a relação desta cinase com o desenvolvimento de WRS, pretende-se analisar o efeito de mutações identificadas em doentes com WRS, que eliminam ou alteram as uORFs, na regulação da expressão da PERK. Resultados obtidos anteriormente no laboratório demonstram que algumas destas mutações podem alterar a taxa de produção de PERK, o que poderá explicar o fenótipo da doença, necessitando, no entanto, de confirmação experimental adicional. Pretende-se também estudar mais extensivamente a função do péptido codificado pela uORF1, visto ser a principal responsável pela regulação da tradução da mORF da PERK.

Resumidamente, neste estudo é apresentado um exemplo de um transcrito cuja tradução da mORF é regulada por uORFs. Adicionalmente, destaca-se a importância de se incluírem as 5'UTRs de genes como a PERK, no rastreio de mutações que possam estar associadas ao desenvolvimento de doenças genéticas, bem como o estudo detalhado das suas consequências biológicas. Este tipo de informações é imprescindível para o desenvolvimento de novas terapias ou para a aplicação adequada das já existentes.

Palavras-chave

PERK, síndrome de Wolcott-Rallison, stress do retículo endoplasmático, uORF, UPR.

Index

List of tables

List of figures

Abbreviations and acronyms

5'UTR	5' untranslated region
5MP	eIF5-mimic protein
aa	amino acid
ABCE1	ATP binding cassette subfamily E member 1
A site	aminoacyl site
ATF	activating transcription factor
ATP	adenosine triphosphate
BiP	binding immunoglobulin protein
BSA	bovine serum albumin
CDS	coding DNA sequence
CFNS	Craniofrontonasal syndrome
CHOP	C/EBP homologous protein
CO_2	carbon dioxide
ddH_2O	bidistilled water
DENR	density-regulated protein
DMSO	dimethyl sulfoxide
eEF	eukaryotic elongation factor
EFNB1	Ephrin B1
eIF	eukaryotic initiation factor
eIF2α-P	phosphorylated eIF2α
ER	endoplasmic reticulum
eRF	eukaryotic release factor
FLuc	Firefly luciferase
FoxO	forkhead box O
GADD	growth arrest and DNA damage-inducible protein
GAPDH	glyceraldehyde 3-phosphate dehydrogenase
GCN2	general control non-derepressible-2

GDP	guanosine diphosphate
GO	gene ontology
GSC	glioblastoma stem cells
GTP	guanosine triphosphate
hCMV	human cytomegalovirus
HCT116	human pre-metastatic colorectal carcinoma-derived cell line
HEK293	human embryonic kidney 293 cells
HRI	heme-regulated inhibitor
IRE1	inositol-requiring enzyme 1
ISR	integrated stress response
Keap1	Kelch-like ECH-associated protein 1
KO	Knockout
LB	Luria-Bertani
Leu	leucine
Met	methionine
mORF	main open reading frame
miRNA	microRNA
MPEP	MYC pre-mRNA encoded protein
mRNA	messenger RNA
NMD	nonsense-mediated mRNA decay
NP-40	Nonidet P-40
Nrf2	nuclear factor erythroid-derived 2-like 2
nt(s)	nucleotide(s)
ORF	open reading frame
PABP	poly(A)-binding protein
PBS	phosphate-buffered saline
PCR	polymerase chain reaction

pDNA	plasmid DNA
PERK	PKR-like endoplasmic reticulum kinase
PIC	pre-initiation complex
PKR	protein kinase double-stranded RNA-dependent
P site	peptidyl site
PTC	premature termination codon
Ribo-seq	ribosome profiling
RLuc	Renilla luciferase
RT	reverse transcription
SDS-PAGE	sodium dodecyl sulfate-polyacrylamide gel electrophoresis
Ser	serine
siRNA	small interfering RNA
SNPs	single nucleotide polymorphisms
SOEing	synthesis by overlap extension
SP	signal peptide
STN1	CST complex subunit 1
TBS	tris-buffered saline
TC	ternary complex
Tg	thapsigargin
TRKB	tropomyosin receptor kinase B
$tRNA_i^{Met}$	methionyl-initiator tRNA
uORF	upstream open reading frame
UPR	unfolded protein response
WRS	Wolcott-Rallison syndrome

1. Introduction

In a simplified way, gene expression can be defined as the flow of genetic information encoded in genes to an intermediate molecule, the messenger RNA (mRNA), which ultimately gives rise to proteins — a process known as the central dogma of molecular biology. However, this expression comprises complex interconnected transcriptional and post-transcriptional processes regulated at different levels to determine which, when, and where a pool of specific proteins is being produced [1]. Despite transcription presenting one of the earliest points for gene expression regulation, post-transcriptional control of gene expression constitutes a faster and reversible way for the cell to adapt to changes in its surrounding environment. This translational oversight plays an important role in ensuring cell homeostasis and driving cell growth, proliferation, and differentiation, so, any dysregulation or error in its steps may lead to disease [2,3].

1.1. Overview of eukaryotic mRNA translation

Translation can be described as the process in which proteins are produced using RNA molecules as a template. It is a complex cyclical mechanism that can be divided into four phases — initiation, elongation, termination and ribosome recycling, all of them requiring their own set of specific factors and conditions [1]. The initiation step is a key component of the translation process, therefore being its most highly regulated phase [4]. Its goal is to assemble the large (60S) and small (40S) ribosomal subunits into an elongation-competent 80S ribosome with the initiator methionyl-transfer RNA (Met-tRNAi) positioned over the start codon of the mRNA [3].

Canonical translational initiation begins with the formation of a ternary complex of Met-tRNAi, eIF2 and guanosine triphosphate (GTP), which assemble with the 40S ribosomal subunit and eukaryotic initiation factors (eIFs) 1, 1A, 3 and 5, forming a pre-initiation complex (PIC). eIFs 1, 1A and 3 bind to the 40S subunit and induce an open conformation, which allows the binding of the ternary complex. The PIC is recruited to the 5′ end of the mRNA, marked by a 7-methylguanosine cap structure, through eIF3, the poly(A)-binding protein (PABP), eIF4B and the eIF4F complex. Once loaded onto the mRNA, the open PIC scans the 5′ untranslated region (UTR) to locate the first start codon in a favourable context. When the start codon is recognized, the PIC ejects eIF1, triggering the conversion of eIF2 to its guanosine diphosphate (GDP)-bound state via gated phosphate release. These events prompt the transition to a closed conformation of the PIC, which stabilizes its interaction with the mRNA and Met-tRNAi, arresting the scanning process. After that, eIF2 and eIF5 dissociate. Finally, eIF5B collaborates with eIF1A to mediate the coupling of the 60S subunit and assemble the 80S initiation complex [5,6] (Figure 1.1).

In contrast to the complex factor requirements in translation initiation, elongation is assisted by a minimal set of factors. It begins when Met-tRNAi is bound in the peptidyl site (P site) of the ribosome [7]. The anticodon of the Met-tRNAi is base-paired with the start codon of the messenger RNA (mRNA), and the second codon of the open reading frame (ORF) is in the aminoacyl site (A site) of the ribosome. Elongation starts with delivery of the cognate elongating aminoacyl-tRNA to the A site of the ribosome. The eukaryotic translation elongation factor 1A (eEF1A) is activated upon binding GTP and forms a ternary complex when bound to an aminoacyl-tRNA. The eEF1A•GTP•aminoacyl-tRNA complex binds in the A site [3]. Base-pairing interactions between the anticodon of the aminoacyl-tRNA and the A site codon trigger GTP hydrolysis by eEF1A. The eEF1A•GDP complex is released and the aminoacyl-tRNA is accommodated into the A site. Following release of eEF1A and accommodation of the aminoacyl-tRNA into the A site, the factor eIF5A binds in the E site, interacts with the acceptor arm of the peptidyl-tRNA and promotes peptide bond formation by inducing a favourable positioning of the

substrates. The nascent peptide is transferred from the peptidyl-tRNA in the P site to the amino group of the A-site aminoacyl-tRNA to form a new extended peptidyl-tRNA [7]. This process is repeated until a stop codon is encountered.

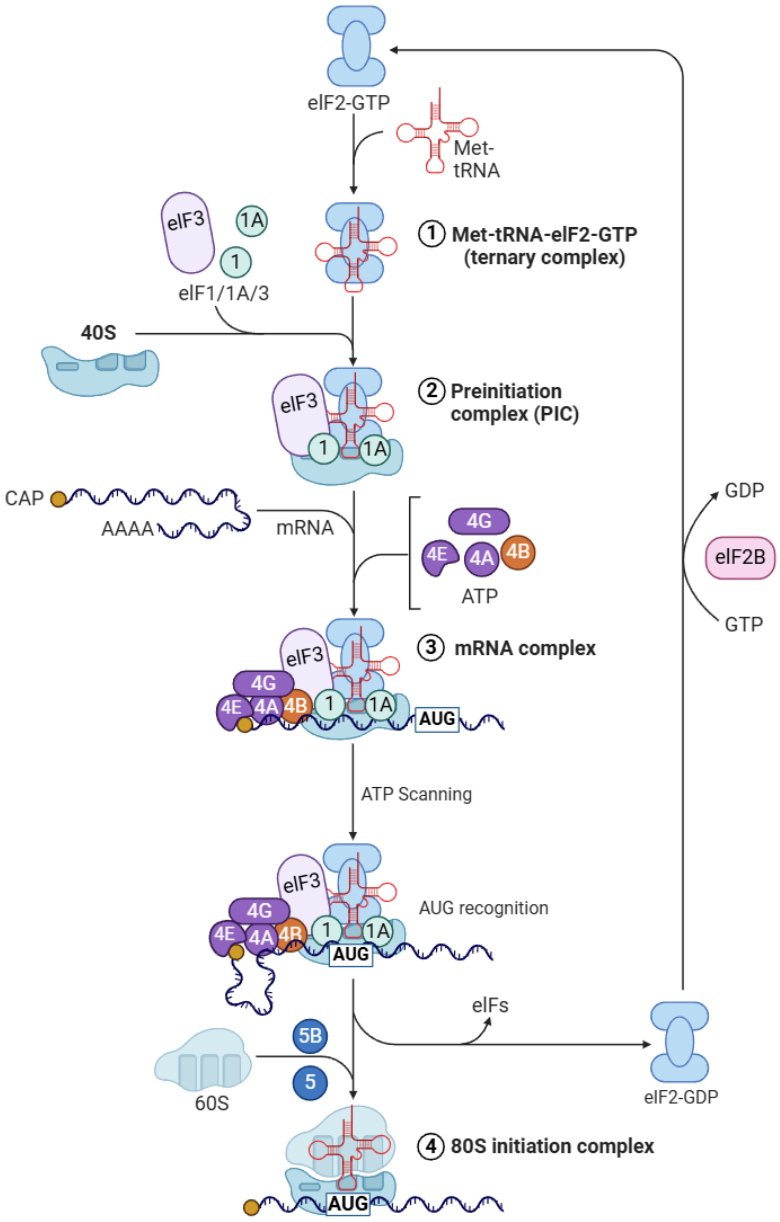


Figure 1.1 - Model of the canonical eukaryotic translation initiation process. The process begins with the assembly of the PIC, with eIFs 1A, 1, 3 and 5 binding initially and stimulating recruitment of the TC, composed of eIF2, GTP and Met-tRNAi Met. The PIC attaches near the 5' cap of the mRNA through interaction of eIF3 and eIF4G, which is part of the eIF4F complex (eIFs 4G, 4E and 4A), forming an activated mRNA. Base-pairing between the start codon and the anti-codon of the tRNAi Met at the P-site promotes conformational changes within the PIC, with the consequent release of eIF1 and eIF2-GDP in complex with eIF5. Then, eIF5B promotes joining of the 60S subunit, with release of eIF5B-GDP and eIF1A to form the 80S ribosome, ready to continue with the elongation phase. The released eIF2-GDP is then recycled to eIF2-GTP by the exchange factor, eIF2B, to start a new round of translation initiation. Image adapted from: Yartseva, V. (n.d.). https://app.biorender.com/illustrations/67c1a957ff028c8328471ee9

Termination is triggered when a stop codon enters the A site of the ribosome and is mediated by the release factors eRF1 and eRF3, forming a ternary eRF1/eRF3–guanosine triphosphate (GTP) complex [8]. eRF1 is responsible for recognizing all stop codons, inducing the release of the nascent polypeptide from the P-site peptidyl-tRNA, whereas eRF3 is a GTPase that enhances polypeptide release. This results in a post-termination complex (post-TC), which is recycled by splitting of the ribosome, through

ATP binding cassette subfamily E member 1 (ABCE1), an NTPase critically required for this phase [3,9]. This step is followed by release of deacylated tRNA and messenger RNA (mRNA) from the 40S subunit via redundant pathways involving initiation factors, ligatin (eIF2D) or density-regulated protein (DENR). The recycling process enables ribosomes and mRNAs to participate in multiple rounds of translation [10]

1.2. Cellular mechanisms of translational control

As mentioned before, translation represents a fundamental part of gene expression and accounts for a significant proportion of the energy budget of a cell, requiring tight regulation [4], which is achievable through two distinct processes: (i) global control, affecting translation of most mRNAs in the cell at a particular moment, and (ii) transcript-specific control, in which translation of a specific or a limited subset of mRNAs is regulated without affecting general protein synthesis [11]. The general mechanism the cell uses to control global mRNA translation comprises the reversible phosphorylation of some initiation factors, as well as the regulators that interact with them. The eIF2 complex, composed of three subunits (α, β, and γ) is a well-characterized example of a factor targeted by this type of regulation, being usually associated with translation impairment [11,12]. There are 4 protein kinases able to phosphorylate eIF2α, each of them activated by specific cellular stresses: the heme-regulated inhibitor (HRI), which responds to heme deprivation; the protein kinase double-stranded RNA-dependent (PKR), which is activated by double-stranded RNAs; the general control non-derepressible-2 (GCN2), activated by amino acid starvation; and the PKR-like endoplasmic reticulum (ER) kinase (PERK), a transmembrane ER enzyme with its kinase domain in the cytoplasm, which is activated during ER stress caused by misfolded proteins [3] (Figure 1.2). Despite the mRNA translation attenuation induced by these kinases, there are mRNAs with unique structural features that can evade and be translated as part of stress-responses [13], as they mediate their specific translational control, according to cellular environment. These structural features include: (i) specific binding sites for molecular ligands to form stable ribonucleoprotein complexes that may inhibit or promote translation; (ii) specific binding sites for small regulatory microRNAs (miRNAs) (iii) RNA secondary or tertiary structures that impair ribosomal scanning; (iv) internal ribosome entry sites (IRES), which are highly structured RNA regions that induce cap-independent translation and (v) upstream open reading frames (uORFs), which usually inhibit translation of the downstream main ORF (mORF) [1-3,11,12].

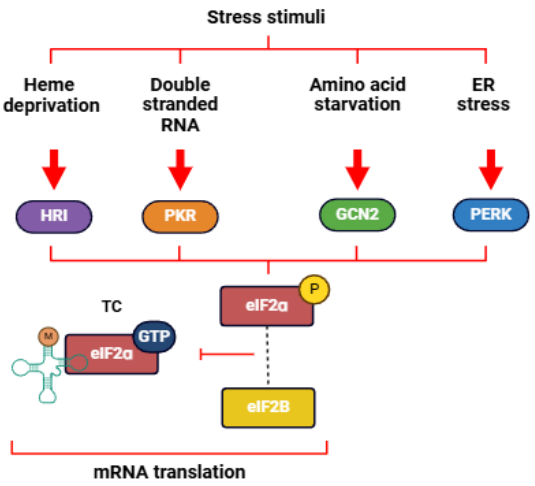


Figure 1.2 - Signalling pathways of global translational control. Different cellular stresses activate specific kinases that phosphorylate eIF2α. This, in turn, forms a stable complex with eIF2B that inhibits ternary complex formation and global mRNA translation. Red arrows indicate inhibitory events. Image adapted from: Fernandes, R. (2020, November 1). uORF-mediated translational regulation of PERK: implications for cell homeostasis and human disease. Repositorio.ul.pt. http://hdl.handle.net/10451/48530

1.3. uORFs as regulators

A uORF can be described as a potentially translated small ORF with its start codon located within the 5'UTR of an mRNA and its in-frame stop codon upstream or overlapped with the mORF [1]. Bioinformatics studies have identified uORFs fully situated within the 5′ UTR or partially overlapped with the main coding sequence (CDS) in multiple animal transcripts [12-14]. Gene ontology (GO) analyses have indicated that uORF-harboring transcripts are enriched in classes like oncogenes, transcription factors, cellular receptors, and genes involved in cell growth and differentiation control [14,15]. This, allied with the fact that many uORF start codons are evolutionarily conserved between species, argues for a translational regulatory role [16-18]. However, there are highly variable transcript specific structural properties that add multiple layers of complexity to the uORF-mediated translational control, which may also result in the opposite effect, that is, induction of mORF translation [19]. These structural features include the sequence context of the initiation codon and the surrounding secondary/tertiary structures of the mRNA, the type of initiation codon, the length and sequence of the uORF, the distance from the uORF stop codon to the mORF start codon, and the existence of multiple uORFs [16,17,20].

The best-known examples of this versatility of the uORF-mediated translational activity are provided by mRNAs that encode stress-responsive proteins. As mentioned before, multiple stress stimuli induce phosphorylation of eIF2α, which consequently impairs global translation initiation. However, eIF2α-P induces translational reprogramming, in which uORFs mediate the translation of specific proteins essential to overcome the stress or, in the case of prolonged or chronic stimuli, induce cell death [11]. While specific uORF-mediated mechanisms allow the expression of these proteins during stress, other mechanisms involving the same uORFs inhibit their translation in homeostatic conditions [21–23]. This highlights the importance of the above-mentioned structural properties of the transcripts and the uORFs to achieve the translational control mechanism that appropriately regulates gene expression according to the cellular context [11]. Overall, translational repression can be achieved by ribosome dissociation and recycling after uORF translation, by ribosome stalling during the elongation or termination processes, and, in case the uORF stop codon is recognized as a premature termination codon (PTC), by inducing degradation of the mRNA through nonsense-mediated mRNA decay (NMD) [1,24,25].

1.4. The mechanism of nonsense-mediated mRNA decay

Several studies reported that some uORFs can induce mRNA decay of the respective transcripts by triggering NMD [3,24]. Since its discovery, NMD has been described as a quality control mechanism that degrades aberrant transcripts derived from mutations or errors in the mRNA processing that introduce premature termination codons (PTCs) [25,26]. However, recent transcriptome-wide studies revealed that NMD degrades many mRNAs from normal, functional protein coding genes, thus arising as a mechanism of post-transcriptional control of gene expression, responsible for the regulation of 3–20% of eukaryotic transcripts [25,27,28]. mRNA degradation by NMD requires the recognition of a translational termination event as 'premature'. Therefore, the presence of uORFs in the 5'UTR may trigger the NMD machinery, as the uORF stop codon can be considered a PTC. However, uORFs are not the only features capable of inducing this mechanism. Despite having no relation with uORFs, other existing NMD-inducing features include the presence of a long 3'UTR or introns downstream of the stop codon [29]. Nevertheless, the rules that define a PTC remain vague and the extent of stop codon recognition as an NMD substrate are still investigated [30]. However, despite the link between uORFs and NMD observed for example in transcripts like CST complex subunit STN1 (STN1) [31] and activating transcription factor 4 (ATF4), where overlapping uORFs results in inhibition of mORF expression and consequent NMD induction, the mRNAs bearing uORFs cannot be generally considered

NMD-targets, since most of the translated uORFs do not affect the transcript mRNA level and so, do not trigger this mechanism [3] (Figure 1.3).

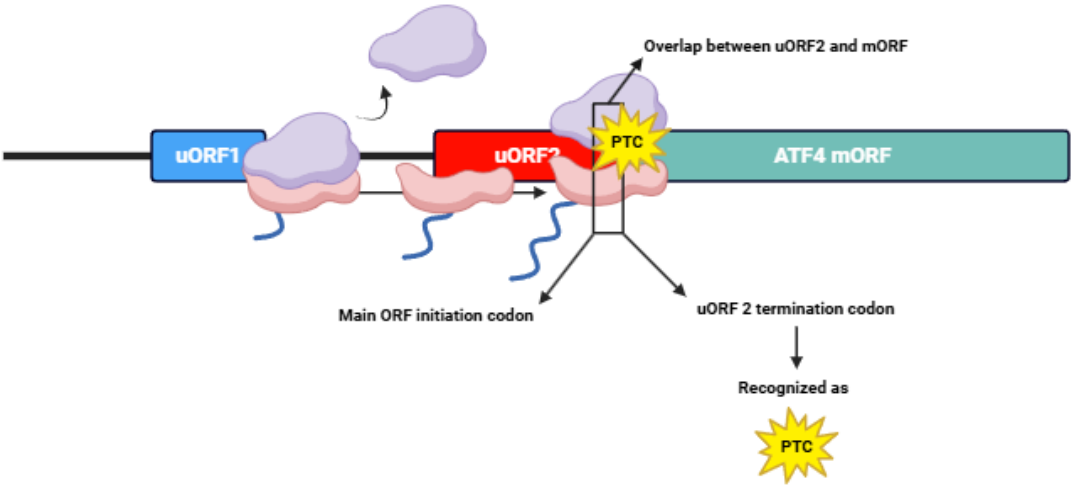


Figure 1.3 - uORF-mediated control of ATF4 expression in normal conditions. Ribosomes scanning the ATF4 mRNA initiate translation at uORF1. After termination, the 40S ribosomal subunits quickly reacquire a new ternary complex and reinitiate translation at uORF2, which overlaps out-of-frame with ATF4 mORF. Translation of uORF2 results in ribosome termination 3′ of the ATF4 initiation codon, which inhibits mORF translation. At the same time, the uORF2 stop codon is recognized as a premature termination codon (PTC), activating the NMD machinery that leads to mRNA decay. Image adapted from: Fernandes, R. (2020, November 1). uORF-mediated translational regulation of PERK: implications for cell homeostasis and human disease. Repositorio.ul.pt. http://hdl.handle.net/10451/48530

1.5. uORFs in human disease

The key role of uORFs in regulating the expression of many proteins is crucial for a variety of biological processes, implying that the dysregulation of their activity may lead to the development of diseases. Many bioinformatics studies mapped genetic variants within the 5'UTRs of transcripts and found alterations that can interfere with the regulatory function of uORFs [14,16]. Single nucleotide polymorphisms (SNPs) that create or delete a uORF were found in 509 genes and multiple SNPs that affect uAUGs and can alter the Kozak context of uORFs were also identified in a group of 2610 human genes [32]. These findings suggest that the expression of several proteins may be subject to specific variability created by polymorphic uORFs with potential clinical significance. In addition to polymorphisms that affect uORFs, rare mutations may have similar effects and lead to disease. An example of a disease that can be related to genetic variants in uORFs is, for instance, Craniofrontonasal syndrome (CFNS), which is a rare X-linked dominant disorder characterized by frontonasal malformations typically caused by loss-of-function mutations in the Ephrin B1 (*EFNB1*) gene. Interestingly, the wild type *EFNB1* transcript contains two uORFs: the first one containing 10 codons, and the second one, uORF2, with 4 codons, responsible for dampening the mORF translation [33,34]. Formerly, a c.-95 T>G variant has been described to disrupt the stop codon of uORF2, placing its start codon in-frame with a stop codon located 44 codons inside *EFNB1* mORF. This extended uORF2 was reported to have an accentuated inhibitory effect on mORF translation, which would explain the CFNS phenotype [34].

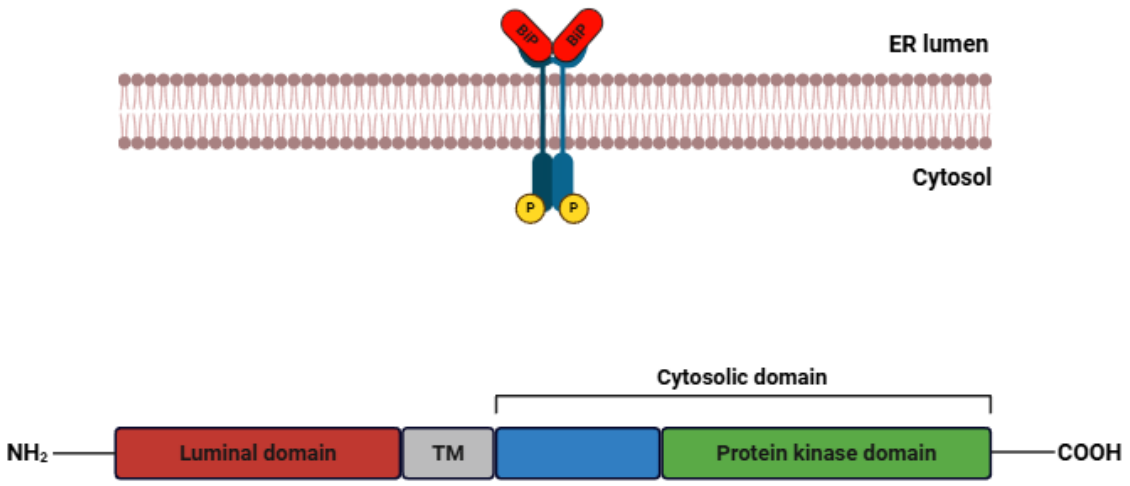
1.6. The dual role of PERK in stress response

As mentioned before, PERK is part of the eIF2 kinases family and primarily activated upon ER stress. This can be caused by abnormal high translation in the ER and/or inefficient protein folding capacity due to decreased expression of chaperones, perturbations in the cell energy levels, calcium homeostasis or redox status [27,35,36]. Structurally, PERK is an 1116 aa-ER transmembrane protein composed of three main domains: (i) the N-terminal domain placed at the ER lumen; (ii) a transmembrane domain; (iii) and the cytosolic C-terminal domain [37] (Figure 1.4 A).

During homeostatic conditions, PERK is thought to be maintained in an inactive state promoted by binding immunoglobulin protein (BiP) chaperones, forming a stable complex with ER luminal sequences flanking the transmembrane domain of PERK. However, when misfolded proteins accumulate in the ER, the chaperones are released to help in the folding process, allowing for homo-oligomerization of PERK lumenal domains, which brings the cytoplasmic domains near each other to induce its activation [38,39]. Once activated, PERK phosphorylates the α-subunit of eIF2 to shut down global protein translation, while allowing the expression of downstream targets harbouring uORFs, as part of the protective mechanisms of the integrated stress response (ISR) [40-43]. When ER stress occurs, PERK is activated and initiates a cascade of events to achieve cell homeostasis as part of another stress response called unfolded protein response (UPR). This cascade of events elicited by ER stress is expected to reduce the burden of misfolded proteins in the ER, enhance its folding capacity, and restore the energy and nutrient levels of the cell. However, in prolonged or chronic ER stress conditions, the UPR enters a maladaptive phase in which the signalling pathways change and culminate in cell death [3] (Figure 1.4 B).

Thus, PERK has a dual role in response to stress: in the initial phase of the UPR/ISR, PERK induces stress-corrective mechanisms that promote cell survival, whereas its persistent activation favours the expression of pro-apoptotic factors. It has been reported that constitutive PERK signalling impairs cell proliferation and promotes apoptosis, stressing the importance of regulating PERK activity to avoid its harmful effects [43].

A



B

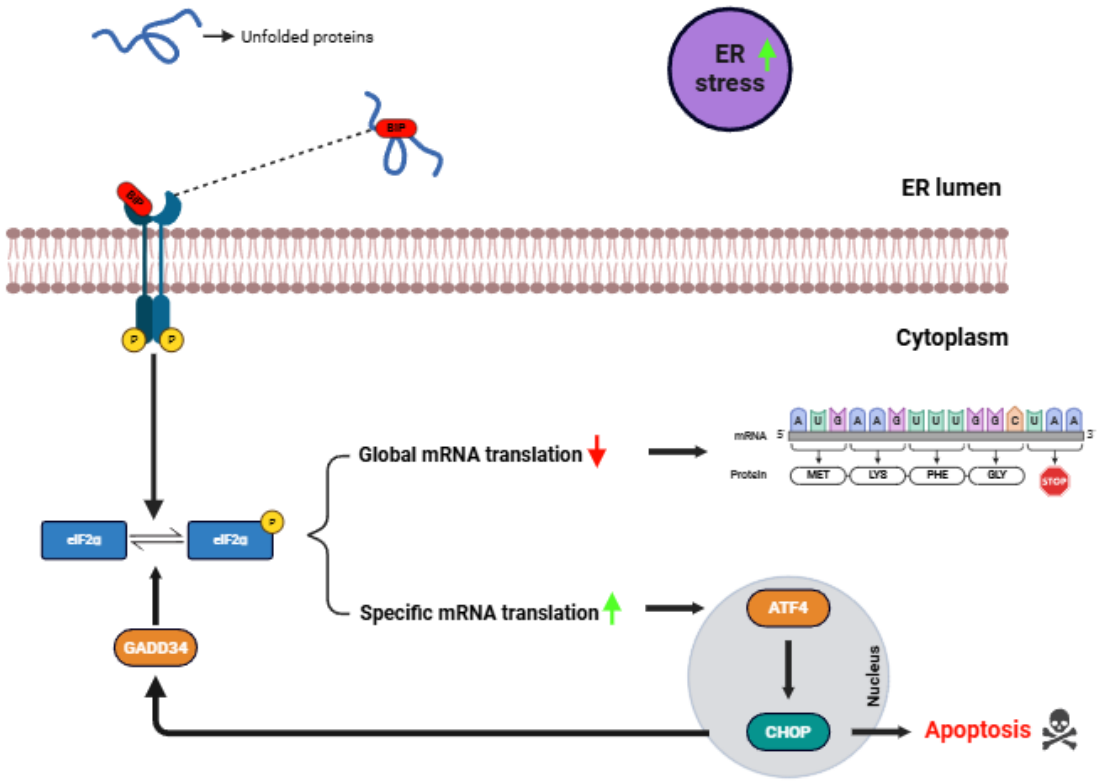


Figure 1.4 - PERK structure and signalling cascade. (A) Physical and chemical structure of PERK **(B)** During endoplasmic reticulum (ER) stress, unfolded proteins accumulate within the ER lumen. These proteins promote dissociation of the BiP chaperones from the transmembrane sensors and PKR-like ER kinase (PERK), to help in the folding process. Consequently, the three branches of the unfolded protein response (UPR) become activated, triggering a cascade of events aiming to solve the stress (indicated by black arrows). PERK activation promotes phosphorylation of the eukaryotic initiation factor 2α (eIF2α) to inhibit global mRNA translation, while allowing the exclusive translation of specific mRNAs, like the one encoding ATF4. This transcription factor promotes expression of autophagy-related proteins and the transcription factor, CCAAT-enhancer-binding protein homologous protein (CHOP). The latter is involved in the expression of the growth arrest and DNA damage-inducible protein 34 (GADD34), which functions in the negative regulation of the PERK branch through dephosphorylation of eIF2α. If stress resolution fails, pro-apoptotic cascades are activated. Image created through BioRender (https://www.biorender.com/).

1.7. PERK and the Wolcott-Rallison syndrome

The fact that PERK is crucial in maintaining cell homeostasis and controlling cell fate during stress suggests that its dysregulated expression or activity may lead to major issues. Indeed, shortly after PERK's discovery, this kinase was implicated in the etiology of the Wolcott-Rallison Syndrome (WRS). This rare autosomal recessive disease is primarily characterized by neonatal insulin-requiring diabetes associated with non-autoimmune-mediated destruction of pancreatic β-cells, followed by skeletal dysplasia, hepatic dysfunction, and growth retardation [44,45]. Genetic studies with tissue-specific and cell-specific knockouts revealed that diabetes, exocrine pancreas dysfunction and skeletal dysplasia are caused by specific loss of PERK in β-cells, acinar cells, and osteoblasts, respectively [46-48].

Correspondingly, mice deficient for PERK have shown similar phenotypes to human WRS patients after birth [49,50]. PERK ablation in β-cells during embryonic development resulted in reduced proliferation and repressed expression of the insulin genes which are key markers of differentiation in these cells [51,52]. This suggests that impaired β-cell differentiation and proliferation during embryonic

development may represent part of the mechanisms by which PERK deficiency leads to neonatal β-cell failure and diabetes in WRS patients [51,53].

Another possible mechanism through which PERK maintains β-cell homeostasis includes its signalling cascade. Mice harbouring a homozygous mutation in Ser51 of eIF2α, inhibiting its phosphorylation, present β-cell deficiency with low production and secretion of insulin, in addition to dying shortly after birth [54]. Additionally, bearing this mutation specifically in β-cells displays increased β-cell death and reduced insulin production [55], suggesting that β-cells require a functional PERK/eIF2α signalling pathway in physiological conditions.

In addition to eIF2α, PERK has another target related to β-cell survival during oxidative stress. The nuclear factor erythroid-derived 2-like 2 (Nrf2) is a transcription factor involved in the upregulation of several antioxidant and detoxifying enzymes required for the biosynthesis and metabolism of glutathione (GSH), a cellular antioxidant [56]. Normally, Nrf2 is constitutively expressed but is inactivated by the Kelch-like ECH-associated protein 1 (Keap1), a component of the ubiquitin ligase complex. During homeostasis, Keap1 anchors Nrf2 in the cytoplasm and favours its degradation by the proteasome [57]. However, during ER stress or oxidative stress conditions, PERK is activated and phosphorylates Nrf2, releasing it from Keap1 and translocating it to the nucleus where it activates transcription of the cytoprotective genes [58]. In β-cells, depletion of this factor results in a reduction of pancreatic islet size [59]. Additionally, suppressed expression of Nrf2 in mouse insulinoma cells increases oxidative stress, with subsequent decreased proliferation and a high rate of apoptosis [60]. These findings support that PERK protects β-cells against oxidative damage through Nrf2 activation.

The fact that several PERK-mediated mechanisms are involved in the development and maintenance of β-cell homeostasis helps explain why loss-of-function mutations in the *EIF2AK3* gene result in the development of WRS.

2. Aims

PERK is a kinase that determines important translational reprogramming during stress conditions, which allows the regulation of stress-related proteins through uORF-mediated mechanisms, while refraining from global protein translation. Its dual facet in determining cell fate during stress is of key importance for cell homeostasis. Therefore, it is not surprising that dysregulated PERK expression is involved in several diseases, specifically the ones accompanied by persistent ER stress conditions. Its complex role in some of these diseases highlights the importance of understanding the mechanisms that mediate PERK expression and the consequences of their dysregulation so PERK can be used as a therapeutic target.

Previously in the laboratory, R. Fernandes (2020) discovered that the *PERK* 5'UTR contains potentially translated uORFs that can repress mORF translation. Furthermore, out of all *PERK*'s uORFs, the first one (uORF1) appeared to be essential for an efficient repression of mORF translation, being a potent inhibitor by itself, and the combined activity of the other uORFs could likely contribute to fine-tuning the outcome of its regulatory role. Not only that, but some described mutations that affect PERK's uORFs were identified in patients with WRS, which, as discussed before, is promoted by insufficient expression/activity of PERK. Of all the mutations studied, the two that were revealed to be more significant both cause some sort of disruption in uORF1. Lastly, R. Fernandes also observed that some factors that contribute to different parts of the translation mechanism — such as eIF1A, eIF5 and eIF5B — can impact and be determinant on PERK's normal expression.

In this work, we aim to study the functional role of uORFs in PERK expression and evaluate their biological impact on cell physiology and human disease. For this purpose, the following goals were established:

I. Determine which are the translation initiation factors involved in the uORF-mediated control of PERK mORF translation;
II. Study the role of ER stress in the uORF-mediated translational regulation of PERK;
III. Determine the consequences of dysregulated PERK expression in cell physiology.

3. Materials and Methods

3.1. *In silico* analysis of *PERK* uORFs

Fernandes, R. *et al.* (2020) observed that PERK contain 8 uORFS in its 5'UTR (Figure 3.1). However, he noticed that uORF1 has its inhibitory effect maximized when expressed in combination with uORF3, uORF4 and uORF8 (this last one seeming to be a key element in the translational repression exerted by the *PERK* 5'UTR) [3], presented in Table 3.1. The sORFs.org database (available at: http://sorfs.org/) was used to identify these uORFs based on experimental data from ribosome profiling studies. uORF sequences were assessed through Benchling (available at: https://www.benchling.com/) and analysed with the online tool SignalP - 5.0 (available at: https://services.healthtech.dtu.dk/services/SignalP-5.0/), to verify if these uORFs encode peptides with a signal peptide (SP).

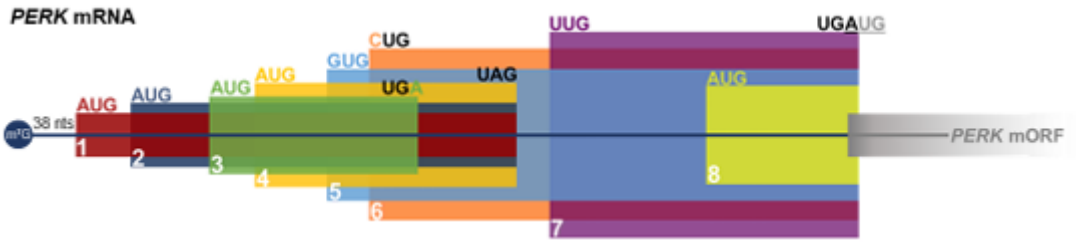


Figure 3.1 - The uORFs in the *PERK* mRNA 5'UTR. Schematic representation of the native arrangement of the five AUG- and three non-AUG-uORFs identified in the 5'UTR of the human PERK mRNA. Each coloured box represents a uORF, numbered from 1 to 8. The coloured letters indicate the start codons of the uORFs, and the black letters represent their stop codons. Image from: Fernandes, R. (2020, November 1). uORF-mediated translational regulation of PERK: implications for cell homeostasis and human disease.

Table 3.1 - Characterization of the uORF1, uORF3, uORF4 and uORF8 identified in the human *PERK* 5'UTR.

uORF ID	uORF nucleotide sequence	uORF length (nts)	Kozak sequence context	Distance to mORF (nts)	Encoded peptide
uORF1	ATGGAGTGGCAGCG GCCGCAGCCAATGA GAGAGCAAACGCGC GGAAAGTTTGCTCAA TGGGCGATGTCCGAG ATAGGCTGTCACTCA GGTGGCAGCGGCAG AGGCCGGGCTGAGA CGTGGCCAGGGGAA CACGGCTGGCTGTCC AGGCCGTCGGGGCG GCAGTAG	165	GGGAACATGG	99	MEWQRPQP MREQTRGK FAQWAMSE IGCHSGGSG RGRAETWP GEHGWLSR PSGRQ
uORF3	ATGGGCGATGTCCGA GATAGGCTGTCACTC AGGTGGCAGCGGCA GAGGCCGGGCTGA	57	TGCTCAATGG	151	MGDVRDRL SLRWQRQR PG

uORF4	ATGTCCGAGATAGGC TGTCACTCAGGTGGC AGCGGCAGAGGCCG GGCTGAGACGTGGCC AGGGGAACACGGCT GGCTGTCCAGGCCGT CGGGGCGGCAGTAG	102	TGGGCGATGT	99	MSEIGCHSG GSGRGRAE TWPGEHGW LSRPSGRQ
uORF8	ATGCCTGCGCGCGGG GCGGGACGCTGA	27	GCCTCCATGC	Overlapped with mAUG	MPARGAGR

In the "Kozak sequence context" column, the start codons are marked in red. When a purine is found at positions -3 and/or +4 relative to the first nucleotide of the start codon, the nucleotide is presented in green.

3.2. *PERK* 5'UTR site-directed mutagenesis and cloning

The pGL2-enhancer plasmid that encodes the firefly luciferase (FLuc) had been previously modified in the laboratory by Fernandes, R. *et al.* (2020) to contain the human cytomegalovirus (hCMV) promoter sequence between BglII and HindIII restriction sites originated from a BglII/HindIII digested plasmid [3]. This reporter vector was named "pGL2" or empty vector. The 302 nt-long 5'UTR of the human *PERK* transcript was inserted into the HindIII/BsrGI restriction sites of the "pGL2" plasmid, forming the "PERK 5'UTR" construct. For this purpose, the Splicing by Overlap Extension Polimerase Chain Reaction (SOEing-PCR) approach was used. Briefly, in the first PCR the 5'UTR of *PERK* was amplified from human genomic DNA using a primer with a HindIII restriction site extension and a primer with a linker for the 5' end of the FLuc ORF. In the second PCR, the 5' end of the FLuc ORF was amplified using a primer with a linker for the 3' end of the *PERK* 5'UTR and a primer located within the FLuc ORF sequence and downstream the BsrGI restriction site, using the "pGL2" vector as template. Then, the two obtained products were joined by the SOEing-PCR using the flanking primers, in a way that the *PERK* 5'UTR was inserted exactly upstream the FLuc start codon. The resulting product and the "pGL2" plasmid were digested with HindIII and BsrGI before being ligated with the T4 DNA ligase to form the "PERK 5'UTR" reporter construct. NZY5α competent cells (NZYTech) were transformed with this ligation reaction and were plated in Luria-Bertani (LB) agar-ampicillin (Sigma Aldrich) plates to select positive colonies. The pDNA was extracted with the NZYMiniprep kit (NZYTech), following the manufacturer's instructions. The sequence of the "PERK 5'UTR" reporter plasmid was confirmed by Sanger sequencing. This construct was subjected to direct or sequential site-directed mutagenesis to obtain several mutant constructs needed to study the translational regulatory function of each uORF, as well as to study the 5'UTR variants of the *PERK* gene listed at the ClinVar database (available at: https://www.ncbi.nlm.nih.gov/clinvar/), obtaining the constructs "c.-172C>T", "c.-195G>T", "c.-201A>G" and "c.-263T>C" [3].

The "PERK 5'UTR" construct was subjected to sequential site-directed mutagenesis to obtain different mutant constructs — in which different tags were inserted to study the peptide function of uORF1. The site-directed mutagenesis reactions were performed using the Q5® Hot Start High-Fidelity DNA Polymerase (New England Biolabs) following the manufacturer's instructions. Thirty ng of the "PERK 5'UTR" plasmid (or a derivative construct) was used as template. PCR cycling was done as follows: initial denaturation at 98 ºC for 30 sec, 25 cycles of 98 ºC for 10 sec, 57 ºC for 30 sec, and 72 ºC for 10 min and a final extension at 72 ºC for 2 min. Then, the PCR product was digested with 1 µl of DpnI (ThermoFisher Scientific) for 1 h at 37 ºC and purified with the DNA Clean & Concentrator™-5 kit

(Zymo Research). NZY5α competent cells were transformed with the purified mutagenesis reaction mixture and the pDNA from overnight cultures of single colonies was extracted as indicated before. All the mutagenesis constructs were confirmed by Sanger sequencing.

To study the uORF1 peptide function and cellular location, site-directed mutagenesis was applied to "PERK 5'UTR" constructs either containing fully functional uORFs (Wild-type) or only one functional uORF (uORF1), intending to insert two different tags (FLAG-tag or Myc-tag) in the uORF1 sequence, right before its stop codon, in frame with the remaining ORF (Figure 3.2). Due to the large size of the tag sequences (24 nucleotides for FLAG and 30 nucleotides for Myc), they are introduced sequentially with the primers listed in Table 3.2. Briefly, for the FLAG-tag, 4 nucleotides are inserted at a time per mutagenesis, and for the Myc-tag, 5 nucleotides are inserted at a time per mutagenesis. This means that 6 different successful mutagenesis per construct are needed to obtain the full tag sequence in the plasmids. All the mutagenesis constructs were confirmed by Sanger sequencing using the primers listed in Table 3.3

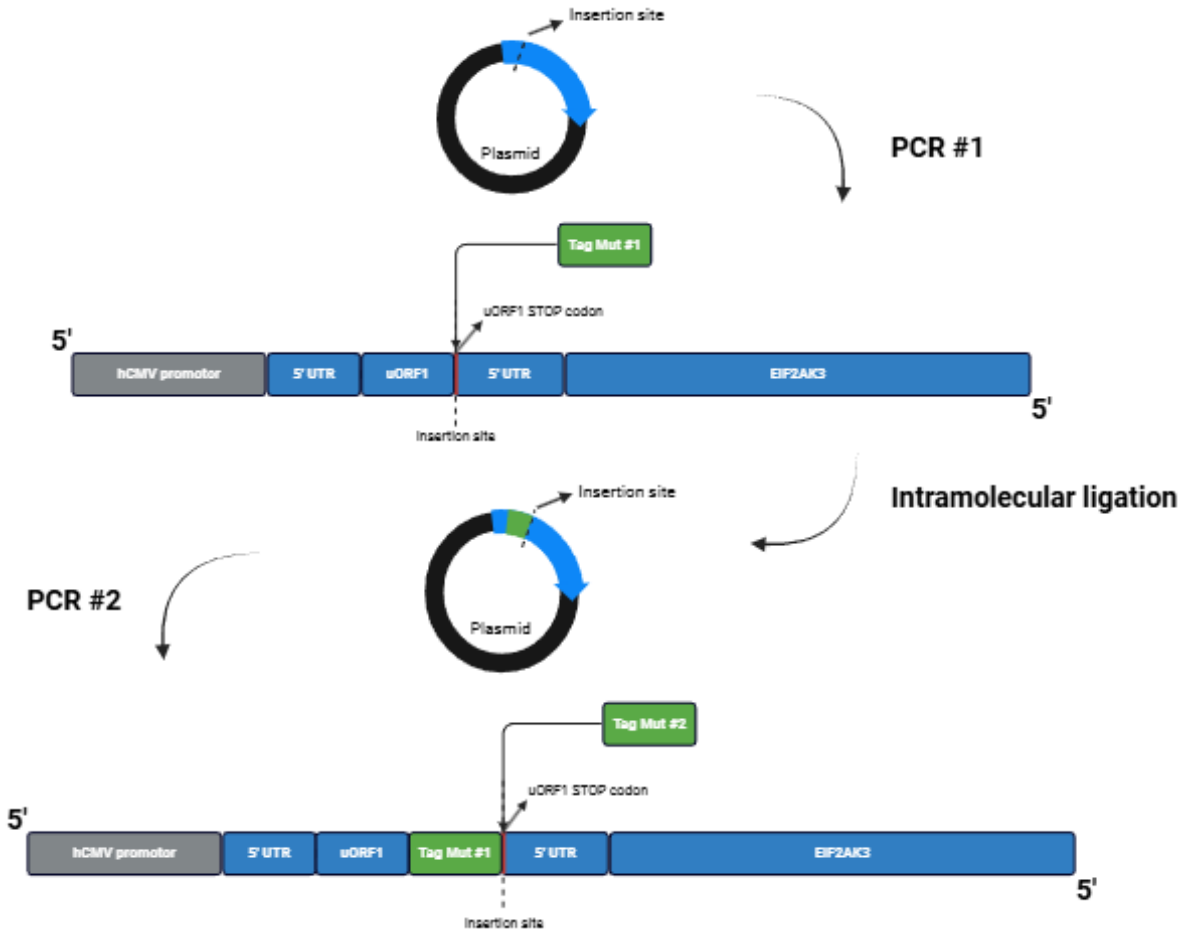


Figure 3.2 - Sequential site-directed mutagenesis. Due to the large size of the tag sequences, a sequential method was used. For each tag, a small number of nucleotides were inserted at a time per mutagenesis. 6 different successful mutagenesis per construct are needed to obtain the full tag sequence in the plasmids. Image created through BioRender (https://www.biorender.com/).

Table 3.2 - List of primers used for site-directed mutagenesis.

Primer	Sequence (5' ⇨ 3')	Function
#1	CAGGCCGTCGGGGCGGCAGGACTTAGGGTCCCTAGCACGTC	**Mutagenesis** Tag insertion in PERK's 5'UTR uORF1: **- Flag tag**
#2	GACGTGCTAGGGACCCTAAGTCCTGCCGCCCCGACGGCCTG	
#3	CCGTCGGGGCGGCAGGACTACAATAGGGTCCCTAGCACGTCC	
#4	GGACGTGCTAGGGACCCTATTGTAGTCCTGCCGCCCCGACGG	
#5	CGGGGCGGCAGGACTACAAGGACTAGGGTCCCTAGCACGTCC	
#6	GGACGTGCTAGGGACCCTAGTCCTTGTAGTCCTGCCGCCCCG	
#7	GCGGCAGGACTACAAGGACGACGTAGGGTCCCTAGCACGTCC	
#8	GGACGTGCTAGGGACCCTACGTCGTCCTTGTAGTCCTGCCGC	
#9	GCAGGACTACAAGGACGACGATGATAGGGTCCCTAGCACGTCC	
#10	GGACGTGCTAGGGACCCTATCATCGTCGTCCTTGTAGTCCTGC	
#11	GACTACAAGGACGACGATGACAAGTAGGGTCCCTAGCACGTCC	
#12	GGACGTGCTAGGGACCCTACTTGTCATCGTCGTCCTTGTAGTC	
#13	CAGGCCGTCGGGGCGGCAGGAACATAGGGTCCCTAGCACGTC	**Mutagenesis** Tag insertion in PERK's 5'UTR uORF1: **- Myc tag**
#14	GACGTGCTAGGGACCCTATGTTCCTGCCGCCCCGACGGCCTG	
#15	CGTCGGGGCGGCAGGAACAAAAACTAGGGTCCCTAGCACGTCC	
#16	GGACGTGCTAGGGACCCTAGTTTTTGTTCCTGCCGCCCCGACG	
#17	GGGCGGCAGGAACAAAAACTCATCTAGGGTCCCTAGCACGTCC	
#18	GGACGTGCTAGGGACCCTAGATGAGTTTTTGTTCCTGCCGCCC	
#19	GCAGGAACAAAAACTCATCTCAGATAGGGTCCCTAGCACGTCC	

#20	GGACGTGCTAGGGACCCTATCTGAGATGAGTTTTTGTTCCTGC	
#21	GAACAAAAACTCATCTCAGAAGAGGTAGGGTCCCTAGCACGTCC	
#22	GGACGTGCTAGGGACCCTACCTCTTCTGAGATGAGTTTTTGTTC	
#23	CTCATCTCAGAAGAGGATCTGTAGGGTCCCTAGCACG	
#24	CGTGCTAGGGACCCTACAGATCCTCTTCTGAGATGAG	

Table 3.3 - List of primers used for sequencing.

Primer	Sequence (5' ⇨ 3')
#1	CACCAAAATCAACGGGACTT
#2	GCCTCAGCGTCCTCCACAGT
#3	CGGATCCCACATCCAAATCCCACTG
#4	ACTCAGTGCATATAGTGGAAAGGTG
#5	TCTTCATCTTCTAAAACATCATCAT
#6	CAGAATTTCAGAAAAGTTTCCTTCA
#7	TTCACCACTTACAGAATCATATTTA
#8	TAAAGCCTTAGCCAAGCTTGAACAC
#9	AAAGGAGTGCCCCTCATCATTGCCA
#10	CTTCCAGTAAAGAAGAGCCGAAAAC
#11	GCCCAGAGCAGATTCATGGAAACAG
#12	TAGAATCAGGCCTAAAGAAAAGATG

#13	ACGGCGTGGAGTTCGAGCTG
#14	CTTCACCGGCATCTGCATCC

3.3. Cell culture

HCT116 (human pre-metastatic colorectal carcinoma-derived cell line) cells were cultured in T75 cell culture flasks containing Gibco McCoy's 5A Medium supplemented with 10% (v/v) fetal bovine serum at 37 ºC in a humidified 5% carbon dioxide (CO_2) incubator.

3.4. siRNA and plasmid DNA transfection

HCT116 cells were seeded at 20–30% confluence in 6-well plates and incubated for 24 h. Then, the cells were transiently transfected with 250 pmol of siRNAs (listed in Table 3.4) diluted in 250 µl of Opti-MEM medium. A scramble siRNA was used as a negative control. Transient transfections were performed using Lipofectamine®2000 (Lipofectamine) transfection reagent, according to the manufacturer's instructions. Per well, 4 µl of Lipofectamine were diluted in 250 µl of Opti-MEM medium and left to rest for 5 min. Both solutions were mixed and left to rest for 20-30 min at room temperature. Meanwhile, the old culture medium present in the cell culture dishes was replaced by a fresh medium. Cells were transfected dropwise and incubated at 37 °C in a humidified 5% CO_2 incubator.

Twenty-four hours after siRNA transfection (48 h after initial seeding), cells were transiently co-transfected with 1000 ng of "PERK 5'UTR" plasmid expressing firefly luciferase (FLuc) — either containing all of PERK's uORFs, or none — and with 500 ng of "pRL-TK", a reporter vector expressing Renilla luciferase (RLuc) used as a transfection efficiency control. Transient transfections were performed using Lipofectamine®2000 transfection reagent according to the manufacturer's instructions. Per well, 4 µl of Lipofectamine were diluted in 250 µl of Opti-MEM medium and left to rest for 5 min. The corresponding amount of each reporter plasmid was also diluted in 250 µl of OptiMEM. Both solutions were mixed and left to rest for 20-30 minutes at room temperature. Meanwhile, the old culture medium present in the cell culture dishes was replaced by a fresh medium. Cells were transfected dropwise and incubated at 37 °C in a humidified 5% CO_2 incubator. Twenty-four hours after plasmid transfection, the cells were harvested.

Table 3.4 – List of siRNAs used in the present work.

Target gene	Sequence (5' ⇨ 3')	Manufacturer
Luciferase	CGUACGCGGAAUACUUCGA	Thermo Scientific
DENR	CAAGUUAGAUGCCGAUUAC	Invitrogen
eIF4G2/DAP5	AAUGUGGGUGUAGAGUCUAAA	Invitrogen

eIF3D	AUCGAAUGAGAUUUGCCC	Thermo Scientific
eIF3H	CCCAAGGAUCUCUCUCACUAA	Invitrogen
eIF2D		Dharmacon
eIF2AK3/PERK	GAAGCUACAUUGUCUAUUU UAGCAAAUCUUCUUCUGAA UAAACUAACUGCUUUCAAG ACUAAUCGAUUGCAUAUUG	Dharmacon

3.5. Cell lysis

Cells were washed with 1 mL of pre-chilled 1× (v/v) phosphate-buffered saline (PBS) and lysed with either 80 µl of 1x (v/v) Passive Lysis Buffer, 80 µl of 1× (v/v) sodium dodecyl sulfate (SDS) lysis buffer, or 80 µl of a 1x (v/v) mix containing Nonidet P-40 (NP-40), protease and RNAse inhibitors. Total lysates were frozen at -80 ºC for at least 2 hours before use.

3.6. Luminometry assays

Relative Firefly Luciferase (FLuc) and Renilla Luciferase (RLuc) activities were assessed using the Dual-Luciferase® Reporter Assay System (Promega), according to the manufacturer's instructions, on a GloMax® 96 Microplate Luminometer. Briefly, 10 µl of the cleared cell lysates were plated in a white, opaque 96-well plate. First, the luminescence signal of the FLuc reporter was measured by adding 40 µl of the Luciferase Assay Reagent — containing FLuc substrate — to each sample. Then, RLuc reporter luminescence was sequentially quantified by adding 40 µl of the Stop & Glo® Reagent that stops the reaction between FLuc and its substrate and contains the substrate for the RLuc reaction. The collected data were expressed in arbitrary light units. Relative luciferase activity was obtained by normalizing FLuc to RLuc luminescence for each sample.

3.7. SDS-PAGE and Western blot

Total protein was mixed with 5x (v/v) sodium dodecyl sulphate sample buffer and denatured at 95 ºC for 20 min. The samples were resolved in a 10% acrylamide/Bis (Bio-Rad) gel at 20 mA. Proteins were then transferred to methanol-activated polyvinylidene difluoride (PVDF) membranes at 200 mA. Membranes were blocked with 5% (w/v) non-fat dry milk or bovine serum albumin (BSA), diluted in 1× (v/v) tris-buffered saline (TBS) supplemented with 0.05% (v/v) Tween-20. Membranes were then probed with gentle shaking overnight, at 4 ºC, with the primary antibody solutions listed in Table 3.5. Detection was carried out using secondary peroxidase-conjugated anti-mouse IgG or anti-rabbit IgG diluted 1:4000 or 1:3000, respectively, in 5% (w/v) non-fat milk in 0.05% (v/v) TBS-Tween-20 for 1h at room temperature, with gentle shaking, followed by enhanced chemiluminescence. When necessary,

membranes were stripped off and re-probed with different antibodies. Densitometry analysis of the obtained bands was performed using the ImageJ software.

Table 3.5 – List of primary antibodies and respective conditions used in the present work.

Primary antibody	Manufacturer, catalog nr.	Antibody dilution	Solution	Secondary antibody
Mouse Anti-α-tubulin	Sigma-Aldrich, #103M4773V	1:50000	5% non-fat milk, 0.05% TBS-Tween-20	Anti-mouse
Mouse Anti-DENR	Santa Cruz Biotechnology, #D1720	1:1000	5% non-fat milk, 0.05% TBS-Tween-20	Anti-mouse
Rabbit Anti-eIF4G2	Cell Signalling, #D88BG	1:2000	5% non-fat milk, 0.05% TBS-Tween-20	Anti-rabbit
Mouse Anti-PERK	Santa Cruz Biotechnology, #J1718	1:2000	5% BSA, 0.05% TBS-Tween-20	Anti-mouse
Mouse Anti-eIF2D	Santa Cruz Biotechnology, #K1317	1:1000	5% non-fat milk, 0.05% TBS-Tween-20	Anti-mouse
Rabbit Anti-eIF3H	Cell Signalling, #D9C1	1:2000	5% non-fat milk, 0.05% TBS-Tween-20	Anti-rabbit
Rabbit Anti-eIF2α-P	Invitrogen, #44728G	1:1000	5% BSA, 0.05% TBS-Tween-20	Anti-rabbit
Mouse Anti-eIF2α	Santa Cruz Biotechnology, #E2003	1:2000	5% non-fat milk, 0.05% TBS-Tween-20	Anti-mouse
Mouse Anti-eIF3D		1:2000	5% non-fat milk, 0.05% TBS-Tween-20	Anti-mouse

3.8. RNA extraction and reverse transcription (RT)-PCR

Total RNA isolation of cleared cell lysates was performed with the Nucleospin® RNA kit according to the manufacturer's protocol. cDNA was synthesized using the NZY M-MuLV Reverse Transcriptase as instructed by the manufacturer. Briefly, 1000 ng of total RNA was added to 10 µl of NZYRT 2× Master Mix containing dNTPs (10 mM) and an optimized RT buffer; 2 µl of NZYM-MuLV RT Enzyme Mix containing NZY M-MuLV Reverse Transcriptase and NZY Ribonuclease Inhibitor; 250 ng of random hexamers and RNAse free ddH_2O water to a final volume of 20 µl. The mixture was incubated in a thermocycler at 25 °C for 10 min, 42 °C for 50 min and 85 °C for 5 min. The cDNA was then stored at -20 ºC for future usage.

3.9. Semi-quantitative RT-PCR

A semi-quantitative reverse transcription-PCR (RT-PCR) was performed using the AmpliTaq® DNA polymerase following the manufacturer's protocol. The amplification reaction was prepared through two mixtures: the first containing 5 µl PCR buffer II (10×; AB), 2 mM of MgCl2, 1 µl of 10 mM of dNTPs mix, 2 µl of primer forward (10 µM) and primer reverse (10 µM) for GAPDH, MCTS-1 and 5MP cDNA amplification (primers #1 and #2; #3 and #4; #5 and #6, respectively, in Table 3.6), 0.3 µl of AmpliTaq® DNA polymerase (5 U/µl), and 3 µl of dimethyl sulfoxide (DMSO). The second mix consisted of three different cDNA dilutions (2:1, 1:1, and 1:2) containing 360 ng cDNA, 180 ng cDNA and 90 ng cDNA respectively, along with nuclease-free water until 39.7 µl. Both mixtures were then joined and proceeded to the thermal cycling, following these conditions:

- For GAPDH: initial denaturation at 95 ºC for 5 min; 28 cycles of 95 ºC for 30 s (denaturation), 57 ºC for 45 s (annealing), and 72 ºC for 1 min (extension); and then a final extension at 72 ºC for 10 min.
- For MCTS-1: initial denaturation at 95 ºC for 5 min; 30 cycles of 95 ºC for 30 s (denaturation), 59 ºC for 45 s (annealing), and 72 ºC for 1 min (extension); and then a final extension at 72 ºC for 10 min.
- For 5MP: initial denaturation at 95 ºC for 5 min; 27 cycles of 95 ºC for 30 s (denaturation), 55 ºC for 45 s (annealing), and 72 ºC for 1 min (extension); and then a final extension at 72 ºC for 10 min.

After amplification, 20 µL of the PCR products were resolved by electrophoresis in a 2% (w/v) agarose gel stained with ethidium bromide. The density of the bands was quantified using ImageJ software, and the mRNA levels were determined using a standard calibration curve performed with serial dilutions of the cDNA samples.

Table 3.6 – List of primers used in the semi-quantitative RT-PCR.

Primer	Target gene	Orientation	Sequence (5' ⇨ 3')
#1	GAPDH	Forward	CCATGAGAAGTATGACAACAGCC
#2		Reverse	GGGTGCTAAGCAGTTGGT
#3	MCTS-1	Forward	TCCTTTTATCCTGCCACACCAG
#4		Reverse	GTTGACTTTCTCAATGTCTTCTGCA
#5	5MP	Forward	CGATGCTGTCGGGGATTCTG
#6		Reverse	GGCCGCAATGCCTTCTTTG

3.10. *PERK* knockdown and thapsigargin treatment

HCT116 cells were seeded at 20–30% confluence in 6-well plates and incubated for 24 hours. Then, the cells were transiently transfected as described above (siRNA and plasmid DNA transfection section), using 250 pmol of siRNA against PERK. A scramble siRNA was used as a negative control. Twenty-four hours later, the cells were transiently transfected with different amounts of "PERK 5'UTR" plasmid

expressing *EIF2AK3* (PERK), or its derivative constructs. Transient transfections were performed as described above (plasmid DNA transfection section). Four hours after transfection, the cells were incubated with 1 µM of thapsigargin (Tg) or DMSO, diluted in fresh medium for 24 hours before being harvested.

3.11. Statistical analysis

All the above experimental analyses were performed in a minimum of three independent experiments. The results are expressed as mean ± standard error of the mean (SEM). Student's t-test, two-tailed and unpaired, was applied for statistical significance. Significance for statistical analysis was defined as $p<0.05$ (*), $p<0.01$ (**) and $p<0.001$ (***).

4. Results

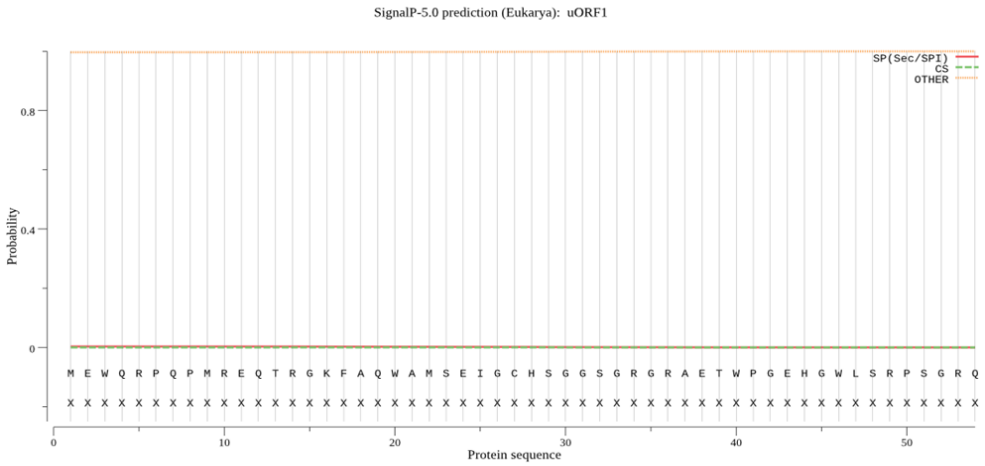
4.1. *PERK* uORFs probably contain no signal peptide

Signal peptides are short peptides located in the N-terminal of proteins, carrying information for protein secretion and targeting proteins to the extracellular environment through the endoplasmic reticulum in eukaryotic cells [61]. Different secretory pathways are determined by the signal peptides (SP's) features, affecting protein localization in the cell as well as post/co-translational modification of protein. These have evolved among organisms as ways to offset the low rate of protein synthesis *versus* the high rate of secretion; and to be able to secrete proteins with different characteristics [62,63].

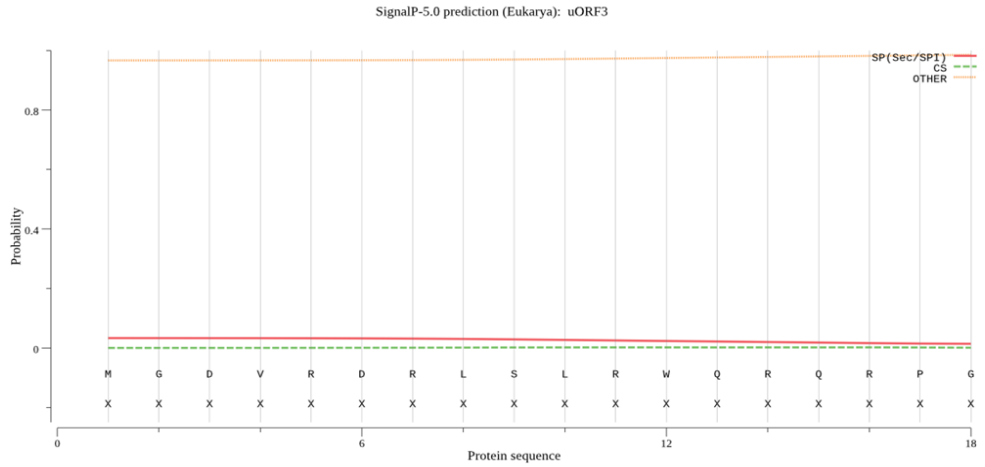
However, SPs are of special importance for aspects other than the production of recombinant proteins. A significant amount of eukaryotic secretory proteins (~20%) possess SPs [64], and many human diseases are caused by mutations in them. As an example, the mutation in the SP of preproinsulin is associated with the onset of diabetes [65].

According to this information, we proposed the possibility that some of *PERK*'s upstream open reading frames (uORF1, uORF3, uORF4 and uORF8) would encode novel peptides with a functional independency of that one of PERK main ORF. Specifically, to determine if these uORFs are encoding SP-harbouring sequences, we used the SignalP - 5.0 online tool. This analysis showed that none of the tested uORFs has a high predicted score — 0.32%, 3.4%, 1.09% and 7.7%, respectively (Figure 4.1). These results suggest that uORFs 1, 3, 4 and 8 do not code for proteins containing signal peptides.

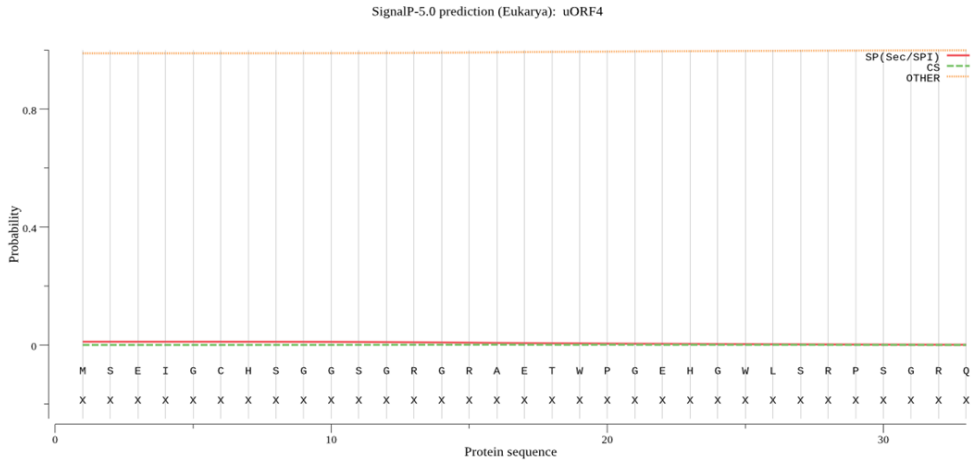
A



B



C



D

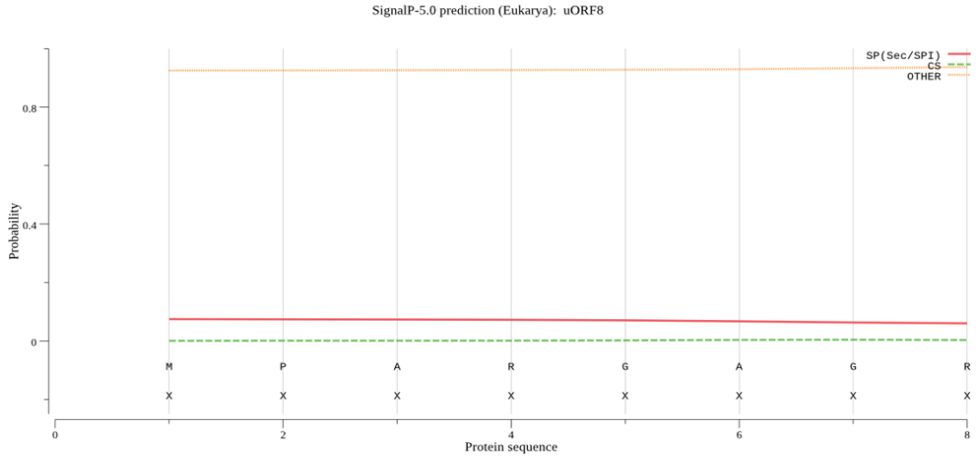


Figure 4.1 - uORFs 1, 3, 4 and 8 do not code for proteins containing signal peptides. The sORFs.org database was used to identify uORFs 1, 3, 4 and 8 based on experimental data from ribosome profiling studies. The uORF sequences were assessed through Benchling and analysed with the online tool SignalP - 5.0. Representative images show the probability of each uORF to contain a signal peptide sequence **(A)** uORF1 has a predicted score of 0,32%, **(B)** uORF3 has a predicted score of 3,4%, **(C)** uORF4 has a predicted score of 1,09% and **(D)** uORF8 has a predicted score of 7,7%.

4.2. eIF2D and eIF4G2 may be required for regulation of PERK normal expression

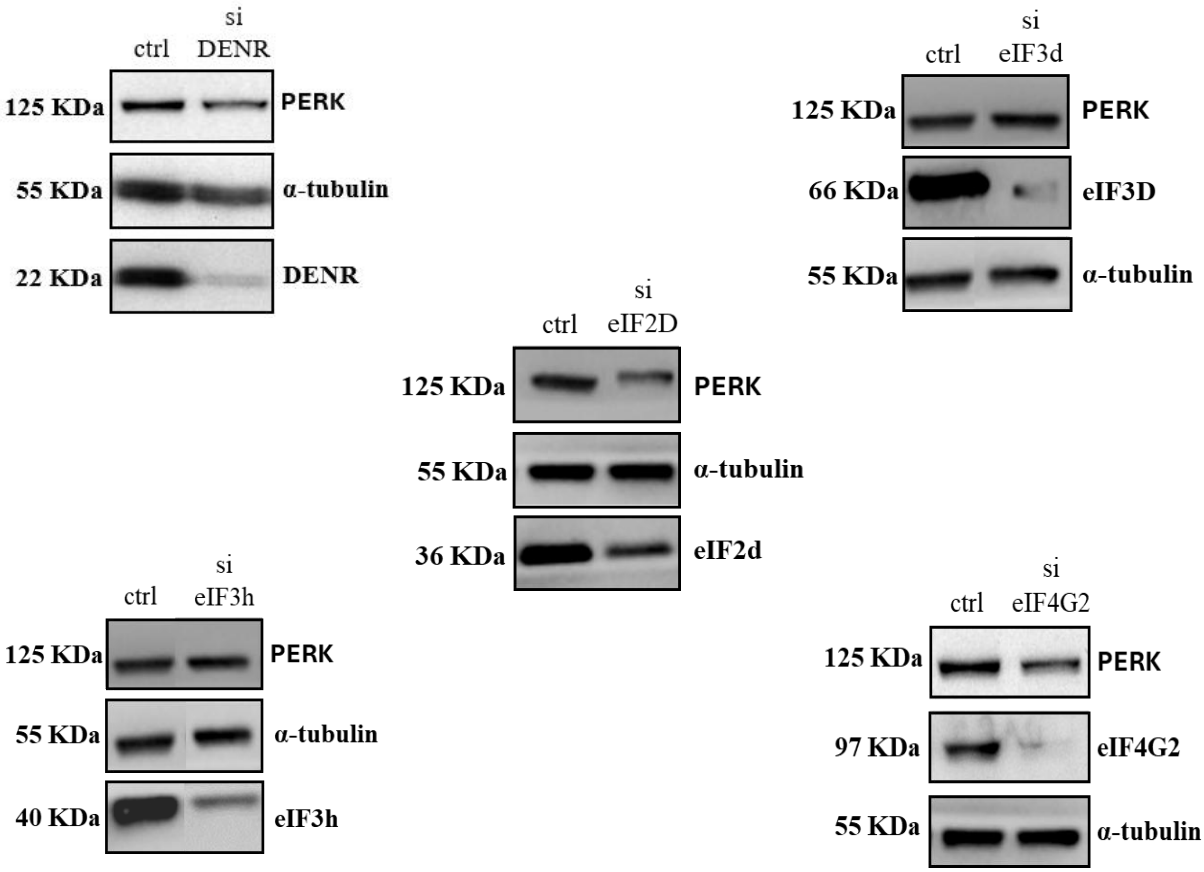
The best studied mRNA regulated by translation reinitiation in animals is *Activating Transcription Factor 4 (ATF4)*, the main downstream effector of the integrated stress response [66]. It has been reported that the DENR•MCTS1 complex and eIF2D factor are required for the translation of *ATF4* and other oncogenes by promoting translation reinitiation [67]. Bohlen *et al.* have found that the knockdown of DENR leads to strongly reduced ATF4 protein levels, both in the absence and in the presence of stress, but not reduced *ATF4* mRNA levels, consistent with impaired *ATF4* translation. They also noticed that DENR knockout (KO) cells still expressed eIF2D — a protein containing the functional domains of both DENR and MCTS1 combined. Knockdown of eIF2D in the $DENR^{KO}$ cells further reduced ATF4 protein levels and strongly impaired the stress-inducibility of ATF4. Knockdown of DENR and/or eIF2D showed reduced ATF4 protein levels in HeLa cells and HT1080 fibrosarcoma cells, indicating lack of cell specificity and meaning that, if these proteins turned out essential for maintaining PERK protein levels, similar observations should be noted in HCT116 cells used in our study.

As eIF4G2 has been assumed to participate in noncanonical translation initiation mechanisms and scanning [68], we included this factor in our study. Under normal conditions, eIF4G2 provides leaky scanning through some upstream open reading frames (the translation of which was canonical cap- and eIF4F-dependent), which are typical for long 5′UTRs and highly frequent uORFs. Shestakova *et al.* affirm that it is likely that the protein can also help the ribosome to overcome other impediments during scanning of the 5′-UTRs of animal mRNAs, explaining the need for eIF4G2 in higher eukaryotes, as many mRNAs that encode regulatory proteins have rather long and highly structured 5′UTRs. Additionally, they can often bind to various proteins, which also hamper the movement of scanning ribosomes. Recently, Schneider's group showed that eIF4G2 is complexed with eIF3d to promote non-canonical, cap-dependent, eIF4E-independent translation of approximately 20% of mRNAs in mammalian cells [69]. Genome-wide transcriptomic and translatomic analyses demonstrated that highly eIF4G2-dependent mRNAs were enriched in cell survival, motility, DNA repair and translation initiation pathways [70]. It would be expected that, in normal conditions, the absence of eIF4G2 would impact negatively PERK protein expression due to its role in mediating leaky scanning mechanisms through uORFs. The same is hypothetized in stress conditions since, with no eIF4G2 and eIF3d to promote non-canonical translation initiation, it would suffer drawbacks, affecting PERK protein expression.

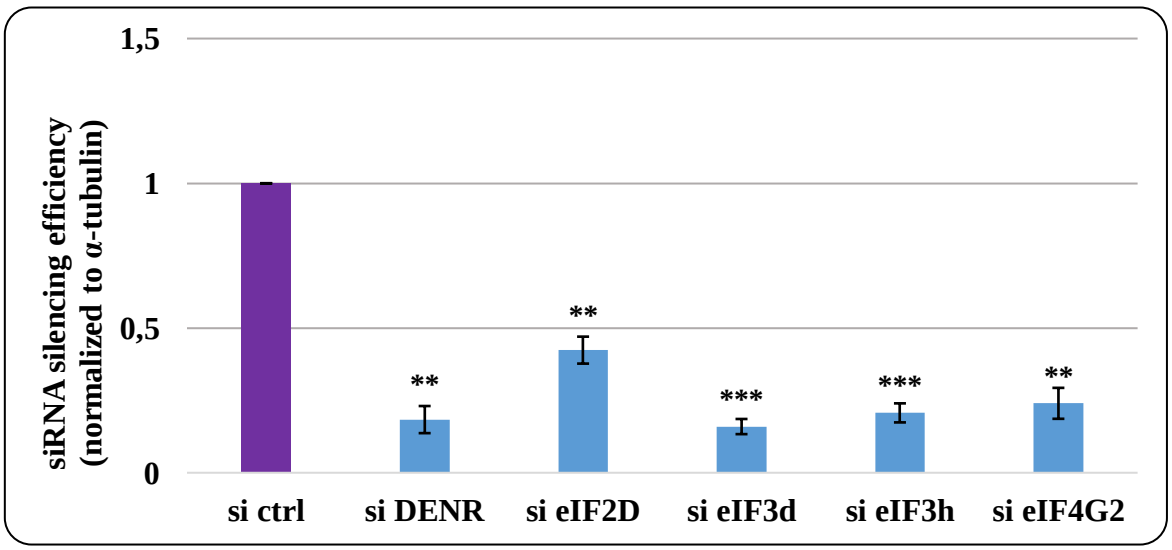
Lastly, it has been shown that eIF3h supports efficient reinitiation after uORF translation, ensuring that a fraction of uORF-translating ribosomes retain their competence to resume scanning [71]. Roy *et al.* caused deficiencies in expression of this factor that could be observed on a panel of mutant 5′ leader sequences derived from their model system, the 5′ leader of *Arabidopsis AtbZip11*, and observed through transient expression assays (using FLuc as a reporter) that the cluster of four uORFs in the *AtbZip11* 5′ leader repressed expression by about twofold in the wild type but by about 10-fold in eIF3h mutant seedlings. The dependence on eIF3h was attributed primarily to direct translational inhibition, rather than reduced abundance or decreased stability of the *AtbZip11*-FLuc mRNA. They obtained similar results using the transformation of Arabidopsis seedlings with expression plasmids. To directly test whether eIF3h is responsible for reinitiation after uORF translation, they altered the uORFs start codons to a strong Kozak context to minimize leaky scanning and tested reinitiation events through spacers between the uORF stop codon and the main ORF's start codon (it is generally poor if the distance between the uORF stop codon and the main ORF's start codon is too short). They demonstrated that reinitiation in the wild type had reached a plateau with a spacer length of 50 nt, whereas the eIF3h mutant showed only marginal expression, regardless of the spacer length, indicating that eIF3h is specifically required for reinitiation downstream of one uORF to another.

Given this, we wanted to test if some of these factors (DENR, eIF2D, eIF3d, eIF3h and eIF4G2) could be involved in uORF-mediated mechanisms of PERK translation. For this purpose, HCT116 cells were transiently transfected with siRNAs against each of the mentioned eIFs, or against luciferase, as control. Twenty-four hours post-transfection, cells were harvested, and PERK protein expression levels were assessed by Western blot. The knockdown efficiency of each eIF was also monitored by Western blot (Figure 4.2).

A


si
ctrl DENR
125 KDa PERK
55 KDa α-tubulin
22 KDa DENR
si
ctrl eIF3d
125 KDa PERK
66 KDa eIF3D
55 KDa α-tubulin
si
ctrl eIF2D
125 KDa PERK
55 KDa α-tubulin
36 KDa eIF2d
si
ctrl eIF3h
125 KDa PERK
55 KDa α-tubulin
40 KDa eIF3h
si
ctrl eIF4G2
125 KDa PERK
97 KDa eIF4G2
55 KDa α-tubulin

B


siRNA silencing efficiency (normalized to α-tubulin)
1,5
1
0,5
0
**
**
***
***
**
si ctrl
si DENR
si eIF2D
si eIF3d
si eIF3h
si eIF4G2

C

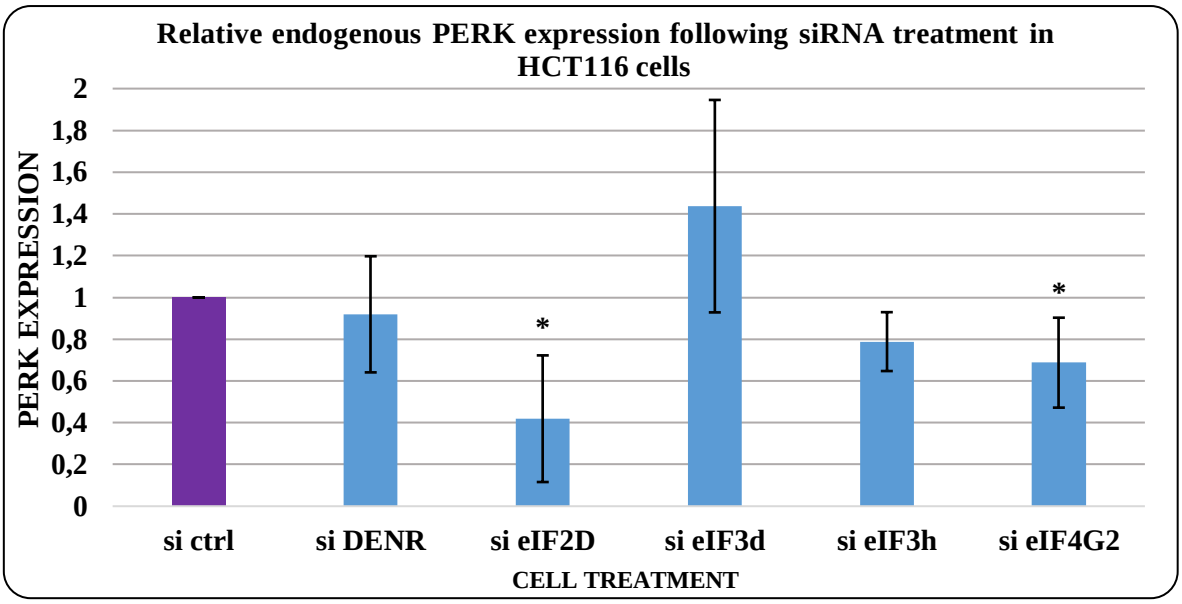


Figure 4.2 - Depletion of DENR, eIF2D, eIF3d, eIF3h and eIF4G2 affect PERK expression at different levels. (A) HCT116 cells were transiently transfected with a control siRNA (siLuc) or specific siRNA pools targeting DENR (siDENR), eIF2D (sieIF2D), eIF3d (sieIF3d), eIF3h (sieIF3h) and eIF4G2 (sieIF4G2), and incubated for 48h before being harvested. Then, 35 µl of total protein were resolved in a 10% SDS-PAGE before performing immunoblotting. Representative images are shown of immunoblots probing for PERK, DENR, eIF2D, eIF3d, eIF3h, eIF4G2, eIF2α-P, or α-tubulin (internal control). **(B)** Quantitation of DENR, eIF2D, eIF3d, eIF3h and eIF4G2, each normalized to α-tubulin and compared to si ctrl, to assess knockdown efficiency. **(C)** The results are expressed as relative PERK expression determined by normalizing PERK activity to that of α-tubulin and then by normalizing the PERK/ α-tubulin ratio from each transfected siRNA to that of the "si ctrl", arbitrarily defined as 1. Data are expressed as mean ± SEM from 3 independent experiments. Statistical significance was determined with two-tailed and unpaired Student's t-test: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$.

According to our results (Figure 4.2 C), the significant reduction in DENR levels had no impact on PERK protein, suggesting that, contrary to what was observed for ATF4, this eIF is not determinant for PERK translation. Regarding eIF2D, its depletion significantly decreased PERK expression, consistent with the existence of eIF2D mediated regulatory mechanisms. Accordingly, a substantial decrease in eIF4G2 induced a significant reduction in PERK, suggesting that loss of eIF4G2 may impair *PERK* translation. In contrast, eIF3d and eIF3h depletion did not significantly affect PERK expression.

In summary, while DENR, eIF3d and eIF3h do not appear to have any regulatory role over PERK expression, tests are still needed in order to check if they do not affect the normal steady-state levels of *PERK* mRNA. On the other hand, eIF2D and eIF4G2 are likely to be involved in the regulation of PERK protein levels. Whether this is related to translation (and to PERK uORFs) or to PERK turnover, remains to be tested. It is also necessary to test if the knockdown of these proteins affects *PERK* mRNA levels.

To further analyse whether the presence of *PERK*'s uORFs control mORF expression levels through a mechanism that involves translation initiation factors DENR, eIF2D, eIF3d, eIF3h and/or eIF4G2, we transiently transfected HCT116 cells with siRNAs against each of the mentioned eIFs, or a scramble siRNA, as control. Twenty-four hours post-transfection, we transiently co-transfected these cells with plasmids expressing FLuc, as a reporter, under the control of the PERK 5'UTR either containing all of PERK's uORFs (WT), or none (no_uORFs). The next day, cells were harvested, and FLuc expression was assessed through luminometry assays using the Dual-Luciferase® Reporter Assay System (Figure 4.3).

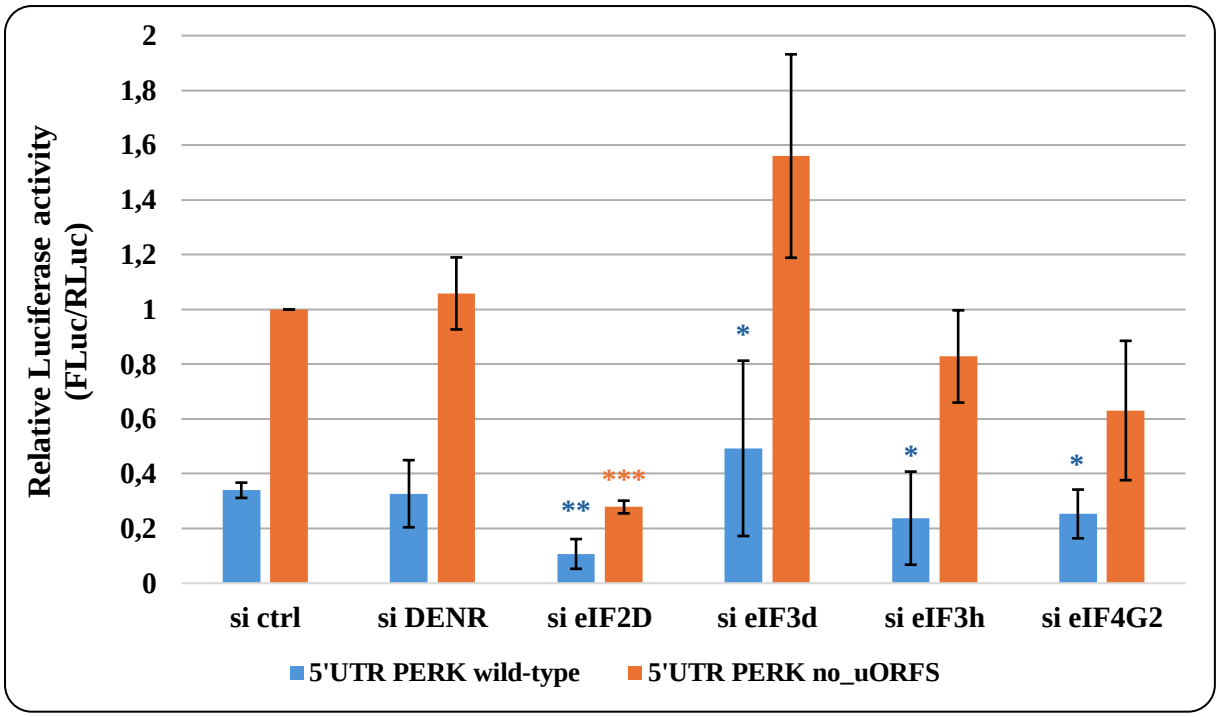


Figure 4.3 - The function of *PERK* uORFs on translational control does not seem to depend on the DENR factor, but seems to involve eIF2D, eIF3d, eIF3h and eIF4G2 factors. HCT116 cells were transiently transfected with a control siRNA (si scramble) or specific siRNA pools targeting DENR (siDENR), eIF2D (sieIF2D), eIF3d (sieIF3d), eIF3H (sieIF3h) and eIF4G2 (sieIF4G2). Twenty-four hours after siRNA transfection, HCT116 cells were transiently co-transfected with plasmids expressing FLuc, as a reporter, under the control of the 5'UTR of PERK, either containing all uORFs (WT, in blue), or none (no uORFs, in orange). The next day, cells were harvested, and FLuc expression was assessed through luminometry assays using the Dual-Luciferase® Reporter Assay System. The results are expressed as relative luciferase activity determined by normalizing FLuc activity to that of RLuc and then by normalizing the FLuc/RLuc ratio from each transfected plasmid to that of the "no_uORFs" construct, arbitrarily defined as 1. Data are expressed as mean ± SEM from 3 independent experiments. Statistical significance was determined with two-tailed and unpaired Student's t-test: * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$. Statistical significance presented in blue is applied in comparison to wild-type control and statistical significance presented in orange is applied in comparison to no_uORFs control.

Our results show that the presence of *PERK* uORFs is needed to refrain protein expression of the downstream main ORF. Interestingly, when compared to what was observed in Figure 4.2 C, the depletion of the factors studied seem to follow a similar trend. However, depletion of eIF3d and eIF3h caused statistically significant effects in mORF activity levels when expressed under the presence of *PERK* uORFs — indicating that these factors may also be involved in the regulation of mORF expression. It is possible that the high band saturation obtained (Figure 4.2 A) caused quantification issues leading to the increased standard deviation values and ultimately to a small discrepancy of values between the two graphs (presented in Figure 4.2 C and 4.3). Additionally, it is also observable that the depletion of eIF2D significantly decreased mORF activity when expressed in the absence of *PERK*'s uORFs, supporting its role in noncanonical translation mechanisms.

4.3. Dysregulated PERK expression and possible relation with WRS

R. Fernandes has shown that *PERK* uORFs are major regulators of mORF translation during basal and stress conditions [3]. This, allied to the fact that PERK is a key protein in the regulation of cell homeostasis with implications for human disease, raised questions about the biological consequences of its dysregulated expression. To address this question, we intended to test the impact of *PERK* 5'UTR mutations that are identified within patients with Wolcott-Rallyson syndrome, which, as discussed before, is promoted by insufficient expression of PERK. The first mutation, c.-195G>T, introduces a

stop codon in-frame with uORF1 start codon, shortening it from 165 to 72 nts. The second mutation, c.-263T>C, is a start codon-disruptive alteration that eliminates uORF1.

To assess the impact of these genetic variants on the uORF-mediated regulation of PERK translation, we cloned these PERK 5'UTR variants in a plasmid expressing the PERK mORF, allowing us to further analyse the effect of these mutations in its expression. The constructs were obtained as described in section 3.2 of Materials and Methods. Next, we intended to test the effect of these mutations on the expression of PERK mORF. For that, we pretended to transfect the plasmids containing these mutations in HCT116 cells — under both normal conditions and under endoplasmic reticulum stress, induced by thapsigargin cellular treatment. The results obtained would be then compared to those obtained in cells expressing the wild-type variant of *PERK* 5'UTR, containing all its functioning uORFs. In order to test how the cell line used would react in normal and stress conditions, in the presence and absence of endogenous PERK, HCT116 cells were transiently transfected with a siRNA against PERK or with the control siLuc. Forty-eight hours post transfection, cells were treated with 1 µM of Tg or DMSO for more 24 h, before lysis. PERK knockdown and eIF2α-P levels were monitored by Western blot. PERK was efficiently knocked-down, and as expected, Tg treatment increased the levels of eIF2α-P (Figure 4.4). Due to the lack of time, this experiment did not advance further than this point.

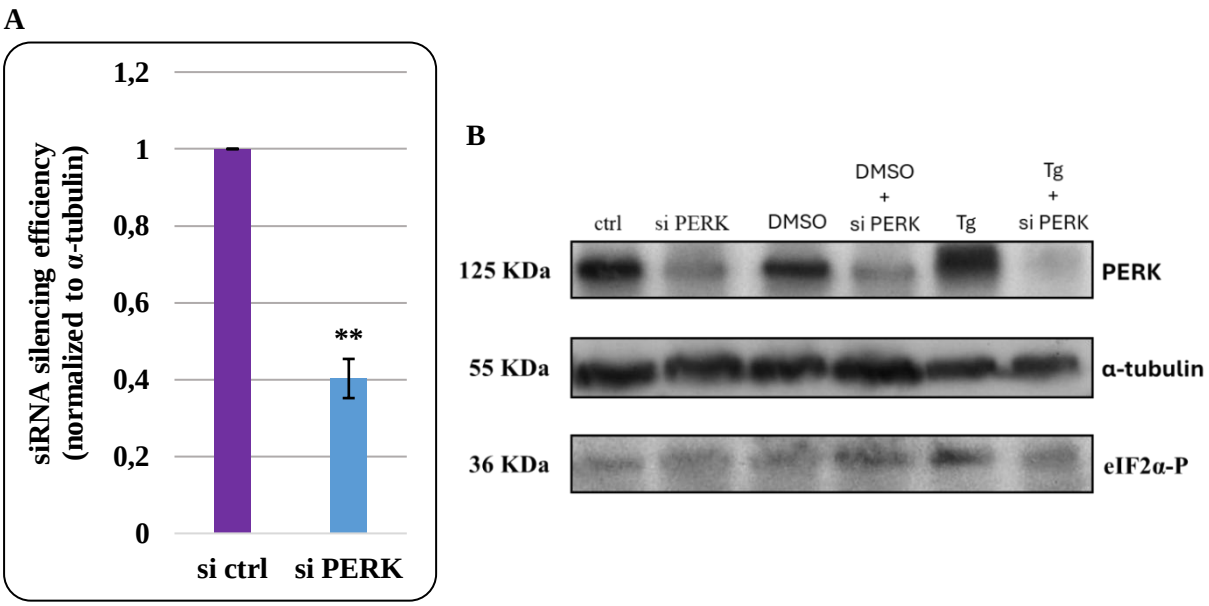


Figure 4.4 - Impact of PERK knockdown and thapsigargin treatment in HCT116 cells. HCT116 cells were transiently transfected with a siRNA against PERK (siPERK), or with a control siRNA (siLuc). Forty-eight hours post-transfection, cells were treated with 1 µM of Tg or DMSO for 24h, before being harvested. **(A)** Quantitation of PERK (*EIF2AK3*), normalized to α-tubulin, to assess knockdown efficiency. **(B)** 35 µl of total protein were resolved in a 10% SDS-PAGE gel prior to performing immunoblotting. Representative images are shown of immunoblots probing for PERK, eIF2α-P, or α-Tubulin (loading control). Statistical significance was determined with two-tailed and unpaired Student's t-test: ** $p < 0.01$.

It is expected that the c. 195G>T mutation, and the consequent shortening of uORF1, will induce PERK expression through compromise of its inhibitory capacity over mORF, promoting translation reinitiation at mORF. Additionally, the c.-263T>C mutation and the consequent disruption of uORF1 start codon is also expected to promote an increase in PERK expression. Given that these uORF-altering mutations are expected to increase mORF expression, it is, by principle, difficult to co-relate them with the pathogenesis of WRS, as this disease is intimately related to the insufficient expression of PERK. Therefore, it would be necessary additional clinical data about the patients, to relate these alterations to this disease.

5. Discussion and Future Perspectives

mRNA translation is tightly controlled during its initiation step by *cis*-acting regulatory elements within the 5'UTR sequence of the transcripts [1]. Out of these elements, we highlight the uORFs, which are potentially translated small ORFs present in about half of the human transcriptome [14]. uORFs are typically described as repressors of the downstream mORF translation and have been frequently described as dynamic molecular switches, as they drive important translational regulatory mechanisms of mORF expression according to the cellular environment [13, 14, 16, 17]. However, translation of uORFs can potentiate mORF translation during stress, to allow the translation of stress-responsive transcripts that can cope with stress [1]. As demonstrated for the uORF-harbouring transcript encoding ATF4, while its uORFs can repress translation of the respective mORF during basal conditions, the same uORFs can promote mORF translation during stress conditions, as part of the stress-responses [22, 72]. In this work we study an example of a transcript containing uORFs that function as molecular switches, as well as translational buffers, differentially fine-tuning the expression of mORF in basal and stress conditions.

The human *PERK* mRNA contains at least eight uORFs with a complex arrangement within the 5'UTR (Figure 3.1). Data analysis from Ribo-seq studies revealed the occurrence of initiation events upstream of the *PERK* mORF, suggesting that these uORFs could be translated and play a role in regulating mORF translation [3]. With this in mind, Fernandes, R. *et al.* (2020) evaluated the translational power of *PERK* uORFs and their role in the regulation of mORF translation. Indeed, he found that *PERK* 5'UTR is capable of significantly repressing translation of the downstream mORF in two different cell lines; the uORFs in the 5'UTR promoted an 89% reduction in mORF expression in HEK293T cells and a 92% reduction in HCT116 cells [3].

However, the number of uORFs of a transcript alone does not determine their regulatory potential. Other sequence features such as a short intercistronic distance, the overlap with the CDS, a strong initiation codon context and a long distance to the 5' cap can favour translational repression [16,17]. When analysing the *PERK* 5'UTR, it becomes clear that each uORF contains a distinct set of structural features that may define their regulatory potential (Table 3.1), which, in the end, may result in a combined strong repressive action over mORF translation. Fernandes, R. *et al.* (2020) observed that the regulation of *PERK* translation seems to depend on four of its eight uORFs: uORF1, uORF3, uORF4 and uORF8. Particularly, uORF1 is a potent inhibitor of mORF translation, as it alone is responsible for about 89% of the repressive activity of the entire 5'UTR. In contrast, uORF3, uORF4 and uORF8 appear to be only modest dampeners of mORF translation, having less than 40% reduction in the expression of the reporter plasmid. These distinct regulatory activities can be related to the unique structural features of the uORFs. The start codon of uORF1, for instance, displays a strong Kozak context, which is expected to increase its recognition rate by scanning ribosomes [73]. On the other hand, uORF3 and uORF4 have start codons in intermediate contexts, while uORF8 has its start codon in a weak context. This can partially explain why uORF1 is a more competent inhibitory uORF than the others. Accordingly, the group found that uORF1 is highly recognized by the ribosomes, while the other repressive uORFs are only modestly recognized, allowing some degree of ribosome bypass to occur. An interesting observation, however, was that uORF8, besides having a start codon in a weak context, has a repressive activity similar to uORF3, whose start codon is in an intermediate context. This suggests that, as mentioned before, there might be more than one structural feature contributing to modulating the regulatory potential of each uORF. Despite being a major regulator of mORF translation, uORF1 alone cannot repress translation at the level of the intact 5'UTR. Therefore, it was proposed that this uORF could act in combination with the other repressive uORFs to achieve the maximum inhibitory activity. R. Fernandes *et al*. found that

only when uORF1 was combined with uORF3 and/or uORF4 and uORF8, the repressive activity of the resulting 5'UTR could equal the one from the wild-type sequence.

Many translatable short ORFs (such as the ones mentioned above) have been identified within 5'UTRs and/or 3'UTRs of mRNA transcripts with known coding DNA sequences (CDS). A fine balance in the homeostasis of mRNA translation is essential for normal tissue maintenance and disease prevention. During these procedures, some genes can drive global reprogramming of translation toward uORFs [74]. Traditionally, uORFs act as cis-regulatory elements that modulate the expression of the same transcript from the main CDS. However, it is still unclear whether the short peptides encoded by uORFs have trans-acting capabilities. Grasso *et al.* have shown that a uORF associated with the MYC oncogene revealed a novel function in glioblastoma stem cells (GSC) maintenance through tropomyosin receptor kinase B (TRKB), and therapeutic targeting of the short peptide MYC pre-mRNA encoded protein (MPEP) encoded by this uORF offered a potential therapeutic strategy for glioblastoma patients. In contrast, our results have shown that neither of our uORFs of interest (uORF1, 3, 4 and 8) code for proteins containing signal peptides (Figure 4.1), indicating a lack of critical functions regarding markers of protein secretion pathway as well as passenger proteins.

The factor eIF2D was recently found to be required for the uORF-mediated repression of ATF4 mORF translation during basal conditions. Bohlen *et al.* have demonstrated that knockdown of eIF2D leads to reduced ATF4 protein levels both in the absence and presence of stress, apparently because this factor is required for ATF4 translation by promoting translation reinitiation after uORFs in the 5′UTRs of this mRNA [67]. Additionally, eIF4G2 has been reported to provide leaky scanning through some uORFs, which are typical for long 5′ UTRs of mRNAs from higher eukaryotes. Shestakova *et al.* affirm that it is likely the protein can also help the ribosome overcome other impediments during scanning of the 5′ UTRs of animal mRNAs. Here we show that the knockdown of eIF2D and eIF4G2 decreased PERK expression at the protein level, suggesting that these factors are normally required for PERK translation. As eIF2D and eIF4G2 are key elements of translation reinitiation and initiation, respectively, it is not surprising that their decreased availability in the cell can compromise the synthesis of a protein [67, 68]. This means that, like the mechanism described for the *ATF4* mRNA, they appear to be part of a uORF-dependent mechanism responsible for regulating PERK translation (Figures 4.2 and 4.3). Surprisingly, on the other hand, depletion of DENR did not reveal to have any effect on PERK expression. This might be due to the fact that this protein forms a complex with another protein, — MCTS1 — as mentioned before. Since they form a complex, it would be expected that they are both needed for its correct functionality, but nevertheless, it is possible that they might mitigate the lack of their counterparts. For this reason, it would be interesting to study the effects of the depletion of each factor separately, as well as the depletion of both factors in the same experiment. Similarly, since eif3h appears to be necessary for productive reinitiation events, it would be expected that the absence of this protein would cause a significant decrease in our uORF-harbouring transcript. However, this may not have been observed considering that eIF3h is naturally overexpressed in colorectal cancer cells [75], which are the cell lines used in our work. For this reason, it would be interesting to replicate this study in different cell lines and compare the results. As for the eIF3d factor, the lack of significant observable events might be related to the absence of cellular stress. Since this protein is mainly involved in non-canonical translation, the knockdown effects might not be noticeable under normal conditions, meaning that, in the future, it is important to run tests in eif3d^{KO} cells under stress conditions and compare the results. Besides this, it is necessary to study the contribution of all the mentioned factors for maintaining normal steady-state levels of *PERK* mRNA, due to the possibility of their impairment affecting the translation of specific transcription factors involved in the induction of the *EIF2AK3* gene.

As discussed before, different uORF-dependent regulatory mechanisms allow several stress-responsive transcripts to be resistant to the global translational inhibition that occurs upon phosphorylation of eIF2α during stress conditions [3, 11-13]. In agreement with this, several reports showed that ER stress induces the expression of these stress-related factors at both mRNA and protein levels [38-41]. R. Fernandes *et al.* verified that PERK is also transcriptionally induced in Tg-treated cells but, contrary to the uORFs of ATF4, CHOP and GADD34, PERK uORFs maintain most of their repressive activity to regulate PERK protein levels in stress conditions. Several mutations in the *EIF2AK3* gene have been implicated in the development of the WRS due to the expression of dysfunctional forms of PERK, which highlights the importance of this kinase for cell development, homeostasis and survival. As we know that *PERK* uORFs have an important role in regulating PERK expression, it raised the possibility that genetic variants altering these uORFs could lead to unbalanced PERK expression and drive the pathogenesis of WRS. To address this question, we will study the impact of some uORF-altering mutations from WRS patients in the regulation of mORF translation. Fernandes, R. *et al.* found the variants c.-195G>T and c.-263T>C, which shortens uORFs 1, 2 and 4, and disrupts uORF1, respectively, promote a significant increase in mORF translation. At first, this would also suggest that these variants are not related to WRS, as this disease is known to be promoted by insufficient production of functional PERK [46-48]. However, Fernandes, R. *et al.* (2020) observed that despite PERK overexpression not affecting cell viability of HCT116 cells, it resulted in the spontaneous activation of PERK and in the consequent phosphorylation of eIF2α, which may compromise the normal cellular functions or even activate the PERK-related pro-apoptotic pathways at long-term. The sustained PERK activity can potentiate insulin resistance through phosphorylation and activation of the forkhead box O (FoxO) family of transcription factors, whose dysregulated function has been associated with diabetes, a hallmark of WRS [76, 77]. It is possible that, if PERK levels exceed the threshold of BiP inactivation, it may lead to the unintentional activation of downstream targets, like FoxO, and eventually promote disease development. Taking this into account, we think it would be important to further study the pathological potential of variants that induce PERK expression, like c.-195G>T and c.-263T>C, not only in the context of WRS, but in other diseases related to dysregulated PERK expression, like Alzheimer's disease or Huntington's disease [78,79].

In conclusion, PERK uORFs have proven necessary to fine-tuning PERK translation. The combined repressive activity of uORF1, uORF3, uORF4 and uORF8 is likely involved in maintaining PERK in an inactive state, depending on the available amount of BiP. It is possible that when its regulatory mechanisms (uORF-mediated or not) fail or are altered, PERK expression becomes unbalanced, which may drive harmful biological processes and ultimately lead to disease. With this being said we consider it important to include the *PERK* 5'UTR in the screening of disease-related mutations and thoroughly study their impact on cell homeostasis and human disease. Hopefully, new discoveries about the complex effects of PERK dysregulated expression and activity will help to design novel therapeutic approaches for the various diseases in which this kinase appears to be a key player.

References

1. Barbosa, C., Peixeiro, I., & Romão, L. (2013). Gene Expression Regulation by Upstream Open Reading Frames and Human Disease. PLoS Genetics, 9(8), e1003529. https://doi.org/10.1371/journal.pgen.1003529
2. Sonenberg, N., & Hinnebusch, A. G. (2009). Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. Cell, 136(4), 731–745. https://doi.org/10.1016/j.cell.2009.01.042
3. Fernandes, R. (2020, November 1). uORF-mediated translational regulation of PERK: implications for cell homeostasis and human disease. Repositorio.ul.pt. http://hdl.handle.net/10451/48530
4. Hershey, J. W. B., Sonenberg, N., & Mathews, M. B. (2012). Principles of Translational Control: An Overview. Cold Spring Harbor Perspectives in Biology, 4(12), a011528 a011528. https://doi.org/10.1101/cshperspect.a011528
5. Aitken, C. E., & Lorsch, J. R. (2012). A mechanistic overview of translation initiation in eukaryotes. *Nature Structural & Molecular Biology*, *19*(6), 568–576. https://doi.org/10.1038/nsmb.2303
6. Sonenberg, N., & Hinnebusch, A. G. (2009). Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets. *Cell*, *136*(4), 731–745. https://doi.org/10.1016/j.cell.2009.01.042
7. Dever, T. E., Dinman, J. D., & Green, R. (2018). Translation Elongation and Recoding in Eukaryotes. Cold Spring Harbor Perspectives in Biology, 10(8), a032649. https://doi.org/10.1101/cshperspect.a032649
8. Jackson, R. J., Hellen, C. U. T., & Pestova, T. V. (2012). Termination and post-termination events in eukaryotic translation. *Advances in Protein Chemistry and Structural Biology*, *86*, 45–93. https://doi.org/10.1016/B978-0-12-386497-0.00002-5
9. Mancera-Martínez, E., Brito Querido, J., Valasek, L. S., Simonetti, A., & Hashem, Y. (2017). ABCE1: A special factor that orchestrates translation at the crossroad between recycling and initiation. *RNA Biology*, *14*(10), 1279–1285. https://doi.org/10.1080/15476286.2016.1269993
10. Hellen, C. U. T. (2018). Translation Termination and Ribosome Recycling in Eukaryotes. *Cold Spring Harbor Perspectives in Biology*, *10*(10), a032656. https://doi.org/10.1101/cshperspect.a032656
11. Gebauer, F., & Hentze, M. W. (2004). Molecular mechanisms of translational control. Nature Reviews Molecular Cell Biology, 5(10), 827–835. https://doi.org/10.1038/nrm1488
12. Jackson, R. J., Hellen, C. U. T., & Pestova, T. V. (2010). The mechanism of eukaryotic translation initiation and principles of its regulation. Nature Reviews Molecular Cell Biology, 11(2), 113–127. https://doi.org/10.1038/nrm2838
13. Young, S. K., & Wek, R. C. (2016). Upstream Open Reading Frames Differentially Regulate Gene-specific Translation in the Integrated Stress Response. The Journal of Biological Chemistry, 291(33), 16927–16935. https://doi.org/10.1074/jbc.R116.733899
14. Ye, Y., Liang, Y., Yu, Q., Hu, L., Li, H., Zhang, Z., & Xu, X. (2015). Analysis of human upstream open reading frames and impact on gene expression. Human Genetics, 134(6), 605–612. https://doi.org/10.1007/s00439-015-1544-7
15. Rodriguez, C. M., Sang Hoon Chun, Mills, R. E., & Todd, P. M. (2019). Translation of upstream open reading frames in a model of neuronal differentiation. 20(1). https://doi.org/10.1186/s12864-019-5775-1
16. Calvo, S. E., Pagliarini, D. J., & Mootha, V. K. (2009). Upstream open reading frames cause widespread reduction of protein expression and are polymorphic among humans. Proceedings

of the National Academy of Sciences of the United States of America, 106(18), 7507–7512. https://doi.org/10.1073/pnas.0810916106

17. Johnstone, T. G., Bazzini, A. A., & Giraldez, A. J. (2016). Upstream ORFs are prevalent translational repressors in vertebrates. The EMBO Journal, 35(7), 706–723. https://doi.org/10.15252/embj.201592759
18. Crowe, M. L., Wang, X.-Q., & Rothnagel, J. A. (2006). Evidence for conservation and selection of upstream open reading frames suggests probable encoding of bioactive peptides. BMC Genomics, 7(1). https://doi.org/10.1186/1471-2164-7-16
19. Wethmar, K. (2014). The regulatory potential of upstream open reading frames in eukaryotic gene expression. Wiley Interdisciplinary Reviews: RNA, 5(6), 765–768. https://doi.org/10.1002/wrna.1245
20. Lin, Y., May, G. E., Kready, H., Nazzaro, L., Mao, M., Spealman, P., Creeger, Y., & McManus, C. J. (2019). Impacts of uORF codon identity and position on translation regulation. Nucleic Acids Research, 47(17), 9358–9367. https://doi.org/10.1093/nar/gkz681
21. Bertolotti, A., Zhang, Y., Hendershot, L. M., Harding, H. P., & Ron, D. (2000). Dynamic interaction of BiP and ER stress transducers in the unfolded-protein response. Nature Cell Biology, 2(6), 326–332. https://doi.org/10.1038/35014014
22. Vattem, K. M., & Wek, R. C. (2004). Reinitiation involving upstream ORFs regulates ATF4 mRNA translation in mammalian cells. *Proceedings of the National Academy of Sciences*, *101*(31), 11269–11274. https://doi.org/10.1073/pnas.0400541101
23. Do Yun Lee, Cevallos, R. C., & Jan, E. (2009). An Upstream Open Reading Frame Regulates Translation of GADD34 during Cellular Stresses That Induce eIF2α Phosphorylation. 284(11), 6661–6673. https://doi.org/10.1074/jbc.m806735200
24. Gardner, L. B. (2008). Hypoxic Inhibition of Nonsense-Mediated RNA Decay Regulates Gene Expression and the Integrated Stress Response. Molecular and Cellular Biology, 28(11), 3729–3741. https://doi.org/10.1128/mcb.02284-07
25. Mendell, J. T., Sharifi, N. A., Meyers, J. L., Martinez-Murillo, F., & Dietz, H. C. (2004). Nonsense surveillance regulates expression of diverse classes of mammalian transcripts and mutes genomic noise. Nature Genetics, 36(10), 1073–1078. https://doi.org/10.1038/ng1429
26. Losson, R., & Lacroute, F. (1979). Interference of nonsense mutations with eukaryotic messenger RNA stability. Proceedings of the National Academy of Sciences, 76(10), 5134 5137. https://doi.org/10.1073/pnas.76.10.5134
27. Goetz, A. E., & Wilkinson, M. (2017). Stress and the nonsense-mediated RNA decay pathway. Cellular and Molecular Life Sciences, 74(19), 3509–3531. https://doi.org/10.1007/s00018-017-2537-6
28. Yepiskoposyan, H., Aeschimann, F., Nilsson, D., Okoniewski, M., & Muhlemann, O. (2011). Autoregulation of the nonsense-mediated mRNA decay pathway in human cells. RNA, 17(12), 2108–2118. https://doi.org/10.1261/rna.030247.111
29. Maquat, L. E. (2004). Nonsense-mediated mRNA decay: splicing, translation and mRNP dynamics. Nature Reviews Molecular Cell Biology, 5(2), 89–99. https://doi.org/10.1038/nrm1310
30. Supek, F., Lehner, B., & Lindeboom, R. G. H. (2021). To NMD or Not To NMD: Nonsense-Mediated mRNA Decay in Cancer and Other Genetic Diseases. Trends in Genetics, 37(7), 657–668. https://doi.org/10.1016/j.tig.2020.11.002
31. Torrance, V., & Lydall, D. (2018). Overlapping open reading frames strongly reduce human and yeast STN1 gene expression and affect telomere function. *PLoS Genetics*, *14*(8), e1007523. https://doi.org/10.1371/journal.pgen.1007523

32. Wethmar, K., Schulz, J., Muro, E. M., Talyan, S., Andrade-Navarro, M. A., & Leutz, A. (2015). Comprehensive translational control of tyrosine kinase expression by upstream open reading frames. Oncogene, 35(13), 1736–1742. https://doi.org/10.1038/onc.2015.233
33. Twigg, S. M., Babbs, C., van, Goriely, A., Taylor, S., McGowan, S. J., Giannoulatou, E., Lonie, L., Jiannis Ragoussis, Elham Sadighi Akha, Samantha, Roseli Maria Zechi-Ceide, Hoogeboom, J., Pober, B. R., Toriello, H. V., Wall, S. A., Maria Rita Passos-Bueno, Brunner, H. G., Irene M.J. Mathijssen, & Lu, C. (2013). Cellular interference in craniofrontonasal syndrome: males mosaic for mutations in the X-linked EFNB1 gene are more severely affected than true hemizygotes. 22(8), 1654–1662. https://doi.org/10.1093/hmg/ddt015
34. Carvalho, V., Kague, E., Camila Manso Musso, Geronimo, T., Freitas, R. G., Débora Romeo Bertola, Stephen R.F. Twigg, & Maria Rita Passos-Bueno. (2019). Craniofrontonasal Syndrome Caused by Introduction of a Novel uATG in the 5′UTR of EFNB1. 10(1-2), 40 47. https://doi.org/10.1159/000490635
35. Harding, H. P., Zhang, Y., & Ron, D. (1999). Protein translation and folding are coupled by an endoplasmic-reticulum-resident kinase. Nature, 397(6716), 271–274. https://doi.org/10.1038/16729
36. Oren, Y. S., McClure, M. L., Rowe, S. M., Sorscher, E. J., Bester, A. C., Manor, M., Kerem, E., Rivlin, J., Zahdeh, F., Mann, M., Geiger, T., & Kerem, B. (2014). The unfolded protein response affects readthrough of premature termination codons. EMBO Molecular Medicine, 6(5), 685–701. https://doi.org/10.1002/emmm.201303347
37. Donnelly, N., Gorman, A. M., Gupta, S., & Samali, A. (2013). The eIF2α kinases: their structures and functions. Cellular and Molecular Life Sciences, 70(19), 3493–3511. https://doi.org/10.1007/s00018-012-1252-6
38. Ma, K., Vattem, K. M., & Wek, R. C. (2002). Dimerization and Release of Molecular Chaperone Inhibition Facilitate Activation of Eukaryotic Initiation Factor-2 Kinase in Response to Endoplasmic Reticulum Stress. Journal of Biological Chemistry, 277(21), 18728–18735. https://doi.org/10.1074/jbc.m200903200
39. Pakos-Zebrucka, K., Koryga, I., Mnich, K., Ljujic, M., Samali, A., & Gorman, A. M. (2016). The integrated stress response. EMBO Reports, 17(10), 1374–1395. https://doi.org/10.15252/embr.201642195
40. Palam, L. R., Baird, T. D., & Wek, R. C. (2011). Phosphorylation of eIF2 Facilitates Ribosomal Bypass of an Inhibitory Upstream ORF to Enhance CHOP Translation * ♦. Journal of Biological Chemistry, 286(13), 10939–10949. https://doi.org/10.1074/jbc.M110.216093
41. Do Yun Lee, Cevallos, R. C., & Jan, E. (2009). An Upstream Open Reading Frame Regulates Translation of GADD34 during Cellular Stresses That Induce eIF2α Phosphorylation. 284(11), 6661–6673. https://doi.org/10.1074/jbc.m806735200
42. Shi, Y., Vattem, K. M., Sood, R., An, J., Liang, J., Stramm, L., & Wek, R. C. (1998). Identification and Characterization of Pancreatic Eukaryotic Initiation Factor 2 α-Subunit Kinase, PEK, Involved in Translational Control. Molecular and Cellular Biology, 18(12), 7499–7509. https://doi.org/10.1128/mcb.18.12.7499
43. Lin, J. H., Li, H., Zhang, Y., Ron, D., & Walter, P. (2009). Divergent Effects of PERK and IRE1 Signaling on Cell Viability. PLoS ONE, 4(1), e4170. https://doi.org/10.1371/journal.pone.0004170
44. Julier, C., & Nicolino, M. (2010). Wolcott-Rallison syndrome. Orphanet Journal of Rare Diseases, 5(1). https://doi.org/10.1186/1750-1172-5-29
45. Delépine, M., Nicolino, M., Barrett, T., Golamaully, M., Mark Lathrop, G., & Julier, C. (2000). EIF2AK3, encoding translation initiation factor 2-α kinase 3, is mutated in patients with Wolcott-Rallison syndrome. Nature Genetics, 25(4), 406–409. https://doi.org/10.1038/78085

46. Wei, J., Sheng, X., Feng, D., McGrath, B. C., & Cavener, D. R. (2008). PERK is essential for neonatal skeletal development to regulate osteoblast proliferation and differentiation. Journal of Cellular Physiology, 217(3), 693–707. https://doi.org/10.1002/jcp.21543
47. Zhang, W., Feng, D., Li, Y., Iida, K., McGrath, B., & Cavener, D. R. (2006). PERK EIF2AK3 control of pancreatic β cell differentiation and proliferation is required for postnatal glucose homeostasis. Cell Metabolism, 4(6), 491–497. https://doi.org/10.1016/j.cmet.2006.11.002
48. Iida, K., Li, Y., McGrath, B. C., Frank, A., & Cavener, D. R. (2007). PERK eIF2 alpha kinase is required to regulate the viability of the exocrine pancreas in mice. BMC Cell Biology, 8(1). https://doi.org/10.1186/1471-2121-8-38
49. Zhang, P., McGrath, B., Li, S., Frank, A., Zambito, F., Reinert, J., Gannon, M., Ma, K., McNaughton, K., & Cavener, D. R. (2002). The PERK Eukaryotic Initiation Factor 2α Kinase Is Required for the Development of the Skeletal System, Postnatal Growth, and the Function and Viability of the Pancreas. Molecular and Cellular Biology, 22(11), 3864–3874. https://doi.org/10.1128/mcb.22.11.3864-3874.2002
50. Harding, H. P., Zeng, H., Zhang, Y., Jungries, R., Chung, P., Plesken, H., Sabatini, D. D., & Ron, D. (2001). Diabetes Mellitus and Exocrine Pancreatic Dysfunction in Perk−/− Mice Reveals a Role for Translational Control in Secretory Cell Survival. Molecular Cell, 7(6), 1153–1163. https://doi.org/10.1016/s1097-2765(01)00264-7
51. Zhang, W., Feng, D., Li, Y., Iida, K., McGrath, B., & Cavener, D. R. (2006). PERK EIF2AK3 control of pancreatic β cell differentiation and proliferation is required for postnatal glucose homeostasis. Cell Metabolism, 4(6), 491–497. https://doi.org/10.1016/j.cmet.2006.11.002
52. Feng, D., Wei, J., Gupta, S., McGrath, B. C., & Cavener, D. R. (2009). Acute ablation of PERK results in ER dysfunctions followed by reduced insulin secretion and cell proliferation. BMC Cell Biology, 10(1). https://doi.org/10.1186/1471-2121-10-61
53. Julier, C., & Nicolino, M. (2010). Wolcott-Rallison syndrome. Orphanet Journal of Rare Diseases, 5(1). https://doi.org/10.1186/1750-1172-5-29
54. Scheuner, D., Song, B., McEwen, E., Liu, C., Laybutt, R., Gillespie, P., Saunders, T., Bonner-Weir, S., & Kaufman, R. J. (2001). Translational Control Is Required for the Unfolded Protein Response and In Vivo Glucose Homeostasis. Molecular Cell, 7(6), 1165–1176. https://doi.org/10.1016/s1097-2765(01)00265-9
55. Back, S. H., Scheuner, D., Han, J., Song, B., Ribick, M., Wang, J., Gildersleeve, R. D., Pennathur, S., & Kaufman, R. J. (2009). Translation Attenuation through eIF2α Phosphorylation Prevents Oxidative Stress and Maintains the Differentiated State in β Cells. Cell Metabolism, 10(1), 13–26. https://doi.org/10.1016/j.cmet.2009.06.002
56. Harvey, C. J., Thimmulappa, R. K., Singh, A., Blake, D. J., Ling, G., Wakabayashi, N., Fujii, J., Myers, A., & Biswal, S. (2009). Nrf2-regulated glutathione recycling independent of biosynthesis is critical for cell survival during oxidative stress. Free Radical Biology and Medicine, 46(4), 443–453. https://doi.org/10.1016/j.freeradbiomed.2008.10.040
57. Soma Jobbagy, Vitturi, D. A., Salvatore, S. R., Turell, L., Pires, M. F., Kansanen, E., Batthyany, C., Lancaster, J. R., Freeman, B. A., & Schopfer, F. J. (2019). Electrophiles modulate glutathione reductase activity via alkylation and upregulation of glutathione biosynthesis. Redox Biology, 21, 101050–101050. https://doi.org/10.1016/j.redox.2018.11.008
58. Cullinan, S. B., Zhang, D., Hannink, M., Arvisais, E., Kaufman, R. J., & Diehl, J. A. (2003). Nrf2 Is a Direct PERK Substrate and Effector of PERK-Dependent Cell Survival. Molecular and Cellular Biology, 23(20), 7198–7209. https://doi.org/10.1128/mcb.23.20.7198-7209.2003
59. Yagishita, Y., Fukutomi, T., Sugawara, A., Kawamura, H., Takahashi, T., Pi, J., Akira Uruno, & Yamamoto, M. (2014). Nrf2 Protects Pancreatic β-Cells From Oxidative and Nitrosative Stress in Diabetic Model Mice. Diabetes, 63(2), 605–618. https://doi.org/10.2337/db13-0909

60. Shi, X., Deng, H., Dai, Z., Xu, Y., Xiong, X., Ma, P., & Cheng, J. (2016). Nr2e1 Deficiency Augments Palmitate-Induced Oxidative Stress in Beta Cells. Oxidative Medicine and Cellular Longevity, 2016, 1–9. https://doi.org/10.1155/2016/9648769
61. Owji, H., Nezafat, N., Negahdaripour, M., Hajiebrahimi, A., & Ghasemi, Y. (2018). A comprehensive review of signal peptides: Structure, roles, and applications. European Journal of Cell Biology, 97(6), 422–441. https://doi.org/10.1016/j.ejcb.2018.06.003
62. D. Thomas Rutkowski, Lippincott-Schwartz, J., Polansky, J. R., & Lingappa, V. R. (2003). Signal Sequences Initiate the Pathway of Maturation in the Endoplasmic Reticulum Lumen. 278(32), 30365–30372. https://doi.org/10.1074/jbc.m302117200
63. Hegde, R. S., & Bernstein, H. D. (2006). The surprising complexity of signal sequences. Trends in Biochemical Sciences, 31(10), 563–571. https://doi.org/10.1016/j.tibs.2006.08.004
64. Kanapin, A., Serge Batalov, Davis, M. J., Gough, J., Grimmond, S. M., Hideya Kawaji, Magrane, M., Matsuda, H., Schönbach, C., Teasdale, R. D., Members, G., & Yuan, Z. (2003). Mouse Proteome Analysis. Genome Research, 13(6b), 1335–1344. https://doi.org/10.1101/gr.978703
65. Bonfanti, R., Colombo, C., Nocerino, V., Massa, O., Lampasona, V., Iafusco, D., Viscardi, M., Chiumello, G., Meschi, F., & Barbetti, F. (2008). Insulin Gene Mutations as Cause of Diabetes in Children Negative for Five Type 1 Diabetes Autoantibodies. Diabetes Care, 32(1), 123–125. https://doi.org/10.2337/dc08-0783
66. McConkey, D. J. (2017). The integrated stress response and proteotoxicity in cancer therapy. Biochemical and Biophysical Research Communications, 482(3), 450–453. https://doi.org/10.1016/j.bbrc.2016.11.047
67. Bohlen, J., Harbrecht, L., Blanco, S., Clemm von Hohenberg, K., Fenzl, K., Kramer, G., Bukau, B., & Teleman, A. A. (2020). DENR promotes translation reinitiation via ribosome recycling to drive expression of oncogenes including ATF4. Nature Communications, 11(1). https://doi.org/10.1038/s41467-020-18452-2
68. Shestakova, E. D., Smirnova, V. V., Shatsky, I. N., & Terenin, I. M. (2022). Specific mechanisms of translation initiation in higher eukaryotes: the eIF4G2 story. RNA, 29(3), 282–299. https://doi.org/10.1261/rna.079462.122
69. Volta, V., Pérez-Baos, S., de, C., Katsara, O., Ernlund, A., Dornbaum, S., & Schneider, R. J. (2021). A DAP5/eIF3d alternate mRNA translation mechanism promotes differentiation and immune suppression by human regulatory T cells. Nature Communications, 12(1). https://doi.org/10.1038/s41467-021-27087-w
70. Liu, Y., Cui, J., Hoffman, A. R., & Hu, J.-F. (2023). Eukaryotic translation initiation factor eIF4G2 opens novel paths for protein synthesis in development, apoptosis and cell differentiation. Cell Proliferation, 56(3), e13367. https://doi.org/10.1111/cpr.13367
71. Roy, B., Vaughn, J. N., Kim, B.-H., Zhou, F., Gilchrist, M. A., & Von Arnim, A. G. (2010). The h subunit of eIF3 promotes reinitiation competence during translation of mRNAs harboring upstream open reading frames. RNA, 16(4), 748–761. https://doi.org/10.1261/rna.2056010
72. Young, S. K., Willy, J. A., Wu, C., Sachs, M. S., & Wek, R. C. (2015). Ribosome Reinitiation Directs Gene-specific Translation and Regulates the Integrated Stress Response. Journal of Biological Chemistry, 290(47), 28257–28271. https://doi.org/10.1074/jbc.m115.693184
73. Kozak, M. (1986). Point mutations define a sequence flanking the AUG initiator codon that modulates translation by eukaryotic ribosomes. Cell, 44(2), 283–292. https://doi.org/10.1016/0092-8674(86)90762-2
74. Li, F., Yang, K., Gao, X., Zhang, M., Gu, D., Wu, X., Lu, C., Wu, Q., Dixit, D., Gimple, R. C., You, Y., Mack, S. C., Shi, Y., Kang, T., Agnihotri, S. A., Taylor, M. D., Rich, J. N., Zhang, N., & Wang, X. (2024). A peptide encoded by upstream open reading frame of MYC binds to

tropomyosin receptor kinase B and promotes glioblastoma growth in mice. Science Translational Medicine, 16(767). https://doi.org/10.1126/scitranslmed.adk9524

75. Jin, H., Huang, X., Pan, Q., Ma, N., Xie, X., Wei, Y., Yu, F., Wen, W., Zhang, B., Zhang, P., Chen, X., Wang, J., Liu, R.-Y., Lin, J., Meng, X., & Lee, M.-H. (2024). The EIF3H-HAX1 axis increases RAF-MEK-ERK signaling activity to promote colorectal cancer progression. Nature Communications, 15(1), 2551. https://doi.org/10.1038/s41467-024-46521-3
76. Zhang, W., Hietakangas, V., Wee, S., Lim, S. C., Gunaratne, J., & Cohen, S. M. (2013). ER stress potentiates insulin resistance through PERK-mediated FOXO phosphorylation. Genes & Development, 27(4), 441–449. https://doi.org/10.1101/gad.201731.112
77. Lee, S., & Dong, H. H. (2017). FoxO integration of insulin signaling with glucose and lipid metabolism. Journal of Endocrinology, 233(2), R67–R79. https://doi.org/10.1530/joe-17-0002
78. Ohno, M. (2018). PERK as a hub of multiple pathogenic pathways leading to memory deficits and neurodegeneration in Alzheimer's disease. Brain Research Bulletin, 141, 72–78. https://doi.org/10.1016/j.brainresbull.2017.08.007
79. Shacham, T., Patel, C., & Lederkremer, G. Z. (2021). PERK Pathway and Neurodegenerative Disease: To Inhibit or to Activate? Biomolecules, 11(3), 354. https://doi.org/10.3390/biom11030354