UNIVERSIDADE DE LISBOA
FACULDADE DE CIÊNCIAS
DEPARTAMENTO DE QUIMICA E BIOQUIMICA

Study of human Argonaute 1 cap-independent translation

Manuel Nuno Condinho Rodrigues Reis de Carvalho

Mestrado Bioquímica
Especialização em Bioquímica

Dissertação orientada por:
Doutora Luísa Romão
Doutora Juliane Menezes

2019

“I have not failed. I´ve just found 10.000 ways that won´t work”

Thomas Edison

Agradecimentos

Para este trabalho ter sido feito tenho de agradecer a diversas pessoas, que seja pela sua ajuda ou pelas suas palavras possibilitaram a conclusão desta etapa tão enriquecedora.

Em primeiro lugar, agradecer à Doutora Luísa Romão por me ter recebido no seu laboratório e pela confiança que sempre me deu enquanto desenvolvi o meu trabalho. Obrigado por todos conselhos e incentivos que tornaram este período cheio de desafios e aprendizagens.

Acima de tudo um grande, grande obrigado à Juliane por tudo o que me ensinaste no laboratório, pela paciência que sempre tiveste mesmo quando fiz alguma coisa errada, e pela força e confiança que sempre me deste, mesmo quando as coisas não estavam a correr como o esperado. Muitas palavras seriam precisas para descrever como foste uma orientadora cinco estrelas, e por isso estou-te muito grato. Muito obrigado por tudo o que aprendi contigo!

E claro que também tenho de agradecer a todas as pessoas espetaculares que conheci durante este ano. Muito obrigado ao Paulo, Joana, Rafael, Rafaela, Patrícia, Bruna e Neuza por tornarem o laboratório de metabolismo de mRNA o melhor sítio para se trabalhar. Agradeço-vos do fundo do coração por terem partilhado o meu dia-a-dia, por me terem ensinado imensas coisas e por fazerem do nosso laboratório um local de bom ambiente, onde dava gosto trabalhar. Graças a vocês todas as dificuldades eram muito mais fáceis de ultrapassar. Muito obrigado!

Também tenho de agradecer às meninas do laboratório de Oncobiologia, em especial, à Patrícia Barros, que sempre se mostrou disponível para ajudar no que fosse preciso. Um bocadinho desta tese também é graças a vocês, já que várias foram as membranas que deixei a incubar no vosso frigorífico.

Igualmente, não posso deixar de agradecer a todas as pessoas que conheci no Departamento de Genética Humana do INSA. Desde as meninas da receção que nos animavam os almoços, às pessoas dos outros laboratórios que sempre se mostraram disponíveis para ajudar, o meu grande obrigado.

Mas como a vida não é só feita de trabalho, também agradeço a todos os amigos que conheci durante a faculdade e que me acompanharam nesta etapa. Muito obrigado por todas as palavras e mensagens de incentivo de que a tese iria correr bem e que os resultados iriam aparecer.

E como não poderia deixar de ser também tenho de agradecer aos meus amigos moiteiros, Afonso, Ruben, Mariana, Inês e Kiko pelo companheirismo e por apesar de já terem passados tantos anos, continuarmos iguais. Obrigado por todas aquelas noites de sexta ou sábado em que ficávamos no café a falar horas de como tinha sido a semana de cada um, a rir, a gozar uns com os outros e a relembrar aventuras do passado. Graças a esses dias conseguia-se meter uma má semana para trás das costas e encarar a próxima com um renovado otimismo.

Aos meus pais agradeço-lhes tudo o que fizeram por mim, por me terem ensinado tanta coisa e por nunca terem deixado de acompanhar o meu percurso. Não me esqueço que todos os dias me perguntavam pelas minhas colónias e se elas estavam a crescer bem. Ensinaram-me que nunca deveria desistir dos meus objetivos e que deveria trabalhar sempre para os conseguir, porque com esforço tudo se consegue e que o sucesso só depende de nós. Obrigado por me terem sempre dito para nunca deixar fugir as boas oportunidades quando elas aparecem, e que às vezes é melhor o desenrascanço do que a habilidade. À minha irmã, obrigado por aturares as minhas parvoíces, especialmente nestes últimos tempos em que tenho estado em casa.

Por último e mais importante, obrigado à minha namorada, Filipa. Obrigado por seres a pessoa que és, obrigado pelo teu amor, companheirismo e por fazeres de mim uma pessoa melhor. Não tenho como agradecer o apoio que foste para mim nesta fase. Lembro-me quando às vezes vinha frustrado porque uma clonagem não tinha saído ou um Western não tinha funcionado, e tu estavas sempre lá para me animar e meter para cima. Também de te agradeço muito por teres corrigido tão bem o inglês desta tese. Espero que estejas sempre ao meu lado e orgulhosa do que fiz!

Resumo

A expressão génica é um processo a partir do qual a informação genética contida no ácido desoxirribonucleico (DNA, do inglês *deoxyribonucleic acid*) é descodificada em proteínas que serão usadas pelos sistemas biológicos, para a sua sobrevivência, manutenção ou desenvolvimento. Esta é dividida em várias etapas, sendo que tudo se inicia no núcleo com a transcrição da informação contida no DNA em moléculas de ácido ribonucleico mensageiro (mRNA, do inglês *messenger ribonucleic acid*). À medida que vão sendo sintetizadas, as moléculas de mRNA sofrem um processamento, em que serão introduzidas modificações que serão importantes para estabilidade do transcrito e para as etapas a jusante. Entre estas modificações estão a introdução da estrutura *cap* m^7G na extremidade 5´ do transcrito, a poliadenilaçao da extremidade 3´ e o *splicing* do mRNA. Uma vez processado, o mRNA maduro será transportado do núcleo para o citoplasma, onde será traduzido para proteína.

Em condições normais, o mRNA de eucariotas é traduzido de uma forma canónica, em que inicialmente é necessário que exista o reconhecimento da estrutura *cap* na extremidade 5´ do transcrito. Desta forma, a tradução inicia-se com o recrutamento da subunidade ribossomal 40S e de fatores de iniciação (eIF, do inglês *eukaryotic initiation factor*) para a extremidade 5´ do mRNA, sendo que será este complexo, pela ação do eIF4E que reconhecerá o *cap*. Uma vez ligados, a subunidade ribossomal inicia o *scan* da região 5´ não codificante (UTR, do inglês *untranslated region*) até encontrar um codão de iniciação em contexto favorável. De seguida, dar-se-á a síntese da proteína que, terminada, a tradução será utilizada para o funcionamento celular.

Contudo durante alguns processos energeticamente dispendiosos para a célula, como a mitose, ou em condições de stress, como a hipoxia, a tradução canónica é afetada. Isto ocorre porque muitas vezes estes estímulos provocam a inibição do reconhecimento da estrutura *cap*, comprometendo assim a iniciação da tradução. No entanto, existem mRNAs eucariotas cuja expressão não é comprometida aquando destes eventos, sendo que estes transcritos normalmente codificam para proteínas fundamentais para a célula ou de resposta a stress. Desta forma em condições de inibição da tradução canónica, existem proteínas cuja síntese não é afetada, sendo mantida por mecanismos alternativos. Entre estes estão incluídos aqueles em que ocorre o recrutamento ribossomal para o mRNA sem que haja o reconhecimento do *cap* e/ou *scan* da 5´UTR. Nestes casos, o recrutamento ribossomal ocorre através de estruturas secundárias presentes na 5´UTR do mRNA, que interagem com o ribossoma, e com o auxílio ou não de fatores proteicos, posicionam este junto do codão de iniciação. Em eucariotas já foram descritos vários mRNAs cuja tradução é mantida em condições de inibição da tradução canónica, sendo os mecanismos alternativos ativados por vários tipos de stress como a hipoxia, a limitação de nutrientes ou a apoptose. Assim, os mecanismos alternativos de iniciação da tradução têm um papel importante para a manutenção da homeostasia celular durante situações de stress. Em sentido contrário estes também podem ter um papel negativo, nomeadamente na promoção do processo de tumorigénese. Geralmente os ambientes tumorais apresentam hipoxia e limitação de nutrientes, que são condições que inibem a tradução canónica e promovem mecanismos alternativos de síntese proteica. Deste modo, as células cancerígenas têm uma vantagem, pois podem manter a síntese de vários oncogenes de uma maneira não-canónica.

No nosso trabalho, colocou-se a hipótese de que as proteínas *Argonaute*, mais propriamente a AGO1 (do inglês *human Argonaute RNA-induced silencing complex catalytic component 1*) poderiam ser traduzidas por mecanismos não-canónicos em condições de inibição da iniciação da tradução dependente do reconhecimento do *cap*. Estas proteínas estão envolvidas na síntese dos miRNAs e no silenciamento génico, que por sua vez participam na resposta a alguns stresses e na manutenção do funcionamento celular. Assim, devido à sua importância propôs-se que a síntese da AGO1 possa ser mantida durante condições de stress celular por um mecanismo não-canónico. Resultados anteriores obtidos no nosso laboratório utilizando um vetor bicistrónico, demonstraram que a sequência correspondente à 5´UTR do mRNA da AGO1 é capaz de mediar a tradução independente da estrutura *cap* de um gene repórter.

Também se verificou que esta atividade da 5´UTR da AGO1 é mantida tanto em condições normais, como em condições de stress celular que inibem a tradução canónica, como o stress do reticulo endoplasmático ou a hipoxia.

Assim o propósito deste trabalho foi prosseguir nesta linha de investigação, continuando a estudar a tradução independente da estrutura *cap* mediada pela 5´UTR da AGO1. Desta forma, o primeiro objetivo foi reforçar os resultados obtidos anteriormente, pretendendo-se mostrar que a 5´UTR da AGO1 medeia um mecanismo de iniciação de tradução alternativo, num vetor bicistrónico diferente do utilizado anteriormente. Para isso o sistema bicistrónico utilizado contem dois genes repórteres, um correspondente à proteína fluorescente verde (EGFP, do inglês *enhanced green fluorescent protein*) e outro correspondente a uma proteína fluorescente vermelha (*mCherry*). A característica deste sistema é que no primeiro cistrão está a sequência da EGFP, que é traduzida de uma forma dependente da estrutura cap, ou seja, tradução canónica, enquanto que no segundo cistrão está a sequência do mCherry, que a ser traduzida, será duma forma independente da estrutura cap. Entre os dois cistrões está clonado um *hairpin*, de modo a prevenir a reiniciação da tradução. Tendo isto em mente, a estratégia foi clonar a sequencia correspondente à 5´UTR da AGO1 a montante do mCherry e verificar se esta conseguia potenciar a iniciação da tradução independente da estrutura cap. Nesta tese apresentamos a estratégia de clonagem desenvolvida para obter este constructo, ainda que sem sucesso aquando do término dos trabalhos. Contudo apresentamos algumas possíveis soluções para completar estas clonagens.

Outro objetivo desta tese foi avaliar como varia a expressão das proteínas *Argonaute*, mais propriamente da AGO1 endógena, durante um stress que inibe a tradução canónica. A abordagem experimental foi tratar células HCT116 (linha celular derivada de cancro colorretal) com tapsigargina, que é uma droga que induz o stress do reticulo endoplasmático e, por consequência, a inibição da tradução canónica. Posteriormente foram analisados os efeitos do tratamento com esta droga nos níveis de mRNA das quatro *Argonautes* (AGO1, AGO2, AGO3 e AGO4), no nível de proteína do *pool* total de AGOs e no nível de proteína de AGO1. Os resultados obtidos demonstram que o tratamento com a tapsigargina induz uma diminuição nos níveis de mRNA da AGO1 e AGO2 para 54% e 62%, respetivamente. Um resultado que nós acreditamos ser uma consequência indireta da inibição da tradução induzida pela droga. Isto é, ao bloquear-se um processo tão determinante como a tradução, poderão ser afetados indiretamente outros processos celulares. Neste caso, parece indicar que a inibição da tradução canónica afeta a disponibilidade de dois transcritos, o da AGO1 e o da AGO2. Adicionalmente, verificou-se que o tratamento com esta droga afeta a quantidade total de proteínas *Argonaute*, uma vez que se observa a diminuição do pool de AGOs. Neste caso este resultado dever-se-á ao facto do tratamento com a tapsigargina induzir a diminuição dos níveis de mRNA da AGO1 e AGO2, fazendo com que haja menos transcrito disponível para a síntese destas proteínas. Com a expectável diminuição da quantidade de proteína AGO1 e AGO2, a consequência é a diminuição do pool de todas as AGOs. Finalmente, verificamos que o tratamento com a tapsigargina provoca a diminuição dos níveis proteicos de AGO1 endógena. De certa forma, este é um resultado inconclusivo para o estudo da tradução não-canónico desta proteína durante condições de stress. No entanto, pensamos que este resultado se deve principalmente a dois fatores. Um tem a ver com facto de o tratamento com a droga fazer diminuir o mRNA da AGO1, havendo por isso menos transcrito disponível para a tradução, o que poderá levar a uma menor síntese de proteína, mesmo que de uma forma alternativa. O outro fator está relacionado com a possibilidade de que o mecanismo de tradução não-canónico não ser tão eficiente como a via canónica, o que explica o decréscimo observado. Assim, em experiências futuras será importante testar outros tipos de stresses que inibam a tradução canónica, mas que não provoquem alterações na quantidade de mRNA e testar outras linhas celulares. Para além disso dever-se-á testar alguns stresses, que induzam maior atividade da 5´UTR da AGO1, em termos de síntese proteica.

Por fim, o terceiro objetivo deste trabalho foi tentar prever quais as características estruturais que fazem com que a 5´UTR da AGO1 consiga promover um mecanismo não-canónico de iniciação da tradução. Desta forma tentar-se-ia aproximar a estrutura secundária adotada por esta região. Para isso fez-se uma análise *in silico* à sequência correspondente à 5´UTR da AGO1. Verificou-se que esta tem um comprimento de 213 nucleótidos, apresenta uma *upstream ORF* (do inglês *open reading frame)* e tem

uma percentagem de guaninas/citosinas correspondente a 72.3%. Especialmente devido a este último fator a 5´UTR da AGO1 tem tendência a organizar-se numa estrutura secundária estável, que possui quatro *stem loops*. Como perspetiva futura, temos como objetivo validar experimentalmente esta previsão estrutural, através de dicroísmo circular. Posteriormente, fizemos uma análise delecional *in silico*, de modo a prever a dinâmica estrutural e verificar se havia alguma sequência/estrutura mínima necessária para manter a conformação espacial da 5´UTR da AGO1 estável. Para isso removeu-se, sequencialmente nucleótidos da 5´UTR da AGO1, tanto no sentido 5´ para 3´ como no sentido 3´ para 5´, e verificou-se qual o impacto que isso teria na estrutura da região comparativamente à estrutura com todo o comprimento. Pela nossa análise verificou-se que quando é removido o stem loop I ou o stem loop IV, a estrutura não parece alterar-se muito em relação à original, indicado que a conformação se mantém estável quando ocorrem pequenas deleções. Para complementar esta análise *in silico*, pretende-se no futuro validar experimentalmente estas previsões. Isto é verificar se a estrutura real adotada pela 5´UTR da AGO1 é semelhante à prevista computacionalmente, e analisar quais os efeitos de deleções na atividade desta sequência para mediar um mecanismo de tradução não-canónico, utilizando um vetor bicistrónico.

Uma vez concluídas todas estas linhas de investigação pretender-se-á perceber de que forma é que o mecanismo de tradução alternativa de AGO1 pode estar envolvido no processo de tumorigénese.

Palavras-chave: tradução do mRNA, iniciação da tradução, tradução independente da estrutura *cap*, AGO1, stress do reticulo endoplasmático

Abstract

Translation of mRNA into protein is one of the most important and costly processes of the cell, since it supports and regulates several cellular events. Under normal conditions, translation initiation takes place through the canonical mechanism, in which first occurs the recognition of the cap structure at the 5'end of the transcript, followed by the scan of the 5'UTR (untranslated region) until the identification of the initiation codon in a favourable context. However, during high energy cost cell events, such as mitosis, or during some stresses such as hypoxia, mechanisms are activated that decrease canonical translation. Thus, under conditions of inhibition of canonical translation, there are proteins whose synthesis is not affected and is maintained by alternative mechanisms. These include those in which ribosomal recruitment occurs without cap recognition and/or 5'UTR scanning. In these cases, ribosomal recruitment occurs through secondary structures present in mRNA 5'UTR, which interact with the 40S ribosomal subunit, and with or without the help of initiation factors, position it near to the initiation codon. Therefore, it was hypothesized that Argonaute proteins, more properly AGO1 (human Argonaute RNA-induced silencing complex catalytic component 1) could be translated by non-canonical mechanisms due to its importance in the biogenesis of the microRNAs and in the gene silencing. Previous results obtained in our laboratory using a bicistronic system, demonstrated that AGO1 5'UTR is able to mediate cap-independent translation of a reporter gene, even under stress conditions that inhibit canonical translation. Thus, the purpose of this work was to complete this research project, continuing to study the cap-independent translation mediated by the 5'UTR of AGO1. In this way, the first objective was to reinforce the previously obtained results. For this it was intended to show that the AGO1 5´UTR can mediate an alternative initiation mechanism in a bicistronic vector different from the one previously used. For this, the proposed bicistronic system contains two reporter genes, one corresponding to enhanced green fluorescent protein (EGFP) and another corresponding to a red fluorescent protein (mCherry). The characteristic of this system is that in the first cistron is the sequence of EGFP, which is translated by the canonical mechanism. The second cistron is the mCherry sequence that can only be translated in cap-independent manner. With this in mind, the strategy was to clone the sequence corresponding to AGO1 5´UTR, upstream of the mCherry and to verify if it was able to potentiate cap-independent translation of this fluorophore. In this thesis we present the cloning strategy developed to obtain this construct, although we were not able to complete this task. However, we present some possible solutions to complete these clonings. Other objective was to evaluate how the expression of Argonaute proteins, more properly the endogenous AGO1, varies during endoplasmic reticulum stress, which inhibits the canonical translation. For this purpose, we treated HCT116 cells (colorectal cancer-derived cell line) with thapsigargin, which is a drug that induces the stress of the endoplasmic reticulum and, consequently, the inhibition of the canonical translation. Our results demonstrated that treatment with the drug induces a decrease in the AGO1 and AGO2 mRNA levels, a decrease in the total protein amount of Argonautes (AGO1, AGO2, AGO3 and AGO4) and a decrease in the AGO1 protein. In a way, this is an inconclusive result for the study of non-canonical translation of this protein during stress conditions. However, we think that this result is mainly due to two factors. One has to do with the fact that the treatment with the drug decreases the AGO1 mRNA level, thus having less transcript available for translation, which may lead to less protein synthesis, albeit in an alternative way. The other factor is related to the possibility that the non-canonical translation mechanism is not as efficient in as the canonical way, which explains the observed decrease. Thus, in future experiments it will be important to test other stresses that inhibit canonical translation, but that do not induce changes in the amount of mRNA and test this experiment in other cell lines. Finally, we aim to predict what structural features make the AGO1 5´UTR succeed in promoting a non-canonical translation initiation mechanism and thus try to estimate the secondary structure adopted by this sequence. For this we submited AGO1 5'UTR sequence to an *in silico* analysis. Among other characteristics, we found that it has a percentage of guanines/cytosines of 72.3% and it is organized into a stable secondary

structure, which has four stem loops. As a future perspective, we aim to experimentally validate this structural prediction through circular dichroism. Next was performed an *in silico* deletion analysis in which nucleotides were removed from the sequence in order to predict the structural dynamics and to check if there was any minimal sequence/structure required to maintain stable the spatial conformation of AGO1 5'UTR. By our analysis it was verified that when stem loop I or stem loop IV is removed, the structure does not seem to change much in relation to the original, indicating that the conformation remains stable when small deletions occur. To complement this *in silico* deletion analysis it is intended in the future to validate experimentally, in a bicistronic vector, the effect of this deletions. We believe that once these research lines have been completed, we will better understand the relevance of alternative translation mechanisms in the maintenance of synthesis of several proteins, including AGO1, during stress situations, and even during the tumorigenesis process.

Key-words: mRNA translation, translation initiation, cap-independent translation, AGO1, endoplasmic reticulum stress

Index

Figure Index

Table Index

List of abbreviations, acronyms and symbols

µM – micromolar
ºC – Celsius degree
4E-BP – eIF4E-binding protein
A – adenine
A site – aminoacyl site
aa-tRNA – aminoacyl-tRNA
ABCE1 – ATP-binding cassette sub-family E member 1
AGO – argonaute
AP-1 – activator protein 1
ATF4 – activating transcription factor 4
ATP – adenosine triphosphate
Bcl2 – apoptosis regulator Bcl-2
Bip – immunoglobulin heavy-chain-binding protein
bp – base pairs
BSA – bovine serum albumin
C – cytosine
CAT-1 – cationic amino acid transporter
cDNA – complementary DNA
CITE – cap-independent translation enhancer
c-Jun – transcription factor AP-1
c-Kit – Kit proto-oncogene receptor tyrosine kinase
CMV – cytomegalovirus
c-Myc – Myc proto-oncogene protein
CO_2 – carbon dioxide
C-terminal – carboxy-terminal
DAP5 – death-associated protein 5
DENR – density-regulated protein
DMEM – Dulbecco's modified Eagle's medium
DMSO – dimethyl sulfoxide
DNA – deoxyribonucleic acid
dNTP´s – deoxynucleotide triphosphates
E site – exit site
ECL – enhanced chemiluminescence
eEF – eukaryotic elongation factor
EGFP – enhanced green fluorescent protein
EGFR – epidermal growth factor receptor
EGR2 – early growth response 2
eIF – eukaryotic initiation factor
EMCV – encephalomyocarditis virus
eRF – eukaryotic release factor
FLuc – firefly luciferase
FMDV – foot and mouth disease virus
G – guanine
GAP – GTPase-activating protein
GAPDH – glyceraldehyde-3-phosphate dehydrogenase
GCN2 – general control nonderepressible 2
GDP – guanosine diphosphate
GEF – guanine nucleotide exchange factor
GTP – guanosine triphosphate
h – hours
HBB – haemoglobin subunit beta
HCT116 – Human pre-metastatic colorectal carcinoma-derived cell line
HCV – hepatitis C virus

HER2 – human epidermal growth factor receptor 2
HIAP2 – baculoviral IAP repeat-containing protein 2
HIF-1α – hypoxia-inducible factor-1α
hnRNP – heterogeneous nuclear ribonucleoprotein
HRI – heme regulated inhibitor
IBC – inflammatory breast cancer
IGF1R – insulin-like growth factor 1 receptor
IGF-II – insulin-like growth factor II
IRES – internal ribosome entry site
ITAF – IRES trans-acting factor
kDa – kilodalton
La – La autoantigen
LEF-1 – lymphoid enhancer-binding factor 1
let-7 – lethal-7
L-myc – L-myc proto-oncogene protein
m – minutes
m^6A – N^6-methyladenosine
m^7G – 7-methylguanosine
Met-$tRNA_i$ – methionyl initiator tRNA
M_I – NZYDNA Ladder III
MID – middle
M_{II} – NZYDNA Ladder VI
miRNA – microRNA
ml – millilitre
mm – millimetre
mRNA – messenger ribonucleic acid
mTOR – mammalian target of rapamycin
ng – nanogram
N-myc – N-myc proto-oncogene protein
NRF2 – nuclear factor erythroid 2-related factor 2
NT – no treatment
N-terminal – amino-terminal
nts – nucleotides
ODC – ornithine decarboxylase
ORF – open reading frame
P site – peptidyl site
p53 – tumour protein p53
PAB – Poly(A)-binding protein
PAZ – PIWI/Argonaute/Zwille
P-bodies – processing bodies
PBS – phosphate-buffered saline
PCBP2 – poly(rC)-binding protein 2
PCR - polymerase chain reaction
pD – digested plasmid
PDCD4 – programmed cell death 4
PDGF – platelet-derived growth factor
pEGFP_AGO1_mCherry – AGO1 5'UTR-containing bicistronic plasmid with two fluorophores
pEGFP_HBB_mCherry – HBB 5'UTR-containing bicistronic plasmid with two fluorophores
pEGFP_mCherry – empty bicistronic plasmid with two fluorophores
pEGFP_Δ133p53_mCherry – Δ133p53-containing bicistronic plasmid with two fluorophores
PERK – protein kinase R (PKR)-like endoplasmic reticulum kinase
PITSLRE – cyclin-dependent kinase 11
PITSLREp58 – cyclin dependent kinase 11A
PKR – protein kinase RNA-activated
PLB – passive lysis buffer
pND - non-digested plasmid

Poly(A) – polyadenylated
p-p90RSK – phospho-p90 ribosomal S6 kinase
pR_AGO1_F – AGO1 5'UTR-containing bicistronic plasmid with two luciferases
pR_EMCV_F – EMCV IRES-containing bicistronic plasmid with two luciferases
pR_F – empty bicistronic plasmid with two luciferases
pR_HBB_F – HBB 5'UTR-containing bicistronic plasmid with two luciferases
pR_MYC_F – c-Myc IRES-containing bicistronic plasmid with two luciferases
p-SMAD3 – phospho-SMAD family member 3
p-SrcY – phospho-proto-oncogene tyrosine-protein kinase Src
p-STAT – phospho-signal transducer and activator of transcription
PTB – polypyrimidine tract-binding protein
PTC – ribosomal peptidyl transferase centre
PVDF – polyvinylidene fluoride
RHA – RNA helicase A
RISC – RNA-induced silencing complex
RLuc – *Renilla* luciferase
RNAi – RNA interference
RT-PCR – reverse transcription-polymerase chain reaction
RT-qPCR – quantitative reverse transcription PCR
s – seconds
S6K – S6 kinase
SDS-PAGE – sodium dodecyl sulfate–polyacrylamide gel electrophoresis
Ser/Thr kinase – serine/threonine protein kinase
siRNA – short interfering RNA
SL – stem loop
SNAT2 – sodium-dependent neutral amino acid transporter-2
SREBP-1 – sterol regulatory element-binding protein 1
SV40 – simian vacuolating virus 40
T – thymine
TBS – tris-buffered saline
TCP80 – translational control protein 80
TNFα-R – tumour necrosis factor α-receptor
tRNA – transfer ribonucleic acid
U – uracil
Unr – Upstream of N-ras
uORF – upstream open reading frame
UPR – unfolded protein response
UTR – untranslated region
UV – ultra-violet
v/v – volume per volume
VEGF-A – vascular endothelial growth factor A
WT – Wilms´ tumour
XIAP – X-linked inhibitor of apoptosis protein
ΔG – Gibbs minimum free energy

1. Introduction

1.1. Translation in Eukaryotic Gene Expression

Gene expression is the process by which the genetic information contained in the DNA (deoxyribonucleic acid) is converted in proteins that will be used by biological systems for their survival, maintenance or development. Experiments done by Francis Crick and his contemporaries in the late 1950's demonstrated that the information encoded in the genome is transcribed into mRNA (messenger ribonucleic acid) molecules, which are then translated into proteins that will be vital for cellular functioning [1]. This is the central dogma of molecular biology, which served as the basis for later scientific work. Since then, many experiments and researches were made in the sense of understand, in detail, how protein synthesis works and how it is regulated. Nowadays the genetic expression can be divided into these steps: transcription of DNA in mRNA; mRNA processing; mRNA transport from nucleus to cytoplasm; translation of mRNA into protein; and processing of the protein.

Translation is one of the most conserved processes midst organisms and the most expensive for a cell, energetically. Translation has a vital role in the cell, namely in its maintenance, development, resistance, among others and it is regulated by the coordinated action of all its intervening agents (mRNA, protein factors, ribosome and tRNA). In eukaryotes, it is divided in four steps: initiation, elongation, termination and recycling [2, 3]. In the first stage, occurs the recognition of initiation codon by the 40S ribosomal subunit and the eukaryotic initiation factors (eIFs). In elongation, the amino acids are added by the ribosome to the nascent polypeptide chain, according to the sequence encoded in the mRNA. The termination begins when the stop codon is found, in which there is the release of the polypeptide chain. Recycling is the final step, in which the ribosomal subunits are dissociated and the mRNA and the deacylated tRNA are released, being available for further translation rounds [4].

In eukaryotes, the initiation of translation of most mRNAs occurs through a mechanism that requires recognition of the cap m^7G structure at the 5'end of the transcript and the scan of the 5' untranslated region (5´UTR). Hence, this process is known as cap-dependent translation or canonical translation (because it is the most frequent and occurs under normal conditions) [5].

1.1.1. Initiation

The starting point for translation is a rather complex process requiring ribosomal recruitment, participation of various initiation factors, formation of ribonucleoprotein complexes and selection of the initiation codon. This makes it a primary target of a strict regulation and a limiting step in protein synthesis (Fig.1) [4].

The first step of translation initiation is the formation of the ternary complex. This step involves the assembly of eukaryotic initiation factor 2 (eIF2) to a guanosine triphosphate (GTP) molecule and a methionyl initiator tRNA (Met-$tRNA_i$). The formation of this complex is a process promoted by guanine nucleotide exchange factor (GEF) eIF2B. After each round of translation this protein catalyzes the exchange of GDP to GTP. This is a fundamental step in the assembly of the ternary complex since GDP has an ~100-fold higher affinity for eIF2 compared to GTP, making the action of eIF2B fundamental at this stage. In turn, the eIF2-GTP complex will establish a positive interaction with Met-$tRNA_i$ [4, 6].

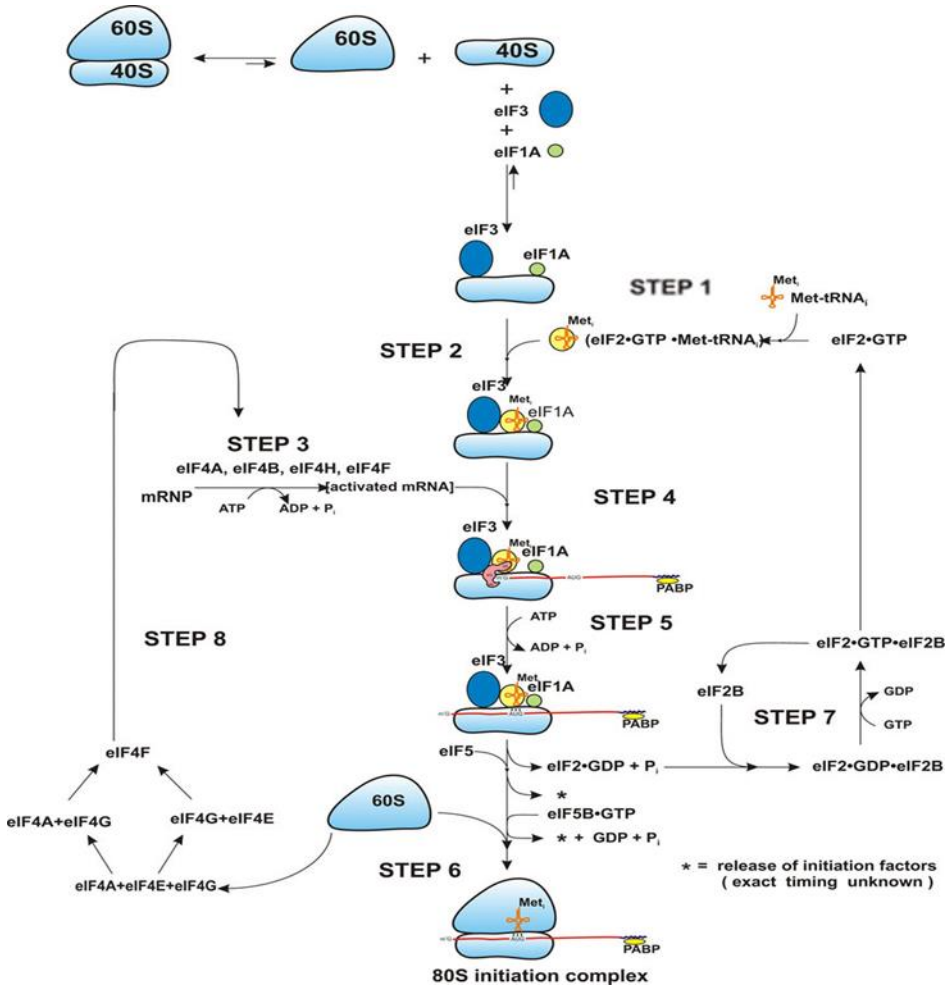

Figure 1. Scheme of canonical initiation of translation: Step 1: Formation of the ternary complex. **Step 2:** Binding of the ternary complex to the 40S ribosomal subunit, mediated by eIF1, eIF1A, eIF3 and eIF5, forming the pre-initiation complex 43S. **Step 3:** Assembly of the eIF4F complex and recognition of the m^7G cap at the 5'end of the mRNA. **Step 4:** Scan of mRNA 5´UTR by 43S pre-initiation complex to find the initiation codon. **Step 5:** Recognition of the initiation codon in optimal context. **Step 6:** Joining of 60S subunit with the aid of eIF5 and eIF5B, thus forming the complex 80S. **Step 7:** Recycling of eIF2-GDP to eIF2-GTP, catalysed by eIF2B. **Step 8:** Reassembly of eIF4F complex after each round of translation. (Adapted from [7]).

After assembly of the ternary complex, it will bind to the 40S ribosomal subunit in a process assisted by eIF1, eIF1A, eIF3 and eIF5 forming the 43S complex or pre-initiation complex (PIC). During this process, eIF3 alters the conformation of the 40S subunit, which then enables the interaction between the ternary complex and the ribosomal subunit [7, 8].

Once the 43S complex is formed it must bind to the mRNA molecule. However, often enough transcripts have secondary structure in their 5´UTR, which hinders ribosome and initiation factors binding. In order to facilitate this process the eIF4F complex - composed by eIF4A, eIF4E and eIF4G - binds to the m^7G cap structure at the 5'end of the mRNA, through the cap-binding factor eIF4E. Simultaneously, the same eIF4F complex, through the eIF4G component, interacts directly with the eIF3 of the PIC and with the poly (A) binding protein (PAB), attached to the 3'poly(A) tail, guiding the mRNA to the complex 43S [4, 9, 10].

Upon binding to the 5'end, the 43S complex begins the scan of the 5'UTR in the search for an initiation codon. The scanning mechanism for the initiation of translation postulates that the 43S complex binds to the 5'end of the mRNA and migrates linearly through the 5'UTR, stopping when it encounters the AUG codon in favourable context, i.e. with a purine in the -3 position and a guanine at the +4 position (where A of AUG corresponds to position +1) [11, 12]. When the 5'UTR is scanned, some of the initiation factors will play an important role, namely, eIF4A which is a DEAD-box helicase (ATP-dependent) that

allows the unwind of secondary structures present in 5'UTR. For this eIF4A helicase activity it will be essential the stimulation of eIF4G and of eIF4B, an independent factor that also stimulates the action of the eIF4F complex [13, 14]. eIF1 and eIF1A will also be important in scanning and detecting of the correct (or cognate) initiation codon, as they can distinguish between cognate and non-cognate AUG codons [9, 15]. Upon recognition of the initiation codon the scanning stops. Immediately, the AUG codon is positioned at the P (peptidyl) site of the ribosome and there is a pairing between the bases of the AUG codon and the anticodon in the tRNA initiator of the ternary complex. In this process, eIF5 plays an important role, as it guarantees the stop of scanning when initiation codon is detected [5, 12]. The recognition triggers the hydrolysis of GTP of ternary complex by eIF2, in a reaction facilitated by the GTPase-activating protein (GAP) eIF5. This results in the release of eIF2-GDP, since the complex in this state has ~10-fold less affinity for Met-tRNA$_i$, leaving it at the P site of the 40S ribosomal subunit. At the same time eIF1 and eIF5 are also dissociated [4, 16].

After detection of the AUG codon, the eIF5B-GTP complex binds to the 48S complex and promotes, in conjunction with other factors, the assembly of the large ribosomal subunit 60S, forming the 80S initiation complex. The eIF5B will also be important for stabilizing codon-anticodon pairing [4, 17]. The subsequent association of the ribosomal subunits leads to the hydrolysis of GTP by eIF5B and to dissociation of the eIF5B-GDP complex, due to its low affinity to the ribosome [16]. Once the 80S initiation complex is formed and the Met-tRNA$_i$ is located on the P site, the ribosome is ready to receive the appropriate aminoacyl-tRNA (aa-tRNA) on the A (aminoacyl) site and form the first peptide bond [9].

Next the translation initiation step, some of the components need to be recycled so that they are available for another translation round. One such case is the eIF2-GDP complex that is released from the ribosome and has to be reconverted into eIF2-GTP to reform the ternary complex. Since eIF2 has a high affinity for GDP, it is eIF2B that promotes the GDP:GTP change. On the other hand, the eIF2-GTP complex is not very stable, having to associate with the Met-tRNA$_i$, thus forming the ternary complex. As this step depends on a nucleotide exchanger and the availability of the components, translation initiation is limited by it [5].

Another of the interveners that must be recycled is the eIF4F complex, which has to be reassembled before each translation round. As already mentioned, this factor is crucial for the canonical initiation of the translation because it recognizes the cap structure and participates in the scan of the 5'UTR. The availability of eIF4F for translation is regulated by the mTOR (mammalian target of rapamycin) via. This pathway, when downregulated, decreases the degree of phosphorylation of various proteins, including 4E-BP (eIF4E-binding protein), which in their hypophosphorylated forms will bind to eIF4E. Once eIF4E is bound to 4E-BP, the reassembly of the eIF4F complex is inhibited. Consequently, this step is also a point of regulation, as the availability of eIF4F is dependent on a phosphorylation pathway [9].

In conclusion, there are two limiting steps in initiating translation: the recycling of the ternary complex and the reassembly of eIF4F complex [7, 14].

1.1.2. Elongation

As previously stated, elongation begins when Met-tRNA$_i$ is positioned at the ribosomal P site, and the A site is available to accept an aa-tRNA. Thus, the ternary complex composed by eukaryotic elongation factor 1A (eEF1A), a GTP molecule and an aa-tRNA can bind to the second codon of the open reading frame (ORF) located on the ribosomal A site. The linkage between aa-tRNA and mRNA is dependent on the hydrolysis of GTP and conformational changes in the small ribosomal subunit. This step will be quite important as it ensures that the cognate tRNA is selected and that there is a correct pairing between the codon and the anticodon. If the match between them is perfect, distortions occur in the ribosome to stabilize the interaction, and eEF1A-GDP frees itself from the complex leaving the aa-tRNA in the A site [18, 19]. The ribosomal peptidyl transferase center (PTC) then catalyzes the formation of a peptide bond between the aminoacyl-tRNA of the A site and the peptidyl tRNA of the P site. This reaction

causes a deacylated tRNA to remain in the P site, and therefore, the ribosomal complex must be translocated. The translocation of tRNAs is guaranteed by eEF2, which in an ATP-dependent manner promotes the passage of the deacylated tRNA to the E (Exit) site, the peptidyl tRNA to the P site and that the A site is free to accept the following aa-tRNA [6]. As in the translation initiation, the complexes containing GDP must be recycled to their active form by a guanine nucleotide exchange factor to be available for subsequent rounds of elongation. In the case of eEF1A-GDP, it is known that the guanine nucleotide exchange reaction is catalyzed by eEF1B. All these mechanisms repeat cyclically until the stop codon is found, which triggers the termination [6, 18].

1.1.3. Termination

Translation termination begins when a stop codon is recognized on the ribosomal A site. When this occurs, the binding between the nascent polypeptide chain and the tRNA of the P site is hydrolysed leading to the release of the polypeptide. In eukaryotes, this process is mediated by a termination ternary complex formed by eukaryotic release factor 1 (eRF1), eRF3 and a GTP molecule. Among these, it is the eRF1 that can decode any of the three stop codons - UAA, UAG or UGA. This factor will be important throughout the process because it promotes the stopping of translation and the hydrolysis reaction, catalyzed by the peptidyl transferase center of the ribosome [4, 20]. The eRF3, in a GTP hydrolysis dependent process, accelerates the release of the nascent peptide as it promotes conformational changes in the ribosomal complex [18].

1.1.4. Recycling

When the entire protein synthesis process is over, it is necessary to recycle the various components of the translation machinery, for them to be available for subsequent translation rounds. At this point in the process, the ribosomal 80S complex and the deacylated tRNAs remain linked to the mRNA [4]. In eukaryotes, one of the central proteins of the recycling is the ATP-binding cassette protein ABCE1, which is a highly conserved cytosolic ATPase that has already been observed to promote ribosomal recycling [21]. This protein belongs to the ABC-family ATPases, and it is proposed that it converts chemical energy from the ATP hydrolysis into mechanical energy that causes dissociation of the 80S ribosomal complex [18, 20, 22]. However, ABCE1 does not act alone and there are other proteins that aid in the process and stabilize the products of the dissociation. For example, eIF3 participates in the dissociation of the ribosomal 80S complex in 60S and 40S subunits and eIF1A stabilizes the ribosomal subunits, preventing the reassociation between the 40S and 60S subunits and/or the dimerization of the 40S subunit. In this way eIF1A also influences translation efficiency, since it regulates the availability of ribosomal subunits for the protein synthesis (Fig. 1) [21, 23]. Subsequent dissociation of deacylated tRNA and 40S subunit of mRNA is mediated by eIF2D, density-regulated protein (DENR) or initiation factors, such as eIF1, eIF1A and eIF3, which rescue the 40S subunit from another translation round [20].

1.2. Non-Canonical Translation Initiation Mechanisms

As already explained above, the canonical mechanism of translation initiation is that often occurs in the cell. However, under conditions that inhibit canonical translation initiation, there are proteins whose synthesis is maintained by alternative mechanisms, thus ensuring cell survival and stress response. Many of these mRNAs that are translated in a non-canonical way often encode for stress response proteins, growth factors or oncogenes, since their production must be maintained, either under stress conditions or disease states. Thus, non-canonical translation initiation mechanisms occur when there is ribosomal recruitment for the mRNA, with no recognition of the cap m^7G and/or scanning of the 5'UTR [5].

1.2.1. Cap-independent Translation Initiation Mechanism

Under any conditions all mRNAs in humans have a cap m^7G structure at their 5'end. However, in situations of high energy expenditure, such as mitosis, or during cellular stress, such as hypoxia, starvation, apoptosis, among others, recognition of cap m^7G may not occur during translation initiation. This inhibition can occur in two ways: via the mTOR pathway, whose downregulation results in hypophosphorylation of 4E-BP, which binds to eIF4E and prevents recognition of cap by this factor [24]; or by the α-subunit phosphorylation of eIF2, that inhibits the activity of its GEF eIF2B and causes the reduction of ternary complex levels available for translation initiation [25].

Thus, the initiation mechanisms of cap-independent translation occur when either one or both inhibitions occur. Among the mechanisms of cap-independent translation initiation, the most common and best described, is that in which there is a recruitment of the 40S ribosomal subunit to the mRNA, without the recognition of the cap structure, and in some cases scanning. This is guaranteed by specific sequences present in the 5'UTR of the transcript which are called Internal Ribosomal Entry Sites (IRESs). These IRESs can functionally replace the cap and (in some cases) proteins needed to recruit the ribosome to the start codon [5, 26].

For many years, it was thought that IRES-mediated translation was the only cap-independent initiation translation mechanism, yet two more have recently been described. Translation mediated by "cap-independent translation enhancers" (CITEs) or by N^6-methyladenosine (m^6A) [5]. The first case was described when it was found that there are structures in 5'UTR able to promote translation of this protein in the absence of m^7G-cap recognition, but also without the involvement of an IRES [27, 28, 29]. The other mechanism was observed in mRNAs containing N^6-methyladenosine (m^6A) residues in their 5'UTR, which apparently can bind to eIF3, which is sufficient for recruitment of the 43S initiation complex, in cases of suppression of the cap-binding factor eIF4E [30]. This shows that even under conditions of inhibition of the cap-dependent translation, some mRNAs have in their sequence the necessary tools to maintain the synthesis of the proteins that they codify, either through IRESs, CITEs or m^6A at 5'UTR.

1.2.1.1. Internal Ribosomal Entry Sites (IRESs)

In 1988, Pelletier and Sonenberg were among the first to describe a translation initiation mechanism mediated by IRES (Fig. 2). This process was observed in poliovirus mRNA, which is naturally uncapped, and is therefore translated via a cap-independent mechanism. Translation initiation was promoted by internal sequences within the 5'UTR, which were able to bind to the ribosome. This was shown by deletion mutagenesis of an internal sequence in the 5'UTR of poliovirus RNA that is required for cap-independent translation, which can also confer cap-independent translation to eukaryotic mRNAs [31]. These findings suggested that translation of this viral mRNA does not need the eIF4E cap-binding protein of eIF4F complex that normally recruits 40S subunits to capped 5'ends.

Since then, several viral mRNAs have been described as having IRES in their 5'UTR, many of them functionally different from each other. In the case of hepatitis C virus (HCV), the 40S ribosomal subunit together with eIF3 and the ternary complex binds directly to the IRES, without the action of the eIF4F complex. In this case, translation initiation is promoted by the secondary structure of IRES HCV, which, having high affinity for the ribosome, positions it near the initiation codon [16, 32, 33]. The IRES of picornavirus mRNA can bind to a conserved HEAT domain of eIF4G, which belongs to the complex to eIF4F, in the absence of cap recognition. In turn, this binding stimulates the linkage between eIF4F and eIF3, which promotes ribosomal recruitment [34, 35]. Another different example is found in Encephalomyocarditis virus (EMCV), in which non-canonical initiation factors alter the conformation of IRES and promote its binding to eIF4A/eIF4G. In turn, these factors will promote the binding of the 43S complex near the initiation codon. This type of IRES-mediated translation in EMCV means that there is no scanning or participation of eIF1, eIF1A and eIF4E [16].

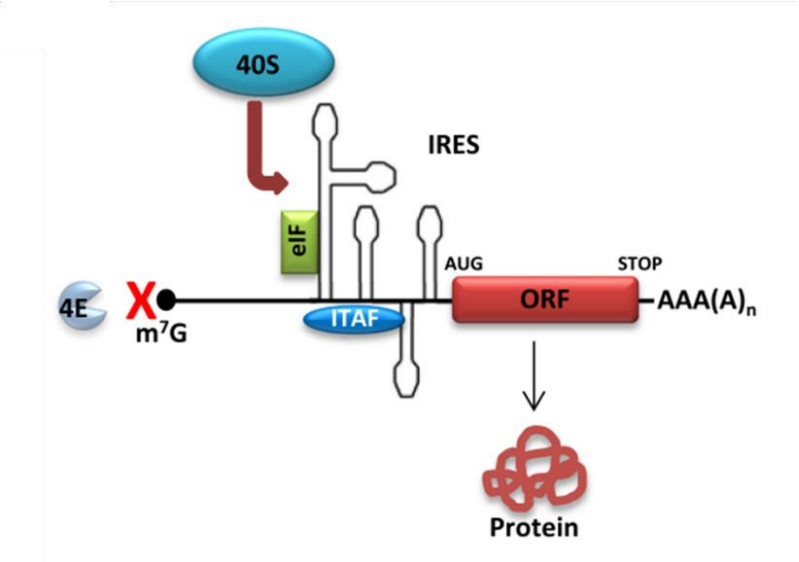

Figure 2. Model of translation initiation mediated by Internal Ribosomal Entry Site. IRES are structures located on the 5'UTR of the mRNA which can recruit the 40S ribosomal subunit into the vicinity of the start codon without recognition of the m^7G cap by eIF4E. This recruitment is often mediated by canonical eIFs and/or ITAFs. (Adapted from [5])

As IRES-driven translation is mediated by RNA itself, the structure of the IRES RNA is one of the main focuses for functional studies. In 2009, Lukavsky et al. demonstrated that HCV IRES folds into four major structural domains, with Domains II, III and IV assuming an essential role in translation initiation. Domain II promoted IRES activity, whereas Domain III (with its various a-f subdomains) was determinant in the interaction and recruitment of the 40S ribosomal subunit. The IV domain contained the AUG initiation codon [10, 33]. Therefore, more studies have been done with other viral IRES (reviewed in Kieft, 2008) but what was verified was that the majority of IRES have a complex secondary and tertiary structure, which influences the action of this to ensure an efficient interaction with the ribosome and/or protein factors [10, 26].

Yet, in most cases, the IRES sequence/structure alone does not guarantee ribosomal recruitment. Often it is necessary direct or indirect participation of canonical initiation factors or other auxiliary proteins - the IRES trans-acting factors (ITAFs). Depending on the IRES involved, the function of the ITAFs may change, seeing that they may stabilize a functional conformation of the IRES or act as a "bridge" between the IRES and the ribosome [36]. One of the ITAFs already described is the La protein that in situations of infection activate translation by IRES of the poliomyelitis virus. The majority of ITAFs are RNA-binding proteins which play a variety of functions in noninfected cells. For example, in the case of La, it is an autoantigen which induces antibody production in several autoimmune disorders. The mechanism by which La mediates IRES-translation is not known, however it has already been observed that this protein can bind to mRNA, and that the cap-independent translation of the poliomyelitis virus is La-dependent [37]. Another very common ITAF is polypyrimidine tract-binding protein (PTB), that in non-infected cells is usually associated with pre-mRNAs participating in the processing of these. It has been shown that PTB activates the IRES-mediated translation of various viruses, such as EMCV, poliovirus or hepatitis A virus [37]. More proteins have been described as mediators of viral IRES such as poly(rC)-binding protein 2 (PCBP2) for Coxsackie virus, Unr for rhinovirus or ITAF45 for foot and mouth disease virus (FMDV) [37]. Although the mechanism by which ITAFs potentiate the activity of viral IRES is still unknown, it has already been found that many of these proteins possess several RNA-binding domains or show a tendency for oligomerization. Owing to these properties, ITAFs are potentially capable of stabilizing or modifying the IRES fold, which can be considered as an RNA-chaperone activity [37].

Naturally, the identification of IRES elements in virus posed the question, whether there would be eukaryotic mRNAs that could be translated via a similar mechanism. In 1991, Macejak described that mRNA encoding immunoglobulin heavy-chain binding protein (BiP) can be translated via a cap-independent mechanism. In this work, it has been demonstrated, with the use of a bicistronic vector, that the 5'UTR of BiP can mediate the initiation of translation by an internal ribosome-binding mechanism in mammalian cells. In addition, it has also been found that the 5'UTR of the poliovirus can mediate a cap-independent translation in both infected and non-infected cells, suggesting that the cell itself has endogenous machinery to promote a non-canonical translation [38]. Since then, several mRNAs have been described as containing IRES in their 5' UTRs. It has even been estimated that about 10% of the cellular mRNAs can be translated through this cap-independent mechanism [39, 40].

In humans, one of the first IRES examples to be described and which functions as a model was mRNA encoding the X-linked inhibitor of apoptosis (XIAP), whose synthesis is maintained under apoptosis conditions by an IRES-dependent translation mechanism. In this specific case, it has been observed that the XIAP IRES supports translation during stress conditions, caused by the phosphorylation of eIF2α and consequent reduction in the levels of available ternary complex. Here, IRES-mediated translation initiation switches to an eIF5B-dependent mode to avoid attenuation due to eIF2α phosphorylation. Thus, the translation of XIAP under stress conditions is eIF2α-independent, but it is eIF5B-dependent, by a mechanism that is not yet understood [41]. Other mRNAs were identified, whose translation was maintained through IRES, such as Bcl2 (apoptosis inhibitor) [42], p53 (transcription factor/tumour suppressor) [43], CAT-1 (cationic amino acid transporter) [44] or c-Myc (transcription factor/oncogene) [45, 46]. These cellular mRNAs can maintain protein synthesis even when inhibition of translation initiation occurs due to the phosphorylation of eIF2α, which can be triggered by different stress conditions, such as starvation, heat shock, UV irradiation, hypoxia, endoplasmic reticulum stress, and virus infection. Interestingly, in the case of CAT-1 IRES, there is a need for eIF2α phosphorylation to occur a cap-independent translation [47]. This has demonstrated that there are different cellular mRNAs that can promote the initiation of translation through internal sequences even under conditions of low concentration of the ternary complex. However, in many cases the mechanism involved in ribosomal recruitment through IRES is still unknown.

As in many viral IRES, non-canonical translation in eukaryotes often requires the action of some canonical eIFs. For example, under hypoxia the synthesis of hypoxia-inducible factor-1α (HIF-1α) and vascular endothelial growth factor A (VEGF-A) is maintained by cap-independent translation in an eIF4G-dependent manner. In these cases, during hypoxia, the activity of the mTOR pathway is reduced, which conducts to hypophosphorylation of 4E-BP. As previously mentioned, the 4E-BP in this state, sequesters the eIF4E, suppressing cap-dependent mRNA translation. In its free form, eIF4G can bind HIF-1α or VEGF-A IRES and promote ribosomal recruitment through a mechanism not fully understood [48]. These observations suggest that different IRES-containing mRNAs might differ in their requirements for the active ternary complex and/or for some canonical factors, which cause them to use different pathway(s) to deliver the ribosome [36].

Although the list of cellular mRNAs containing IRES elements is growing, little is known about the mechanism by which cellular IRES elements capture 40S subunits [49]. Like the viral homologues, the cellular IRES present secondary and tertiary structure that allow an efficient interaction with the translating machinery (canonical initiation factors, ITAFs and 40S ribosomal subunit). Eukaryotic IRES are normally located at the 5'UTR immediately upstream of the initiation codon; they are relatively long and structured sequences, rich in GC, and which may contain upstream initiation codons. There are also cases where IRES extend to the coding region and contain the initiation codon itself. Cellular IRES also have stem loops and domains like viral IRES [36].

Little is known about the three-dimensional conformation that the eukaryotic IRES adopts for their activity, only realizing that there is a structure-function relationship, and there is no obvious conservation

of secondary structure among the various IRES. Currently, several studies have been done to determine the structure of several cellular IRES, often using chemical and enzymatic probing, coupled with computer predictions [50]. In one case, Le Quesne et al. determined by chemical probing the structural model for the human c-Myc IRES. Their results suggest that the IRES is divided into two structural domains linked by a long unstructured region. It was verified that the presence of the two domains is necessary for the full functioning of the IRES, however it was verified that specific deletions do not affect its action, but only diminish its activity. Similar structures and modes of action have been described for N-myc and L-myc IRES [51, 52]. In addition, was determined/approximated IRES structures of insulin-like growth factor II (IGF-II) [53], VEGF [54] or mTOR [55], among others.

It is interesting to note that small deletions and point mutations have a more toxic effect on the structure and function of viral IRES than on cellular IRES. These observations imply that the structure-function relationship is not rigid in these cellular IRES as in the viral IRES [50]. In this way, eukaryotic IRES appear to be less structured and have more plasticity than viral IRES, which presupposes that translation through this mechanism needs more action from ITAFs that "aid" the canonical eIFs [47]. Already several proteins have been described as functioning as ITAFs. Some are proteins that are already associated with mRNA, since they participate in cellular processes upstream (ex: splicing, export), others only bind to the mRNA at the time of translation [49].

Very little is understood about the exact mechanism of ITAFs, knowing only that they function as RNA chaperones. As previously mentioned, the most common and most studied ITAF in both viruses and eukaryotes is PTB. This protein in addition to its role in IRES-dependent translation, it also influences the splicing, stability and localization of mRNA [56]. It is ITAF of a large heterogeneity of IRES, however it is curious to verify that PTB is ITAF of most IRES that are active during apoptosis, such as c-Myc, PDGF (growth factor) or TNFα-R (receptor) [57]. In addition to this condition, PTB was described as the ITAF of CAT-1 mRNA under conditions of amino acid starvation [58], early growth response 2 (EGR2) mRNA under conditions of inflammation [59], insulin mRNA under conditions of nitrosative stress [60] or p53 under DNA damage conditions [61]. In this way it is possible to verify that PTB is a prominent ITAF in many cellular IRES in various stress conditions, however without realizing its mechanism of action in many of them [62].

Other proteins have already been identified as ITAFs of cellular IRES such as, La autoantigen, which has been shown to be necessary for XIAP cap-independent translation [63], the Unr which has been described as the ITAF of the IRES of PITSLRE kinase [64] or DAP5 which is ITAF for Bcl2 IRES and for IRES of its own mRNA [65].

However, ITAFs may also be translation inhibitory elements and may act as negative mediators of IRES functioning [56]. It has been demonstrated that PTB functions as a negative regulator of Bip IRES-dependent translation. It was found that the inhibition of endogenous PTB enhanced the translational initiation directed by Bip IRES happening the opposite when this protein is overexpressed. It is unclear how PTB inhibits the translation of the Bip mRNA, only knowing that the inhibition is promoted by the interaction between the PTB and the central region of the 5'UTR of the Bip mRNA [66]. This shows that some ITAFs, depending on the mRNA in question, may function as a positive or negative regulator of an IRES-dependent translation.

PTB, like many of the ITAFs already described, belongs to the group of heterogeneous nuclear ribonucleoproteins (hnRNP A1, C1/C2, I, E1/E2, K and L) known to shuttle between the nucleus and the cytoplasm. Therefore, it is proposed that these proteins regulate the functioning of IRES by their availability. Several studies of overexpression and/or depletion of specific ITAFs have demonstrated that cellular IRES activity is modulated by the intracellular concentration of these proteins. Thus, it is suggested that the subcellular location of the ITAFs (nucleus or cytoplasm) influences the functioning of IRES and that this is also a form of regulation of this process [36, 47, 67].

For example, the translation via IRES of the transcription factor SREBP-1 (sterol regulatory element-binding protein 1) is mediated by the ITAF hnRNP A1. It has been shown that under stress conditions in liver cells the hnRNP A1 mediates the activity of IRES and allows maintenance of the protein levels of SREBP-1. In addition, it has been found that in normal conditions SREBP-1 is translated in a cap-dependent manner and the hnRNP A1 is in the nucleus unbound to the SREBP-1 5'UTR. However, the induced endoplasmic reticulum stress triggers the cytosolic relocation of hnRNP A1 and causes the increase in hnRNP A1 bound to the SREBP-1 5'UTR [68].

Other reports have also emphasized the importance of subcellular relocation for ITAF and IRES function, which demonstrates a putative mechanism of regulation utilized in IRES-dependent translation mechanism. Still, there are other proposed mechanisms for ITAFs that are not delocalized during cap-independent translation. Some exist in the cytoplasm and only bind to the mRNA at the time of translation or others that are already linked to the mRNA because they also participate in processes upstream of the translation [67].

Alternative translation initiation mechanisms play a key role in some cellular processes when cap-dependent translation is compromised, allowing the cell to maintain its viability regardless of environmental changes. Several cellular mRNAs can be translated in a cap-independent manner, however it is curious to verify the heterogeneity of proteins that can be synthesized through this mechanism. Therefore, Lacerda et al. (2017) compiled by functional classes the proteins whose transcripts possess IRES in order to realize which can be synthesized by this cap-independent mechanism. It has been shown that among mRNAs that have IRES, 21% encodes for transcription factors, 15% for growth factors or 22% for transporters, receptors or channels. This shows that IRES-containing mRNAs often code for proteins that play an important role in maintaining cellular homeostasis and stress response by participating in various regulatory, cell growth, proliferation or differentiation processes [5, 69].

In contrast the deregulation of IRES-mediated translation leads to the development and progression of disease, since these gene families have pivotal roles in cellular processes that demand fine-tuned regulation. Additionally, changes in expression patterns may lead to the synthesis of oncoproteins in various types of cancer. Since the production of many of these oncoproteins is maintained by an IRES-mediated translation this leads to the conclusion that this mechanism operates in conditions such as mitosis, hypoxia, DNA damage, osmotic shock, starvation and apoptosis, which are characteristics of a tumorigenic situation [5, 70].

1.3. Translation Regulation

Gene expression is regulated at multiple levels, including the translation of mRNAs into proteins, that it is subjected to extensive points of regulation. In bacteria, control of gene expression occurs mostly at the level of transcription, while in eukaryotic systems about 30% of the proteins produced are predominantly regulated at the translational level [10]. Regulation at this stage has the advantage of producing a faster response of the cell, i.e. conversion of an "inactive" mRNA into an "active" mRNA without the need for transcription, splicing or transport. This type of control is very useful in various conditions of cellular stress - heat shock, hypoxia, nutrient, deprivation - as it guarantees immediate changes in protein levels [10, 71]. In fact, during various cellular events (ex: mitosis, apoptosis) or in response to most cellular stresses the overall translation is reduced. This action allows to prevent the synthesis of proteins unnecessary for the survival of the cell or that interfere in the response to stress. Effectively, the regulation of protein synthesis occurs at the level of initiation, elongation, termination and recycling, however it is commonly considered that the initial step is the most controlled and the limiting step of the whole process [7, 9, 71].

One of the control mechanisms of translation initiation is the availability of the ternary complex, making the recycling of eIF2-GDP into eIF2-GTP a limiting step in this process. Several types of stress can inhibit this step by inducing the phosphorylation of eIF2 on Ser51 of α-subunit. This causes eIF2α-

Ph function as a competitive inhibitor of eIF2B, which as previously stated is a guanine nucleotide exchange factor, which catalyzes the passage of eIF2-GDP to eIF2-GTP. Consequently, the assembly of the ternary complex and the initiation of the translation will be inhibited [7, 72]. In mammals there are four different kinases capable of phosphorylating eIF2α, which are activated by different stresses. They are PKR (double stranded RNA in virus infection), PERK (unfolded proteins in the endoplasmic reticulum), HRI (heme deprivation), and GCN2 (amino acid starvation), which phosphorylate the same residue in eIF2α and, hence, elicit the same "integrated stress response" involving downregulation of general translation [3, 9].

The second mechanism in eukaryotes that occurs to control the initiation of translation involves the recognition of the 5'cap by the eIF4F complex, more specifically eIF4E. As previously stated, the eIF4F assembly is inhibited by 4E-BP because it competes with eIF4G for the eIF4E binding site. In turn, the binding of 4E-BP to eIF4E is regulated by the degree of phosphorylation of the first. If hypophosphorylated, 4E-BP binds strongly to eIF4E, whereas phosphorylation of 4E-BP weakens their interaction with eIF4E. The phosphorylation of 4E-BP is controlled by the mTOR protein, which is the downstream Ser/Thr kinase in the PI3K/Akt signaling pathway, which senses and integrates signals from extracellular stimuli, amino acid and oxygen availability, and energy status of the cells [7, 9]. mTOR may also indirectly inhibit the assembly of eIF4F by suppressing eIF4A activity. In this case the mTOR phosphorylates the S6 kinase (S6K), which in turn phosphorylate PDCD4, which is a tumor suppressor that binds eIF4A inhibiting it [9, 73].

In these situations, translation initiation is controlled by trans-elements, which is much better characterized than that which is made by functional cis-regulatory elements [9]. However, there are structural aspects in eukaryotic mRNAs that are important for the regulation of translation initiation. Some of them are the presence of cap m^7G, the context around the start codon and the position of this in the transcript, the existence of upstream ORFs (uORFs) or the structural composition of 5´UTR [12, 72, 74, 75, 76].

1.3.1. IRES-mediated translation in health and disease

The discovery of IRES in cell transcripts has introduced a new character to understand the expression of some proteins in certain environmental conditions. IRES elements have emerged as important regulators of selective mRNA translation, under conditions of reduced cap-dependent translation. The question of the biological relevance of IRES is answered by the conditions by which this mechanism occurs as starvation, mitosis, hypoxia, endoplasmic reticulum stress, among others. Thus, internal initiation represents a cellular backup plan for survival under these conditions [47]. However, such conditions are also frequently experienced by cancer cells whose survival relies on IRES-dependent translation of key pro-angiogenic, hypoxia-response and survival mRNAs [36]. The complex nature of regulation of cellular mRNA translation under different pathophysiological conditions suggests that there may be several diverse pathways leading to cellular IRES-mediated initiation.

One example occurs during angiogenesis, in which hypoxic stress is induced in endothelial cells to activate certain signalling pathways, to reduce levels of canonical translation. During this process the IRES-mediated translation plays a key role since it maintains the expression levels of VEGF which is the main character of this process as it stimulates the proliferation and migration of endothelial cells. It has been described that during angiogenesis the VEGF mRNA translation is cap-independent and is mediated by two independent internal ribosome entry sites that are present in 5'UTR (IRES A and IRES B). In addition, hypoxia results in rapid inhibition of translation by eIF2α phosphorylation by PERK and the dephosphorylation of 4E-BP, inhibiting eIF4E. This will cause the increase of free eIF4G, which in turn mediates the cap-independent translation of VEGF. This demonstrates that IRES-mediated translation can maintain VEGF levels under hypoxia conditions, functioning as a compensatory mechanism during

angiogenesis. In this process the PTB will function as positive ITAF, interacting with the VEGF IRES [54, 77, 78].

In contrast, IRES-mediated translation also occurs in angiogenesis during tumour development and progression, which is characterized by low oxygen concentration environments. Although VEGF itself is not an oncogene, it is upregulated in tumorigenesis and is important in blood vessel formation in solid tumours. In breast cancer is induced an inflammatory state that causes in the tumour cells a hypoxic-stress response. This will lead to the downregulation of the mTOR pathway and, consequently, to increased levels of hypophosphorylated 4E-BP which sequesters eIF4E, inhibiting cap-dependent translation. Thus, levels of free eIF4G increase what induces translation through IRES of VEGF mRNA, which as previously seen is an eIF4G-dependent mechanism [36, 48, 71]. Notably, the VEGF IRES activity was higher in metastasizing tumor cells in lymph nodes than in primary tumors, most likely because lymph vessels in these lymph nodes were severely hypoxic [79]. Other mRNAs are translated via IRES during hypoxic stress, such as HIF-1α [80] or ATF4 [81], and it has already been demonstrated that PTB plays a stimulatory role in the IRES-mediated translation of HIF-1α when oxygen supply is limited [82].

Regulated IRES-mediated translation events have also been shown to play important roles in controlling the progression of cells through mitosis. During the cell cycle, more specifically, in the G_2/M transition often the cap-dependent translation is inhibited, causing the synthesis of many proteins required for this process to be maintained from IRES-mediated events. One such case is PITSLREp58, which is involved in spindle formation. In this case, the inhibition of cap-dependent translation by mitosis causes the cytoplasmic relocation of the ITAF hnRNP C1/C2, which stimulates IRES-mediated translation of another ITAF, Unr. Unr, in turn, is required for enhanced IRES-mediated translation of PITSLREp58. Therefore, progression through mitosis occurs via regulation of IRES-mediated translation involving changes in the relative cytoplasmic levels activities of positive regulatory ITAFs. The exact mechanism leading to these changes is unknown [83]. Other IRES elements have been described as present in the mRNAs of proteins of different phases of the cell cycle, such as LEF-1 (lymphoid enhancer-binding factor 1), which is involved in cell growth during mitosis [77] or ODC (ornithine decarboxylase), that affect chromatin organization [84]. These findings support the possibility that cell cycle progression signalling events influence ITAF/IRES activity and function [36].

The nutrient availability can also be regulated by cap-independent translation initiation mechanisms, there being some examples such as the insulin receptor mRNA and some amino acid transporters which have IRES. Spriggs et al. (2009) demonstrated that insulin receptor mRNA contains IRES in its 5'UTR, and that it can be translated through this pathway. The authors observed that IRES activity was stimulated by insulin, suggesting that IRES-mediated translation in vivo might be regulated in a manner dependent on blood glucose levels. In addition, IRES-mediated insulin receptor synthesis has been shown to be more prominent in the brain and in neuronal cell lines compared to other tissues and cell lines. This may indicate that IRES translation is important for maintaining INR expression, and other mRNAs containing IRES in tissues with reduced cap-dependent translation, such as a brain. In addition to these results it has been demonstrated that the PTB is required for the function of the human INR IRES both in vitro and in vivo [85]. The nutritional control modulated by IRES activity also occurs at the level of regulation of amino acid availability. Amino acid depletion induces GCN2 kinase-mediated phosphorylation of eIF2α, leading to a global decrease in canonical protein synthesis. This will lead to IRES-mediated translation of various amino acid transporters mRNAs, such as CAT-1 or SNAT2, which are required to promote recovery of amino acid balance [36, 86, 87, 88].

Many other cellular events and/or stresses induce cap-independent translation, such as HIAP2, that has an ER-stress-inducible IRES, c-Myc, whose IRES translation is activated following the induction of apoptosis or mitosis, or NRF2, whose synthesis is promoted by an IRES element during oxidative stress [71, 77, 89, 90, 91, 92, 93].

However as seen from some previous examples, the abnormal functioning of IRES-mediated translation may play an important role in tumorigenesis. Many transcripts with relevance in cancer, such oncogenes, growth factors and proteins involved in the regulation of programmed cell death are translated via IRES elements, under stress situations induced within the tumour microenvironment.

It has been shown that under hypoxic conditions, IRES-mediated translation promotes cell survival and formation of tumour emboli in inflammatory breast cancer (IBC). Very high levels of hypophosphorylated 4E-BP and free eIF4G were observed in these tumour cells. On the other hand, excess eIF4G promoted cancer cell survival and formation of IBC tumour emboli by enhancing translation by IRES of VEGF and p120 catenin which are responsible for maintaining high rates of angiogenesis and membrane associated E-cadherin during emboli formation, respectively [48, 94].

Another study demonstrated that ovarian cancer cells and ovarian cancer xenografts can survive through IRES-mediated translation, being resistant to inhibition of PI3K/mTOR-pathway. In this work, drug treatment has been shown to induce resistance in these cells, causing upregulation and/or activation of multiple prosurvival proteins, including several receptors (EGFR, HER2, c-Kit, and IGF1R), cytoplasmic kinases (p-p90RSK, p-SrcY), antiapoptotic proteins (Bcl-2, XIAP1), and transcription factors (p-STAT3, p-STAT6, pc-Jun, p-SMAD3) [95].

One of the proteins that is most deregulated in tumour cells is the p53 tumour suppressor [96]. This is a transcription factor that controls the expression of protein coding genes as well as micro-RNAs (miRNAs). It plays a critical role in cellular responses to DNA damage and other stresses by inducing cell-cycle arrest and programmed cell death. In the opposite field, p53 mutations contribute to tumorigenesis [70]. The p53 transcript has two IRESs that control the translation of two isoforms of the protein, the full-length p53 and an N-terminally truncated isoform (Δ40p53), both of which are enhanced in different stress conditions that induce DNA damage, endoplasmic reticulum stress or cancer. A mechanism that links IRES-mediated p53 translation with tumorigenesis has recently been described. For this, two novel p53 ITAFs were identified, the translational control protein 80 (TCP80) and RNA helicase A (RHA), which positively regulate p53 IRES activity. In addition, it has been shown that in two breast cancer cell lines the levels of TCP80 and RHA are extremely low, causing the reduction of IRES-dependent translation of p53 and a defective response following DNA damage. These findings reveal a mechanism of p53 inactivation that links deregulation of IRES-mediated p53 translation with tumoral cells [97, 98].

c-Jun is an oncoprotein whose cap-independent translation is required for tumor progression. c-Jun is a component of activator protein 1 (AP-1), which is a transcription factor involved in regulation of proliferation, differentiation, growth, apoptosis, cell-migration and transformation. Among others, c-Jun stimulates transcription of components of the cell cycle, repress transcription of tumor suppressor genes such as p53, and induces expression of metalloproteinases, which are proteolytic enzymes that promote growth, invasion, and metastasis of cancer cells [99, 100]. IRES-mediated translation of c-Jun can be induced by loss of cell-cell contacts, such as when there is a loss of E-cadherin, which causes disruption or restructuring of the cytoskeletal network. Disruption of the cytoskeletal network activates a signalling pathway that upregulates IRES-mediated translation of c-Jun and induces an invasive program. Thus, IRES-mediated translation of c-Jun likely plays an important role in tumour progression [70, 101, 102]. In this way, high c-Jun protein levels have been observed in glioblastoma, malignant melanoma, invasive breast cancer, and colorectal cancers [103, 104].

Other proteins have been described as upregulated during tumour maintenance and progression, through IRES-mediated translation which allows to conclude that cap-independent translation has an important role for the survival and proliferation of cancer cells under stress conditions, thus contributing to the process of tumorigenesis. Thus, the biological implications of the different initiation mechanisms are clearly important. The stress conditions that cells must respond to - such as hypoxia, starvation, toxins or drug exposure - are often an underlying cause of human diseases, including cancer. Understanding how

translational control is involved in cellular response to stress will provide insight into many human disorders and ultimately lead to the development of new therapeutic modalities.

1.4. Role of Argonaute proteins, more specifically AGO1 in the cell

Argonaute (AGO) proteins are key elements in gene silencing processes by small-RNAs. These proteins interact mainly with microRNAs (miRNAs) or short interfering RNAs (siRNAs) that guide Argonaute proteins to target mRNA molecules for silencing or degradation. These proteins are then involved in post-transcriptional gene silencing [105]. Argonaute family proteins are highly conserved and their members are found in all eukaryotes, being phylogenetically and functionally very close to the family of Piwi proteins [106].

The Argonaute family was first identified in plants, being highly conserved between organisms and species [105, 107]. In humans there are eight Argonaute genes, four that encode for four Piwi proteins (HIWI1, HIWI2, HIWI3 and HILI), and four others that encode for four Argonaute (1-4) proteins, with the AGO1, AGO3 and AGO4 genes being clustered on chromosome 1, whereas the AGO2 gene is located on chromosome 8 [105].

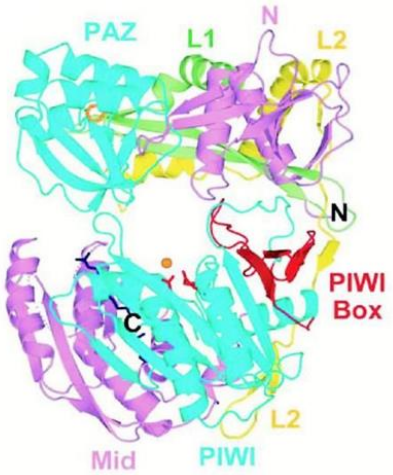

Figure 3. Structure of Argonaute protein. X-ray crystal structure of the Argonaute protein from the archaeon Aquifex aeolicus. Argonaute is divided in two lobes, one containing the amino-terminal domain (N, magenta) and the PAZ domain (blue), and the other containing the MID domain (magenta) and the Piwi domain (blue). Linker 1 (L1, green) bind N domain to PAZ domain and Linker 2 (L2, yellow) connects the PAZ domain with the MID domain. The various domains play different roles during Argonaute action. PAZ and MID domain anchor, respectively, the 3' and 5'end of the miRNA. The N domain is responsible by the unwind of the small RNA duplex during the RNA-induced silencing complex (RISC) assembly. The domain Piwi adopts an RNase H fold indicating that Argonaute proteins are responsible for the 'slicer' activity of the RISC. However, only AGO2 has catalytic activity. (Adapted from [105])

In humans, the Argonautes are preferentially involved in post-transcriptional gene silencing by destabilizing the mRNA or repressing the translation, highlighting its role in embryonic development, cell differentiation and tumorigenesis [105]. Of this family only AGO2 has endonucleolytic activity, whereas AGO1, AGO3 and AGO4 are non-catalytic proteins [108].

These proteins are characterized by having four domains: amino-terminal (N), PAZ, MID (middle) and Piwi (Fig. 3). Most of these structural studies were done on bacteria and archeas, however there is evidence that there are similarities with higher organism due to the high conservation of these proteins and the basic principles of small RNA pathways [109]. Recently, crystal structures of human AGO2 have been reported, being that it was verified that there are elevated appearances between this one and the homologous ones of bacterium and archea [110]. Each one of the domains of Argonautes has different functions, highlighting the Piwi domain which adopts an RNaseH fold suggesting that Argonaute proteins

are responsible for the 'slicer' activity of the RNA-induced silencing complex (RISC). Yet, only AGO2 has catalytic activity [111, 112].

The Argonautes are just one element within a large process that is the biogenesis of small RNAs, being an integral part of the RISC. These proteins come into play when a mature miRNA duplex is already present in the cytoplasm. In the next step, the duplex is unwound in a reaction mediated by Argonaute N domain, and the passenger strand of the miRNA is degraded. One important determinant for guide strand selection by AGO lies in the small RNA duplex itself, whereupon the strand with the less stably paired 5'end is preferentially loaded into AGO proteins. The mature miRNA strand is subsequently incorporated into the RNA-induced silencing complex, where it binds directly to a member of the AGO protein family [108]. Next, the miRNAs conduct the RISC, to the target mRNAs, serving Argonaute as a platform that facilitates the binding between the miRNA and the mRNA. When base complementarity binding between the miRNA and the target occurs, the mRNA is silented or cleaved (Fig. 4) [105].

Often in research work "takes advantage" of the RNA interference (RNAi) mechanism to silence genes of interest. In this case siRNAs (short double stranded RNAs) are introduced into the cell by transfection and are soon processed by the Dicer without nuclear processing, linking directly to the RISC.

It is not clear whether there is a preference for binding of the different miRNAs to the different Argonaute proteins, since several miRNAs bind to different Argonautes. However, it has already been found that several small RNA classes possess specific sequence at the 5'end, and it has become clear that Argonaute proteins are able to sense the 5'terminal nucleotide of the small RNA. It has recently been shown that the rigid loop in the MID domain of human AGO2 allows specific contacts with a 5'terminal uridine or adenine, while small RNAs with Gs or Cs at the 5' end bind to human AGO2 with low affinities [113]. Recently, Werfel et al. demonstrated by immunoprecipitation and in vivo competitive binding assays that miRNAs-21, -199-3p and let-7 bind preferentially to AGO2 [114]. Another example is the case of let-7a-3p which is normally activated by AGO3 [115].

In recent years, novel cellular Argonaute functions are emerging, including roles for Argonaute in transcription, alternative splicing and even DNA repair. In several organisms it has been observed that the AGO proteins are involved in epigenetic alterations and that they cause modifications in the chromatin. In plants it was established that AGO proteins, such as AGO4, are associated with nascent transcripts and mediate DNA methylation and heterochromatin silencing [108, 116]. In mammals there is still little data to show that Argonautes mediate the transcriptional gene silencing. RNAi machinery has also been implicated in alternative splicing in human cells. In this study in HeLa and hepatoma cells it was shown that siRNAs targeting intronic or exonic sequences close to an alternative exon, which regulate the splicing of that exon. Additionally, the authors verified that this effect is dependent on AGO1 [117]. Finally, AGO proteins have also been observed in double-strand break repair mechanisms. As for this process very little is known, however, it has already been observed in human HeLa cells, that small RNAs are produced from the vicinity of the double strand breaks. These then associate with Dicer and AGO2, which mediates double-strand break repair, suggesting an active role of these factors in this process [118].

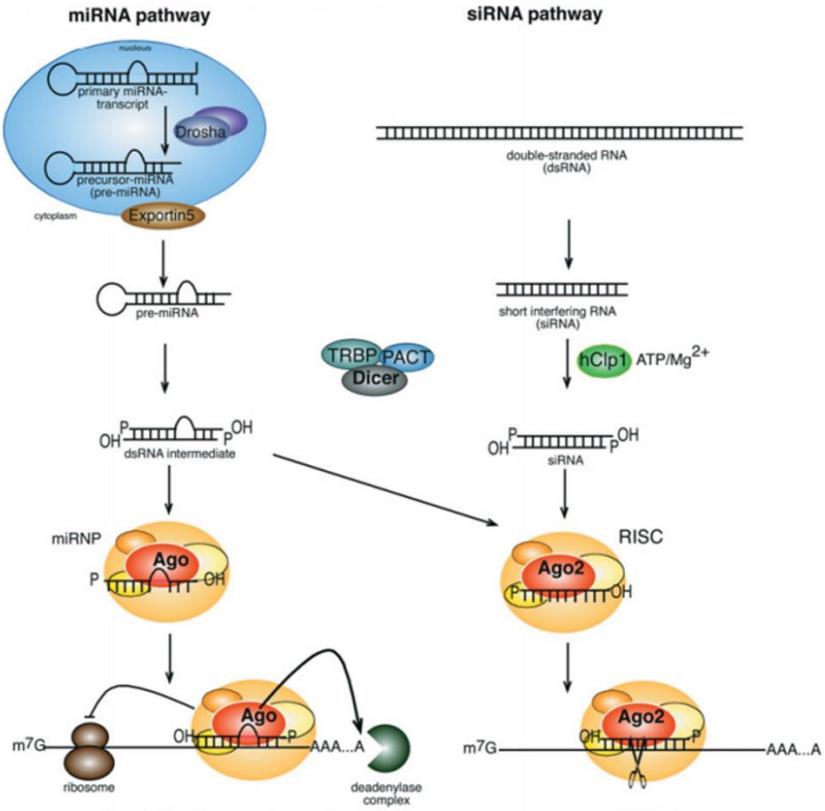

Figure 4. Schematic representation of the biogenesis of microRNAs and short interfering RNAs. The precursors of the miRNAs are synthesized in the nucleus and processed by the Drosha complex and are subsequently exported to the cytoplasm where they are associated with the Dicer complex, which cleaves the hairpin and generates a mature miRNA. This is associated to the RISC, through the Argonaute protein, and guides this complex to the target mRNA. If AGO2 is in question the mRNA is cleaved, otherwise the RISC may recruit exonucleases or otherwise block translation, transiently. The processing of the siRNAs is similar, however they do not undergo a nuclear experiment, linking to the Dicer that delivers them to the RISC. (Adapted from [119])

Argonaute family proteins are ubiquitously expressed and are mostly located in the cytoplasm often associated with P-bodies or stress granules [105, 120, 121]. Argonaute (1-4) proteins are expressed in many tissues and have also been identified in several cell lines. Völler et al. have demonstrated that in many tissues, such as the kidney, urinary tract, thymus or fetal kidney, AGO1 and AGO2 are the most expressed proteins within argonaute family, whereas AGO4 is generally expressed in low amounts in most tissues. AGO3 has a more variable expression between tissues. Interestingly, AGO2 is the protein of this family, whose expression is more constant between tissues and cell lines, which points out its importance in the cell, since it is the only one with catalytic activity [122].

Argonaute-mediated gene silencing has an impact on several cellular and physiological processes. It has been described in several organisms that argonaute proteins participate in the development of several tissues, and it is therefore proposed that they have a role in the stem cell fate. Several Argonaute proteins appear to be involved in stem cell fate decisions, and mutations in these genes cause stem cells to lose their character and differentiate instead of undergoing self-renewing division. For example, drosophila Argonaute 1 is essential for normal development, particularly in the nervous system. Mutations in AGO1 cause embryonic lethality marked by a severe decrease in all types of neurons and glial cells, a phenotype that suggests may arise from a defect in cell cycle progression or cell survival [123]. A mouse study showed that the depletion of AGO1 and AGO2 of the global expression of microRNAs is significantly compromising and causes severe defects in skin morphogenesis [124].

Particularly in relation to Argonaute 1, it is present in humans on chromosome 1, being encoded by the EIF2C1 gene. This protein has a similar function to the other Argonautes in post-transcriptional gene silencing, showing no catalytic activity to cleave the target mRNAs. Structural comparisons between AGO1 and AGO2 from human demonstrated that there is 84% identity as to the primary sequence of

proteins. Faehnle et al. found that the lack of AGO1 slicer activity is due to structural changes in Piwi domain loops compared to AGO2, once Piwi domain is the main responsible for the catalytic activity of Argonaute [125]. Little is yet known about the specific function of AGO1 in relation to the other Argonautes, it seems often that the four proteins have redundant functions. However, some studies have been done to understand if AGO1 is involved in the biogenesis of specific miRNAs [126].

Notably, AGO1 is expressed at low levels in most tissues, but its expression is particularly high in embryonic kidney and lung. Protein levels are also increased in tumours that lack the Wilm's tumour suppressor gene WT1 [106].

Additionally, it has been described that AGO1 is overexpressed in colorectal cancer when compared to adjacent non-cancer tissue. It has also been found that AGO1 in neuroblastoma cells may play a role of tumour suppressor, since its overexpression causes the cell cycle to slow down, decrease in cellular motility and stronger apoptotic response upon UV irradiation [127, 128]. Thus, these studies support the hypothesis that Argonaut deregulation may lead to defects in the RNA interference machinery, which in some cases may play an important role in tumour progression.

2. Previous Results

The research developed in this master's thesis was based on the results obtained by Rafaela Lacerda Santos in her PhD thesis – Non-canonical translation initiation of proteins with potential relevance in colorectal cancer (2016).

In this work, it was hypothesized that the Argonaute proteins, more specifically AGO1, can be translated through mechanisms of cap-independent initiation under conditions of inhibition of canonical translation. Argonaute are involved in the synthesis of miRNAs, which in turn are involved in the pathways of stress response and maintenance of cell viability. Thus, it was proposed that the synthesis of these proteins can be maintained under conditions of cellular stress by a cap-independent mechanism.

To test the initial hypothesis, Lacerda Santos first did an *in silico* analysis to understand if the AGO1 mRNA had characteristics that favour the existence of a non-canonical translation initiation. A thorough analysis of AGO1 5'UTR revealed an overall GC content of 72.3%. This high GC content indicates that it is likely to fold into elaborate RNA secondary structures, which in several cases promote the cap-independent translation.

To prove this experimentally it was used a bicistronic reporter vector with dual-luciferases, which constitutes the most common method to test cap-independent translation activity [129]. The particularity of this system is that he first cistron, the *Renilla* Luciferase (RLuc) ORF, is translated by a cap-dependent mechanism, and the second cistron, the Firefly Luciferase (FLuc), if translated, will be by a cap-independent mechanism. To prevent translation reinitiation, a stable hairpin was inserted downstream of RLuc [130]. It was subsequently cloned the 5'UTR of the AGO1 upstream of the FLuc ORF to mediate cap-independent translation of this reporter. The positive controls were the cellular c-Myc and the EMCV IRES sequences cloned in the same plasmid [45, 131]. As negative controls, it is used the empty plasmid (pR_F) and the human β-globin (HBB) 5' UTR cloned in the same plasmid (Fig. 5).

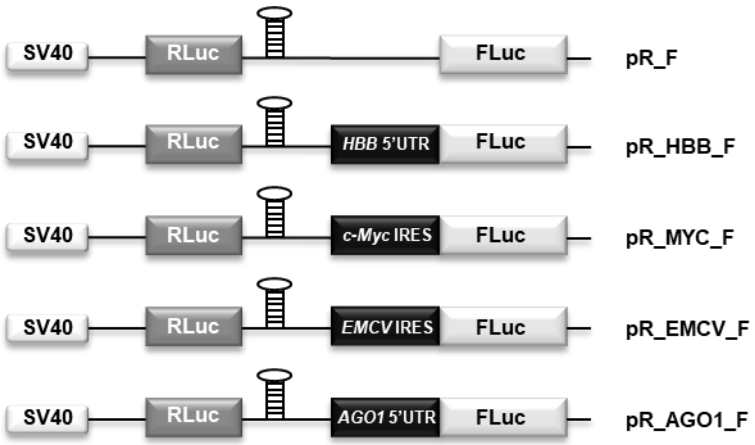

Figure 5. Schematic representation of the constructs used to check whether AGO1 5'UTR is able to drive cap-independent translation initiation. RLuc is the Renilla luciferase cap-dependent translated cistron (grey box) and FLuc the firefly luciferase cap-independent translated cistron (white box). SV40 box represents the promoter. Black boxes represent the different sequences cloned upstream FLuc ATG: empty vector (pR_F); human β-globin 5'UTR (HBB) is the negative control for cap-independent translation initiation activity (pR_HBB_F); c-Myc IRES is the cellular positive control for cap-independent translation initiation activity (pR_MYC_F); EMCV IRES is the viral positive control for IRES activity (pR_EMCV_F); AGO1 5'UTR (pR_AGO1_F). All constructs contain a stable hairpin (represented by a stem loop) downstream RLuc of the cistron to prevent translation reinitiation.

Then, HeLa cells were transfected with each one of the plasmids shown in Fig. 5 and luminometry assays were performed to measure the relative FLuc/RLuc ratio from each construct. The obtained results show that relative FLuc/RLuc luciferase expression levels from pR_AGO1_F are 2.8 fold those from the empty plasmid, pR_F. As for the relative FLuc/RLuc luciferase expression levels from pR_HBB_F, they are similar to those from pR_F, whereas those from pR_MYC_F and pR_EMCV_F are significantly greater than those from the empty plasmid: 5.8 and 13.4 fold the expression levels from the empty plasmid, respectively (Fig. 6). These results show both positive controls are driving FLuc expression via a cap-independent mechanism, which is in concordance with previously published data [45, 131]. From this experiment, we can also conclude that AGO1 5'UTR is able to drive FLuc expression in a bicistronic context, which suggests a non-canonical mechanism of internal translation initiation may be responsible for such expression. Such activity is, however, lower than that measured from c-Myc IRES-containing plasmid (Fig. 6), indicating that the mechanism by which AGO1 5'UTR is driving cap-independent translation is less efficient than that used by cellular c-Myc IRES [89].

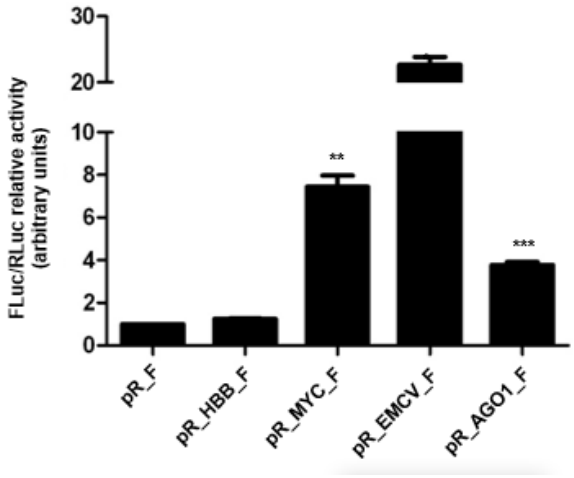

Figure 6. AGO1 5'UTR in a bicistronic context is able to mediate cap-independent translation of FLuc reporter protein. HeLa cells were transfected with pR_F, pR_HBB_F, pR_MYC_F, pR_EMCV_F and pR_AGO1_F plasmids. Relative luciferase activity is shown as the luminescence ratio between FLuc and RLuc normalized to that obtained from the pR_F, arbitrarily set to 1. Presented data are the result of, at least, three independent experiments. Statistical analysis was performed using the Student's t test (unpaired, two-tailed). *P<0.05, **P<0.01, ***P<0.001.

It was also important to understand whether AGO1 5'UTR could mediate FLuc cap-independent translation under stress conditions. With this goal in mind, HeLa cells were transfected with either pR_AGO1_F or the controls (pR_F, pR_HBB_F, pR_MYC_F and pR_EMCV_F) and treated with 1 μM of thapsigargin, which is a drug that induces the stress of the reticulum, or the corresponding vehicle (DMSO), and then was measured relative luciferase activity (Fig. 7). In transfected cells treated with 1 μM of thapsigargin, Western blot analysis shows an increased amount of phosphorylated eIF2α protein, indicating that there was an inhibition of cap-dependent translation (Fig. 7a). In these conditions, relative FLuc/RLuc luciferase expression levels from pR_AGO1_F in the same conditions were also maintained and were significantly greater than those from the negative controls (Fig 7b). Altogether, these results show that AGO1 5'UTR is able to maintain protein synthesis under conditions impairing cap-dependent translation initiation.

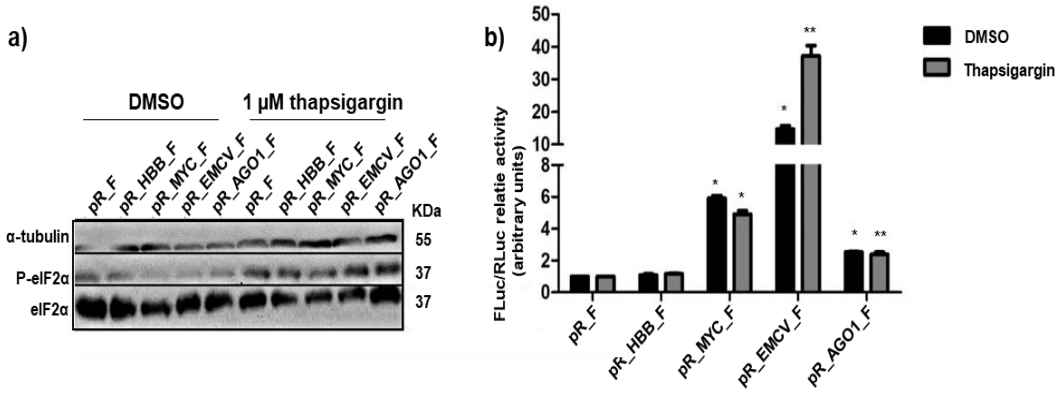

Figure 7. FLuc expression mediated by AGO1 5'UTR after treatment with 1µM of thapsigargin that suppresses global protein synthesis. HeLa cells were transfected with pR_F, pR_HBB_F, pR_MYC_F, pR_EMCV_F or pR_AGO1_F plasmids and treated with 1 µM of thapsigargin. **a)** Western blot analysis against phosphorylated and non-phosphorylated eIF2α proteins. α-tubulin is the sample loading control. **b)** Relative luciferase activity, which is represented as the luminescence ratio between FLuc and RLuc compared to that obtained from the pR_F construct in each control condition, arbitrarily set to 1. Black bars indicate cells in control conditions (DMSO) and grey bars represent cells treated with 1µM of thapsigargin. Presented data are the result of, at least, three independent experiments. Statistical analysis was performed using the Student's t test (unpaired, two-tailed). *P<0.05, **P<0.01.

3. Aims

Argonaute proteins play an important role in the cell, since they participate in the biogenesis of microRNAs, whose function is essential for the regulation of several biological events. Given its importance, it is natural that its synthesis is maintained under conditions that are inhibitory to the canonical translation initiation. The alternative mechanisms of translation initiation guarantee an adaptation when the environment changes, being able to maintain the translation of essential proteins in response to stress.

Previous results show that the 5'UTR of human AGO1 is able to mediate a cap-independent translation in a bicistronic reporter vector. The purpose of this thesis is to continue this work and continue to study cap-independent translation of AGO1. Thus, our specifics aims are:

i. Support the previous results using a different bicistronic system;

ii. Evaluate how the expression levels of endogenous Argonaute proteins, especially AGO1, vary during inhibition of cap-dependent translation;

iii. Simulate the secondary structure adopted by the 5'UTR of Argonaute 1;

4. Material and Methods

4.1. Plasmid Constructs

The pEGFP_mCherry bicistronic plasmid was a gift from Dr Marco Candeias (Kyoto University) [130]. It contains two reporter genes, enhanced green fluorescent protein (EGFP), which is translated from a cap-dependent manner, and mCherry which is translated from a cap-independent manner. A stable hairpin was cloned downstream of EGFP stop codon to prevent translation reinitiation. Upstream of mCherry AUG is cloned Δ133p53 that is an isoform of p53.

The pEGFP_mCherry was digested with two restriction enzymes, EcoRI/XhoI, to remove Δ133p53, generating the vector to ligate the inserts (HBB 5´UTR and AGO1 5´UTR) For that, the human β-globin 5'UTR (HBB; NM_000518), negative control for cap-independent translation, was PCR amplified, using primers #1 and #2 with linker for EcoRI in primer forward and linker for XhoI in primer reverse, respectively (Table 1). The Argonaute 1 5´UTR (AGO1; NM_012199) was PCR amplified, using primers #3 and #4 with linker for EcoRI in primer forward and linker for XhoI in primer reverse, respectively (Table 1). The ligation reaction between the vector and the inserts was performed using T4 DNA Ligase (NZYTech), according to manufacturer's instructions, to generate the vectors pEGFP_HBB_mCherry and pEGFP_AGO1_mCherry.

Table 1. Sequences of primers used in this work

Primer	Sequence (5´→3´)
#1	CCGGAATTCACATTTGCTTCTGAC
#2	CCGCTCGAGGGTGTCTGTTTGAGG
#3	CCGGAATTCACTGGCAGCTGGCCG
#4	CCGCTCGAGCCCATATACCCGTGC

4.2. Cell Culture, Drug Treatments and Cell Lysis

Human HCT116 cells were cultured in Dulbecco's modified Eagle's medium (DMEM), supplemented with 10% (v/v) fetal bovine serum. Cells were kept at 37°C in a humidified atmosphere containing 5% CO_2. Cells were seeded in 35 mm plates and 24 hours after they were treated with DMSO (vehicle) or 4 μM of thapsigargin (Sigma-Aldrich) for other 24 hours. After treatment, cells were harvested with 100 μl of 1x (v/v) passive lysis buffer (PLB) (Promega). Lysates were stored at -80 ºC until needed.

4.3. RNA Isolation and Reverse Transcription-Polymerase Chain Reaction (RT-PCR)

Total RNA from treated cells was extracted using Nucleospin RNA extraction II (Macherey-Nagel), according to the manufacturer's instructions. cDNA was synthesized from 1μg RNA using NZY Reverse Transcriptase (NZYTech) and random hexamers (Invitrogen), according to manufacturer's instructions.

4.4. Reverse Transcription-quantitative PCR (RT-qPCR)

For analysis of relative mRNA levels, RT-qPCR was performed on an ABI 7500 Real-Time PCR System (Applied Biosystems® by Life TechnologiesTM), using SybrGreen Master Mix (Applied Biosystems® by Life TechnologiesTM), cDNA as a template and 1μM of specific primers for the genes of interest (Table 2). The PCR program was performed as described in Table 3. GAPDH was used as internal control, serving as a marker of cytoplasmatic gene expression. The relative mRNA expression of each gene was calculated using the ΔΔCt method.

Table 2. Sequence of the primers used in RT-qPCR to measure the relative mRNA levels of GAPDH, AGO1, AGO2, AGO3 and AGO4.

Gene	Primer sequence (5´→3´)
GAPDH	Forward: CCATGAGAAGTATGACAACAGCC
	Reverse: GGGTGCTAAGCAGTTGGTG
AGO1	Forward: GCACTGCCCATTGGCAACGAA
	Reverse: CATTCGCCAGCTCACAATGGCT
AGO2	Forward: CGCGTCCGAAGGCTGCTCTA
	Reverse: TGGCTGTGCCTTGTAAAACGCT
AGO3	Forward: GGAATTAGACAAGCCAATCAGCA
	Reverse: AGGGTGGTCATATCCTTCTGGA
AGO4	Forward: CTAACAGACTCCCAGCGTGTCA
	Reverse: GACTGGCTGGCCGTCTAGTCA

Table 3. PCR programme used in RT-qPCR

Stage	Temperature (ºC)	Time
Holding stage	95	10 m
Cycling stage (40 cycles)	95	15 s
	62	30 s

4.5. Quantification of the Total Protein by the Bradford Method

The Bradford method was used for the quantification of the total protein, and thus to predict the effect of thapsigargin treatment on protein synthesis and on protein level. Standard calibration curve and quantification of total protein amount was made using NZYBradford reagent (NZYTech) according to manufacturer's instructions. Absorbance was measured at 595nm in a NanoDrop 1000 Spectrophotometer (Thermo Fisher Scientific).

4.6. SDS-PAGE and Western Blot

Cell lysates were denatured with SDS sample buffer 5x (v/v) at 95°C for 20 minutes. Samples are resolved in a 10% polyacrylamide gel by SDS-PAGE and transferred to PVDF membrane (Bio-Rad), previously activated with methanol. The membranes were blocked according to the proteins to be identified, as will be described further below. Thus the blocking was done with: 5% (w/v) bovine serum albumin (BSA, Sigma-Aldrich), 1x (v/v) tris-buffered saline (TBS) and 0,05% (v/v) Tween 20 (Sigma-Aldrich); 5% (w/v) non-fat dry milk, 1x (v/v) TBS and 0,1% (v/v) Tween 20 (Sigma-Aldrich); 5% (w/v) non-fat dry milk, 1x (v/v) TBS and 0,05% (v/v) Trinton x-100 (Sigma-Aldrich); 5% (w/v) non-fat dry milk, 1x (v/v) tris-buffered saline (TBS) and 0,05% (v/v) Tween 20 (Sigma-Aldrich). Membranes were

probed using the following antibodies: rabbit polyclonal anti-Phospho-eIF2α (Ser52) (Invitrogen) at 1:250 dilution in 5% (w/v) BSA (Sigma-Aldrich), 1x (v/v) tris-buffered saline (TBS) and 0,05% (v/v) Tween 20 (Sigma-Aldrich); rabbit polyclonal anti-eIF2α (Cell Signaling) at 1:500 dilution 5% (w/v) non-fat dry milk, 1x (v/v) TBS and 0,1% (v/v) Tween 20 (Sigma-Aldrich); mouse monoclonal anti-α-tubulin (Sigma-Aldrich) at 1:50,000 dilution in 5% (w/v) non-fat dry milk, 1x (v/v) TBS and 0,05% (v/v) Trinton x-100 (Sigma-Aldrich); mouse monoclonal anti-eIF2C (Argonaute1-4) (Santa Cruz Biotechnology) at 1:250 dilution in 5% (w/v) non-fat dry milk, 1x (v/v) tris-buffered saline (TBS) and 0,05% (v/v) Tween 20 (Sigma-Aldrich); rabbit polyclonal anti-Ago1 (Abcam) at 1:250 dilution in 5% (w/v) non-fat dry milk, 1x (v/v) tris-buffered saline (TBS) and 0,05% (v/v) Tween 20 (Sigma-Aldrich). Detection was carried out by incubating the membranes with the secondary antibodies, peroxidase-conjugated anti-mouse IgG (Bio-Rad) or anti-rabbit IgG (Bio-Rad) antibodies, followed by enhanced chemiluminescence (ECL).

4.7. *In silico* Analysis

Human AGO1 (NM_012199) 5'UTR sequences were curated in NCBI (http://www.ncbi.nlm.nih.gov/) database as the most common variant. GC content (%) of AGO1 5´UTR was calculated with Endmemo software (http://www.endmemo.com/). mFold software (http://mfold.rna.albany.edu/?q=mfold) was used to predict the secondary structure of human AGO1 5'UTR, applying the standard parameters defined by the software.

4.8. Data Analysis and Statistics

In order to evaluate mRNA expression of Argonautes under stress conditions, the relative levels of Argonaute mRNA and protein under drug treatment was normalized to the relative levels of Argonaute mRNA and protein treated only with DMSO, which is the vehicle.

Results are expressed as mean ± standard deviation resulting from at least three independent experiments. Unpaired two-tailed Student´s t-test was used for estimation of statistical significance. Significance for statistical analysis was defined as $p<0.05$.

5. Results

5.1. Construction of pEGFP_AGO1_mCherry bicistronic reporter vector

To support the results presented in section 2 it was thought to test whether AGO1 5'UTR promotes cap-independent translation in a different bicistronic system. For this purpose, a bicistronic vector was used, in which the first cistron is the enhanced green fluorescent protein (EGFP) and in the second cistron is the mCherry, a red fluorescent protein. In this system, the EGFP is translated in a cap-dependent manner and represents an internal control and mCherry is translated in a cap-independent manner. Similar to the bicistronic vector with the luciferases a stable hairpin was cloned downstream of EGFP ORF to prevent translation reinitiation, that has already been shown to efficiently inhibit ribosome scanning [130]. Transcription of the bicistronic plasmid is under the control of CMV promoter.

In this way we attempted to clone the AGO1 5'UTR upstream of mCherry ORF, forming the vector pEGFP_AGO1_mCherry to analyse its cap-independent translation (Fig. 8). As negative controls, we used the empty plasmid (pEGFP_mCherry) and the human β-globin 5'UTR cloned in the same plasmid (pEGFP_HBB_mCherry), which has already been demonstrated does not mediate the cap-independent translation (Fig. 6 and 7).

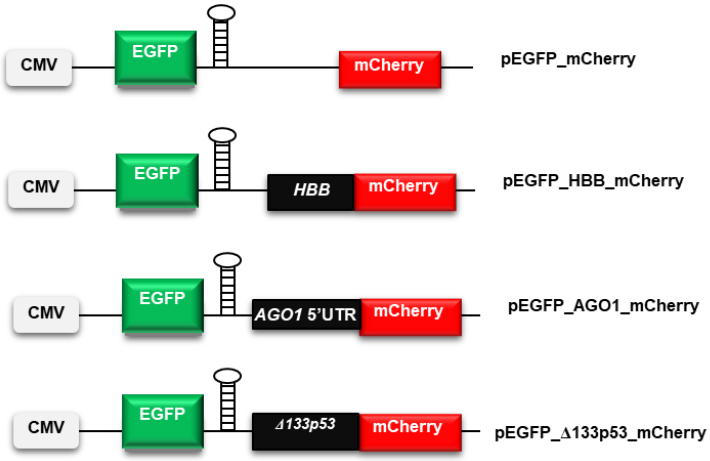

Figure 8. Schematic representation of the vectors that will be used to check whether AGO1 5'UTR is able to drive cap-independent translation initiation. EGFP is the enhanced green fluorescent protein cap-dependent translated cistron (green box) and mCherry the protein cap-independent translated cistron (red box). CMV box represents the promoter. Black boxes represent the different sequences cloned upstream mCherry ATG: empty vector (pEGFP_mCherry); human β-globin 5'UTR (HBB) is the negative control for cap-independent translation initiation activity (pEGFP_HBB_mCherry); AGO1 5'UTR (pEGFP_AGO1_mCherry); the vector containing Δ133p53 (pEGFP_Δ133p53_mCherry) was used as the base to generate the other plasmids. All constructs contain a stable hairpin (represented by a stem loop) downstream EGFP cistron to prevent translation reinitiation.

To generate the plasmids to be tested the base plasmid was the bicistronic vector with reporters EGFP and mCherry, with Δ133p53, which is an isoform of p53, cloned upstream of mCherry (pEGFP_Δ133p53_mCherry) (Fig. 8). Next, the cloning strategy was to remove by enzymatic digestion the Δ133p53 of the bicistronic plasmid and clone our sequences of interest upstream of mCherry ATG.

Thus, as already mentioned in Materials and Methods (Section 4.1), the vector pEGFP_Δ133p53_mCherry was digested with EcoRI and XhoI (Fig. 9a), which results in two plasmid DNA fragments. One corresponding to Δ133p53 (810 bp) and another corresponding to the pEGFP_mCherry vector (6871 bp) (Fig. 9b – lane pD). Subsequently, the vector band were extracted from the gel and purified (using NZYGelpure kit, NZYTech). In parallel, the AGO1 5'UTR (213 bp) and HBB 5'UTR (50 bp) were amplified by PCR, using the primers showed in Table 1, extracted and purified (using NZYGelpure kit, NZYTech) (Fig. 9c).

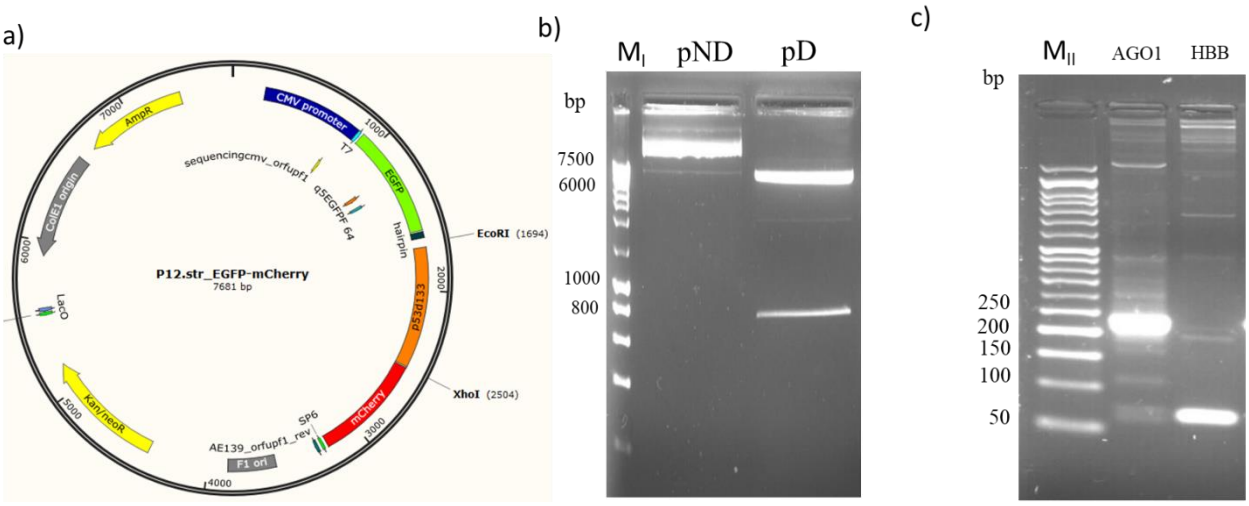

Figure 9. Cloning strategy to construct the vectors pEGFP_AGO1_mCherry and pEGFP_HBB_mCherry. a) Schematic representation of pEGFP_Δ133p53_mCherry vector. In green is represented the EGFP, orange the Δ133p53 and red the mCherry. On the right side are represented the position of the EcoRI and XhoI restriction sites. **b)** Agarose gel corresponding to the digestion of vector the pEGFP_Δ133p53_mCherry with EcoRI and XhoI. M_I – DNA marker (NZYDNA Ladder III); pND – non-digested pEGFP_Δ133p53_mCherry; pD – pEGFP_Δ133p53_mCherry digested with EcoRI/XhoI. **c)** PCR amplification of inserts. M_{II} – DNA marker (NZYDNA Ladder VI); AGO1 – Amplification of AGO1 5´UTR; HBB – Amplification of HBB 5´UTR

These inserts were then cloned into a commercial vector of known sequence (NZY-A PCR cloning kit, NZYTech) and was sequenced to confirm if the sequence of our fragments corresponded to AGO1 5'UTR and HBB 5'UTR. Effectively when comparing the reference sequences of AGO1 5'UTR and HBB 5'UTR to the sequences present in the plasmids obtained, it is verified that exist a match (Annex, Fig. 1 and Fig. 2). Then, the inserts were digested using EcoRI and XhoI, and ligated to the the vector (pEGFP_mCherry) using a T4 DNA ligase (NZYTech) and a 1:5 molar ratio of vector:insert. The ligation product was amplified using bacterial competent cells. Subsequently, the colonies were submitted to a screening by a colony PCR using the primers of Table 1. Effectively, the amplified fragments appeared to have a size corresponding to the respective 5'UTRs, 213 bp for AGO1 and 50 bp for HBB (Fig. 10a and b).

Finally, to verify if the cloning of AGO1 5'UTR and HBB 5'UTR in pEGFP_mCherry was successful, the respective plasmids were sequenced. However, the results obtained from sequencing for the two cases are inconclusive, which may indicate that cloning was not done (Fig. 10c). Indeed the results of the chromatogram indicate that there was no sequence reaction, most likely because the primers did not hybridize in the sequence analysed. Unfortunately, we were not able to confirm these results by Sanger sequencing (Fig. 10c), probably because these clones were false positive.

In order to solve this problem, it was tried to optimize the ligation protocol between the vector and the inserts. For this, several conditions were tested for the ligation reaction, such the reaction temperature, reaction time, the amount of vector used or the vector:insert molar ratio. The effect of the temperature was tested by attempting to make the reactions at 18 °C and at 23 °C. As for the reaction time, a ligation was tested for 3h or overnight (~16h). Furthermore, we performed a ligation between the vector and inserts, in which we used different amounts of vector, being tested 50 ng or 100 ng vector. We also tried to potentiate the binding by changing the vector:insert molar ratio, where we tested 1:5, 1:7, 1:10 and 1:25.

In some of these optimization attempts, we had growth of transformant colonies and an apparent positive result in PCR screening. However, when sequencing the plasmids obtained, the results were not

confirmed by Sanger sequencing (similar to the results observed in Figure 10). Unfortunately, until the end of this thesis we have not been able to complete the cloning of the vectors pEGFP_AGO1_mCherry and pEGFP_HBB_mCherry, however, we will present ahead some hypotheses on how these cloning can be done in the future.

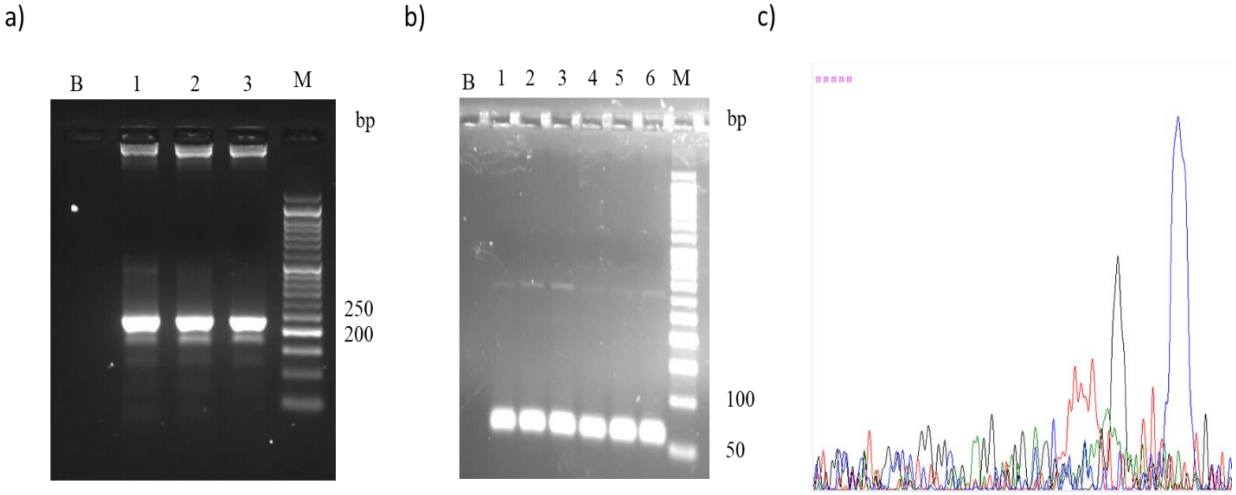

Figure 10. Analysis of pEGFP_AGO1_mCherry and pEGFP_HBB_mCherry cloning. a) Screening PCR of pEGFP_AGO1_mCherry. B – blank; Lanes 1, 2 and 3 – plasmid samples purified from transformant colonies; M – DNA marker (NZYDNA Ladder VI). **b)** Screening PCR of pEGFP_HBB_mCherry. B – blank; Lanes 1, 2, 3, 4, 5, and 6 – plasmid samples purified from transformant colonies; M – DNA marker (NZYDNA Ladder VI). **c)** Sequencing results of pEGFP_AGO1_mCherry and pEGFP_HBB_mCherry. At the top of the figure are represented some N's (pink), which means that the sequencing analysis program, ape (http://en.bio-soft.net/plasmid/ApE.html), cannot identify the nucleotides present. This indicates that the primers did not hybridize with the sequence during the sequencing reaction.

5.2. Expression levels of endogenous Argonaute proteins during stress conditions

Previously it has been shown that Argonaute 1 5'UTR can drive cap-independent translation of a reporter gene, in a bicistronic context, under normal and stress conditions. Now, we want to understand how the expression of endogenous Argonautes (AGO1, AGO2, AGO3 and AGO4) varies during stress conditions that inhibit cap-independent translation. For this purpose, we treated HCT116 cells with thapsigargin, which inhibits the canonical initiation of translation. This drug promotes endoplasmic reticulum stress, which activates the unfolded protein response (UPR), and consequently, the PERK pathway. In turn, this protein phosphorylates eIF2α leading to the inhibition of cap-dependent translation, as shown previously [3, 9].

Cells were treated with 4 μM of thapsigargin, while controls were only exposed to the thapsigargin vehicle, DMSO. For both cases the phosphorylation of eIF2α was followed by Western blot analysis, that showes an increased amount of phosphorylated eIF2α protein in cells treated with thapsigargin, compared to DMSO treated cells (Fig. 11a). In this way, it is possible to ensure that the cap-dependent translation of these cells is inhibited by the drug effect.

The effects of the thapsigargin were also confirmed by the decrease in total protein content, after 24h of drug treatment (Fig. 11b)

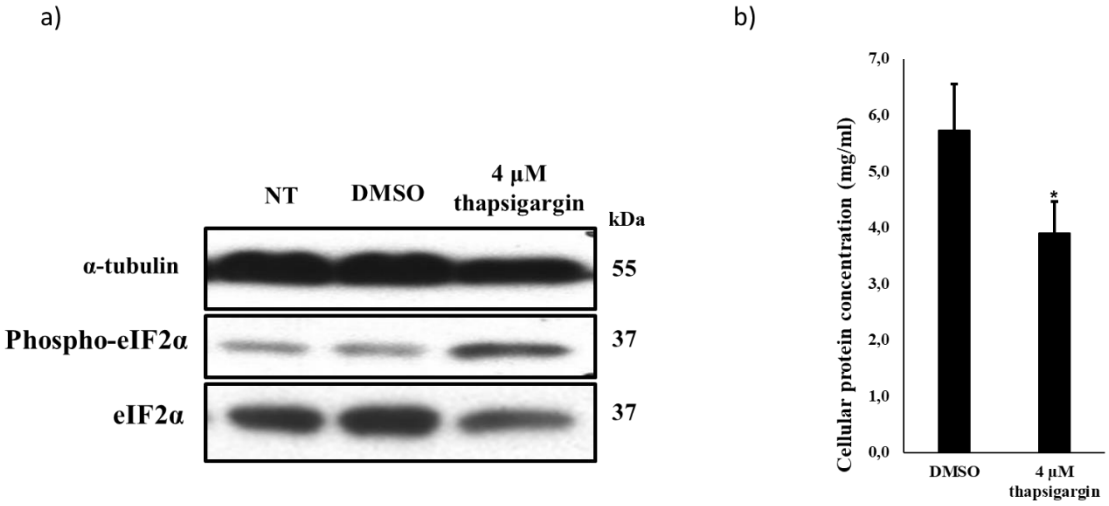

Figure 11. Effect of cellular thapsigargin treatment on cap-dependent translation in HCT116 cells. a) Consequence of treatment with 4 μM of thapsigargin on phosphorylated and non-phosphorylated eIF2α proteins of HCT116 cells. In lane 1 is represented the Western blot analysis of protein extracted from cells without treatment (NT – lane 1); In lane 2 is represented Western blot analysis of protein extracted from cells treated with DMSO (vehicle – lane 2); In lane 3 is represented Western blot analysis of protein extracted from cells treated with 4 μM of thapsigargin (lane 3). α-tubulin is the sample loading control. The Western blots shown are a representative of at least three independent experiments. **b)** Variation of the cellular protein concentration (mg/ml) between control conditions (DMSO) and treatment with 4 μM of thapsigargin, indicating global cap-dependent translation inhibition in the thapsigargin treated cells. Presented data are the result of at least three independent experiments. Statistical analysis was performed using the Student's t test (unpaired, two-tailed), * $P < 0.05$

Next, we evaluated the effect of cap-dependent translation inhibition, in the expression of Argonaute mRNA, using a Real Time PCR. This method will allow us to determine how the levels of expression and the relative amount of Argonaute mRNA vary under conditions of canonical translation inhibition. As shown in Figure 12, treatment with 4 μM of thapsigargin appears to affect the mRNA levels of Argonaute 1 and Argonaute 2 and have no effect on the mRNA levels of Argonaute 3 and Argonaute 4. In fact, treatment with the drug decreases the mRNA levels of AGO1 (54%) and AGO2 (62%) for almost half, when compared to control (DMSO treatment). As for the levels of AGO3 and AGO4 mRNA, they appear to be stable (Fig. 12).

These results suggest that although this drug acts at the level of translation, the stress caused also indirectly affects other cellular processes. In this case the cellular stress inherent to inhibition of translation affect the availability of two transcripts, the AGO1 and the AGO2. So, it is possible to infer that the treatment with thapsigargin and, consequently, the inhibition of cap-dependent translation promotes changes in the turnover or transcription rate of AGO1 and AGO2 mRNAs, in which case these changes lead to a decrease in the amount of AGO1 and AGO2 transcript. In this case, the stimulus with the drug may be promoting, indirectly, an increase in the degradation rate of mRNA or limiting the synthesis *de novo* mRNA, which leads to a decrease in AGO1 and AGO2 transcript levels.

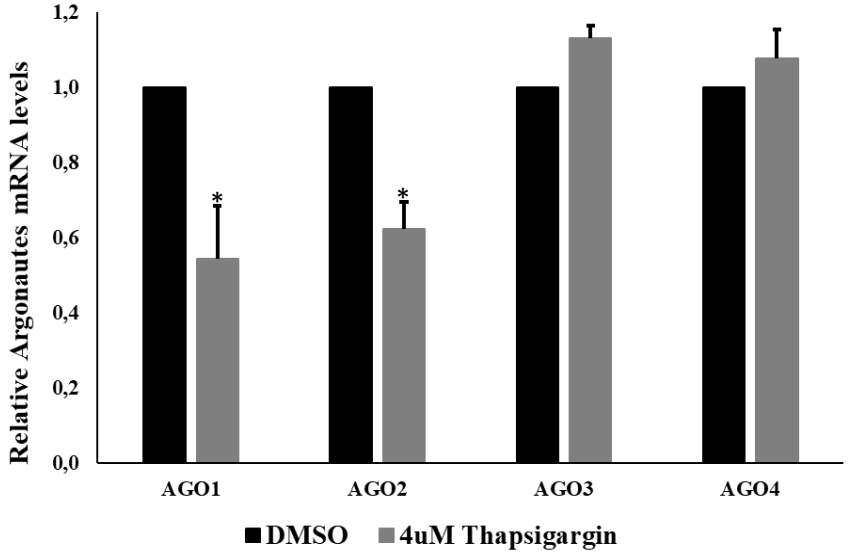

Figure 12. Effect of cellular thapsigargin treatment on Argonaute mRNA levels. AGO1, AGO2, AGO3 and AGO4 mRNA levels expressed in cells treated with 4 μM of thapsigargin were normalized to AGO1, AGO2, AGO3 and AGO4 mRNA levels expressed in DMSO treated cells. Treatment with DMSO represents the control situation, thus having a value of 1. Presented data are the result of at least three independent experiments. Statistical analysis was performed using the Student's t test (unpaired, two-tailed), * P <0.05

Subsequently, we wanted to understand how the inhibition of cap-dependent translation, induced by thapsigargin, influences the synthesis of Argonaute proteins and whether cap-independent translation can maintain the protein levels of AGO1 under stress conditions of the endoplasmic reticulum. To this end we reanalysed by Western blot the lysates of the first experiment, whose thapsigargin induced phosphorylation of eIF2α. At this stage we intend to analyse how the treatment with the drug affects the synthesis of Argonaute proteins, and for that we use a specific antibody that recognizes all proteins of this family. This particularity of the antibody is that it recognizes the C-terminal of Argonaute proteins, which has a high identity among the four proteins. It should be noted that all Argonautes have a similar molecular weight (~97 kDa), so it is impossible to distinguish them by Western blot using this antibody.

By Western blot results, we observe that treatment with thapsigargin induces a decrease in the amount of Argonaute proteins (Fig. 13). Treatment with this drug decreases the levels of AGO1 and AGO2 mRNA, which will mean that there is less transcription available for translation and may lead to a lower level of synthesis of these proteins (Fig. 12 and 13). Thus, the decrease of AGO1 and AGO2 levels due to treatment with thapsigargin, will result in a decrease in the total pool of of AGOs.

However, it is important to note that this result has limitations, since it is not possible to determine whether treatment with the drug affects all AGOs, or only a few, and this is what decreases the Argonaute protein levels.

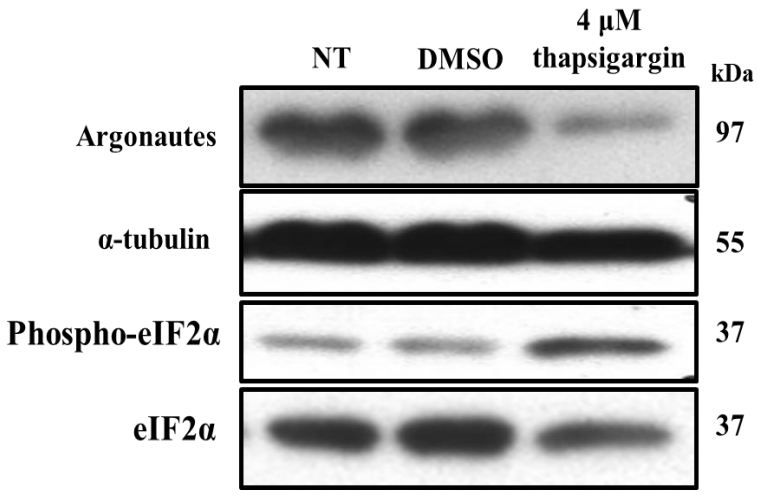

Figure 13. Effects of cellular thapsigargin treatment on Argonaute protein levels in HCT116 cells. Consequence of treatment with 4 µM of thapsigargin on Argonaute, phosphorylated and non-phosphorylated eIF2α protein levels of HCT116 cells. In lane 1 is represented the Western blot analysis of protein extracted from cells without treatment (NT – lane 1); In lane 2 is represented Western blot analysis of protein extracted from cells treated with DMSO (vehicle) – lane 2; In lane 3 is represented Western blot analysis of protein extracted from cells treated with 4 µM of thapsigargin – lane 3. α-tubulin is the sample loading control. The Western blots shown are a representative of at least three independent experiments.

Our next goal was to understand if the treatment with thapsigargin induced the decrease of all the Argonautes or if the protein levels of AGO1 could be maintained by a cap-independent translation under stress conditions. For this, we did a Western blot analysis that evaluated the inhibition of canonical translation on AGO1 protein levels, using an antibody that specifically recognizes human AGO1, since its immunogen is located at the N-terminal, which is the most heterogeneous zone between the Argonautes.

The results in Figure 14 a), show that treatment with 4 µM of thapsigargin decreases the protein levels of Argonaute 1, as compared to the DMSO control. To further understand how drug treatment affects AGO1 protein levels, we performed an analysis by Image J software (https://imagej.nih.gov/ij/), where we measured the intensity of the Western Blot bands corresponding to AGO1, with DMSO and drug (Fig. 14b). In order to do this, we used the Western Blot in which we have the four independent experiments that we did, and we calculated by the software the mean densitometry of AGO1 in a control and drug situation (Annex, Fig. 3). Altogether, these results suggest that treatment with thapsigargin affects the relative protein levels of Argonaute 1.

However, it has already been observed that the 5'UTR of AGO1 is able to mediate a cap-independent translation under stress conditions. So, we believe, that the explanation for decrease of proteins levels is derived from different factors. One of our hypotheses has to do with the fact that alternative translation mechanisms, namely cap-independent translation, may not have the same efficiency as the canonical via, in the case of AGO1. If we look at the results in Figure 7, it is found that cap-independent translation mediated by AGO1 5'UTR under stress conditions is less efficient than that mediated by c-Myc IRES, which is also a cellular transcript with an alternative translation mechanism. Thus, we believe that cap-independent translation mediated AGO1 5´UTR under stress conditions be less efficient in terms of protein level than the canonical way, which may explain the decrease of protein levels observed (Fig. 14).

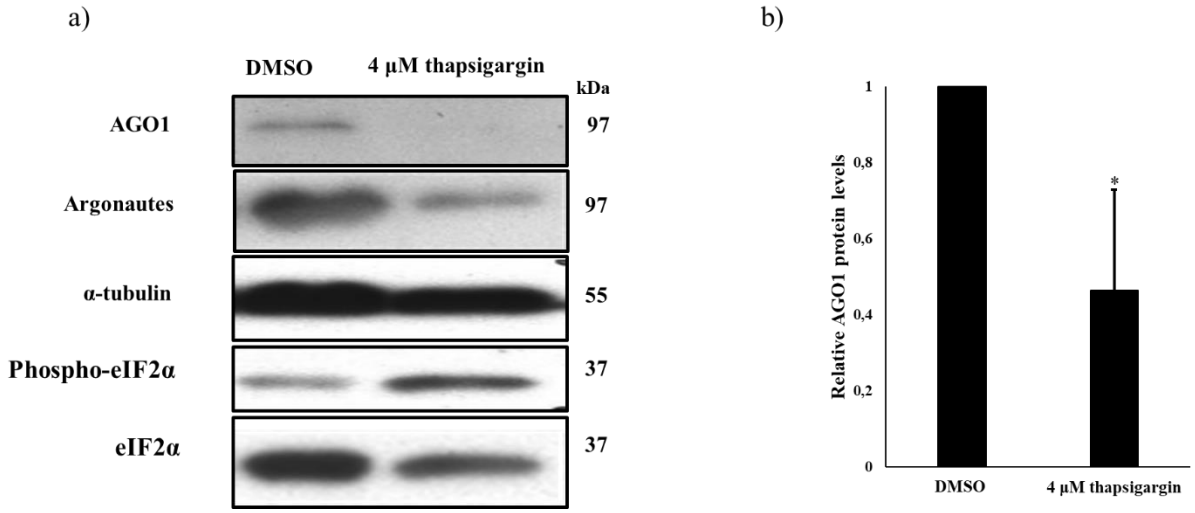

Figure 14. Effect of cellular thapsigargin treatment on Argonaute 1 protein levels in HCT116 cells. a) Consequence of treatment with 4 µM of thapsigargin on Argonautes, phosphorylated and non-phosphorylated eIF2α protein levels of HCT116 cells. In lane 1 is represented Western blot analysis of protein extracted from cells treated with DMSO (vehicle) – lane 1; In lane 2 is represented Western blot analysis of protein extracted from cells treated with 4 µM of thapsigargin – lane 2. α-tubulin is the sample loading control. The Western blots shown are a representative of at least three independent experiments. **b)** Analysis by image J software of the intensities of the AGO1 bands, corresponding to the treatment with DMSO and with 4 µM of thapsigargin. Treatment with DMSO represents the control situation, thus having a value of 1. Presented data are the result of at least three independent experiments. Statistical analysis was performed using the Student's t test (unpaired, two-tailed), * P <0.05.

Another factor that may be contributing to these results is related to the fact that treatment with 4 µM of thapsigargin, which inhibits cap-dependent translation, leads to the decrease of AGO1 mRNA levels (Fig. 12). In this case there would be less transcript available for translation and possibly less protein synthesis. Therefore, we thought that the decrease in the AGO1 protein levels due to thapsigargin treatment may also be a consequence of a smaller amount of mRNA available for translation.

Finally, it is important to bear in mind that there are other regulatory mechanisms to which endogenous mRNAs are subjected during translation, unlike reporter genes, and these processes may alter the expected results. For example, Chen and colleagues reported that AGO1 mRNA is silenced under hypoxia, and this leads to a decrease in protein levels. In this case during this stress are activated miRNAs that act on the 3'UTR of AGO1 mRNA and cause their silencing, and consequently, the decrease in the AGO1 synthesis [132]. Thus, making a parallel with our results, the stress with thapsigargin lead to less available AGO1 transcript which will probable lead to decreased AGO1 protein synthesis.

5.3. Secondary structure adopted by the 5'UTR of Argonaute 1

Cap-independent translation is often reliant on the presence of a 5'UTR rich in secondary and tertiary structures that facilitates in most situations, the ribosomal recruitment to the vicinity of the main AUG, possible via internal entry of the ribosome. On the other hand, these structures may impair the regular scanning of the 5'UTR and, hence, promote the mechanism of ribosome shunting that forces the ribosome to bypass them and reach the AUG in a non-canonical way.

In this way, we aim to understand the structure adopted by AGO1 5'UTR that allows it to mediate cap-independent translation. We submitted the human AGO1 5'UTR sequence (NM_012199) to an *in silico* analysis to predict the structure and the characteristics that favour non-canonical translation initiation.

The 5'UTR of AGO1 has a length of 213 nts, with an upstream AUG within the untranslated region, located at position -5 compared to the main AUG. Interestingly, this suggests the possibility of a uORF regulating AGO1 protein expression. AGO1 5'UTR contains on average a GC content of 72.3%, and the areas with more GC content (up to 87.5%) are located in the regions adjacent to the 5'end, whereas the areas with lower GC content (53.8% minimum) are located in the regions adjacent to 3'end (Fig. 15 a). This high GC content indicates that the 5'UTR fold is likely to form stable RNA secondary structures.

We also used the mFold software (http://unafold.rna.albany.edu/) to predict the spatial conformation adopted by this region. According to mFold, AGO1 5´UTR tends to fold in a relatively stable structure – with a minimum free energy of ΔG = -111.95 kcal/mol (Fig. 15b).

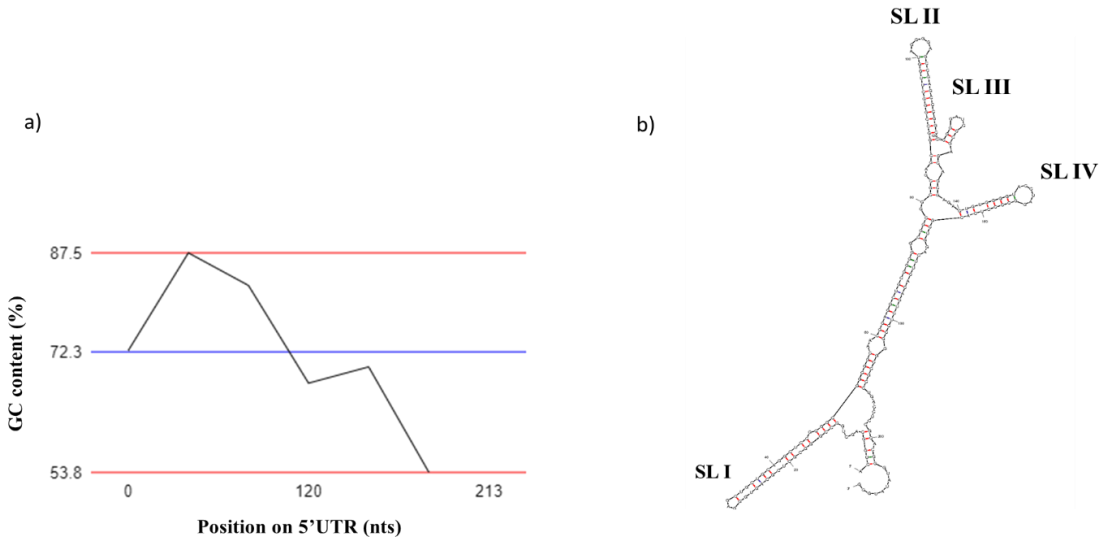

Figure 15. Structural analysis of human Argonaute 1 5´ untranslated region (AGO1 5´UTR). a) Calculation of the GC content (%) of AGO1 5'UTR (data obtained from http://www.endmemo.com). The GC content (average 72.3%) ranges from 87.5% to 53.8%, while the highest GC percentage is localized along the first 120 nucleotides (nts). **b)** Most stable secondary structure of AGO1 5'UTR, predicted by mFold software (http://unafold.rna.albany.edu/) reveals the formation of four stem loops (SLI, SLII, SLIII and SLIV), with a minimum free energy of ΔG = -111.95 kcal/mol.

AGO1 5'UTR is organized into a secondary structure formed by four stem loops (SL) – I, II, III and IV. (Fig. 15 b). The SLI is located between nucleotides 12-51 and corresponds to the richest GC region, suggesting a structure with great stability. The SLII is located between 88-119 nts which is also a region with a high GC content, about 80%. SLIII and SLIV are located between 120-129 nts and 140-163 nts, respectively, the regions with lowest GC content of the 5'UTR AGO1, which may reflect a more dynamic structure during translation. In the future, we intend to experimentally confirm these predictions and see if the AGO1 5´UTR folds into these four stem loops.

Subsequently, we tried to understand the structural dynamics of AGO1 5'UTR, and for this we checked by *in silico* analysis if there were minimal sequences necessary to maintain the structure of AGO1 5'UTR stable. For this, we performed a deletional *in silico* analysis of 5´UTR sequence that consisted of sequentially removing 50 nucleotides at a time, both in a 5' to 3' direction or in a 3' to 5' direction. Using mFold software, we performed an *in silico* analysis of the predicted secondary structures formed in the absence of each deleted sequence and compared it to the secondary structure predicted to the full-length sequence. It should be noted that in this analysis, we selected the predicted structures with lower free energy, assuming that these will be the most stable. Therefore, we evaluated how the original structure was affected by the deletions and, specifically whether the stem loops predicted to be formed by the full-length sequence were maintained or disrupted (Fig. 16 and 17)

By making the deletions in the 5' to 3' direction it appears that the removal of the first 50 nucleotides completely eliminates the SLI, maintaining the SLII, SLIII and SLIV with a spatial conformation similar to the full length structure (Fig. 16a and b). The deletion of 1-100 nts disrupts SLI and SLII, however SLIII remains with a conformation identical to the original that is a short stem loop (Fig. 16c). On the other hand, SLIV, seems to change its original structure. Deletion of the first 150 nts of 5'UTR completely abolishes SLI, SLII, and SLIII, however it appears that SLIV tends to form a structure similar to the previous deletion, although very different from the original SLIV (Fig. 16d). Of note, that even after the deletion of 1-150 nts, the sequence is able to form another stem loop, most likely as a reflection of the high GC content of the AGO1 5'UTR. These predictions show that the disruption of some loops, especially when deleting the first 50 or 100 nts, tends not to cause major changes in the following loops, this being especially observable when SLI removal does not affect the spatial conformation of the SLII, SLIII and SLIV (Fig. 16a). This may indicate some independence of these structures within the 5'UTR AGO1.

As for the deletions in the 3' to 5' direction, we verified that the deletion of 51-213 nts causes the disruption of the entire original structure, only two small stem loops are formed, which are not identifiable in the full length conformation (Fig. 17b). The spontaneous formation of these structures should be derived from the high GC content in this region of the 5'UTR. By removing 101-213 nts, occurs SLI formation, with a spatial conformation comparable to the predicted full-length structure (Fig. 17c). The deletion of the last 63 nts and analysis of the sequence between 1-150 nts shows that SLIV is disrupted, while there is a maintenance of the structure of SLI, SLII and SLIII with respect to the original conformation (Fig. 17a and d).

In sum, this deletional *in silico* analysis leads us to believe that the spatial conformation adopted by the various stem loops of the AGO1 5'UTR is relatively independent from each other. This is because when small deletions occur, as in Figures 16 b) and 17 d), no significant changes happen in the conformation of the other loops, maintaining a structure similar to the original. To couple this *in silico* analysis, it will be important to carry out an experimental validation in order to understand which deletions affect and if there is a minimum sequence/structure required for AGO1 5'UTR mediated cap-independent translation.

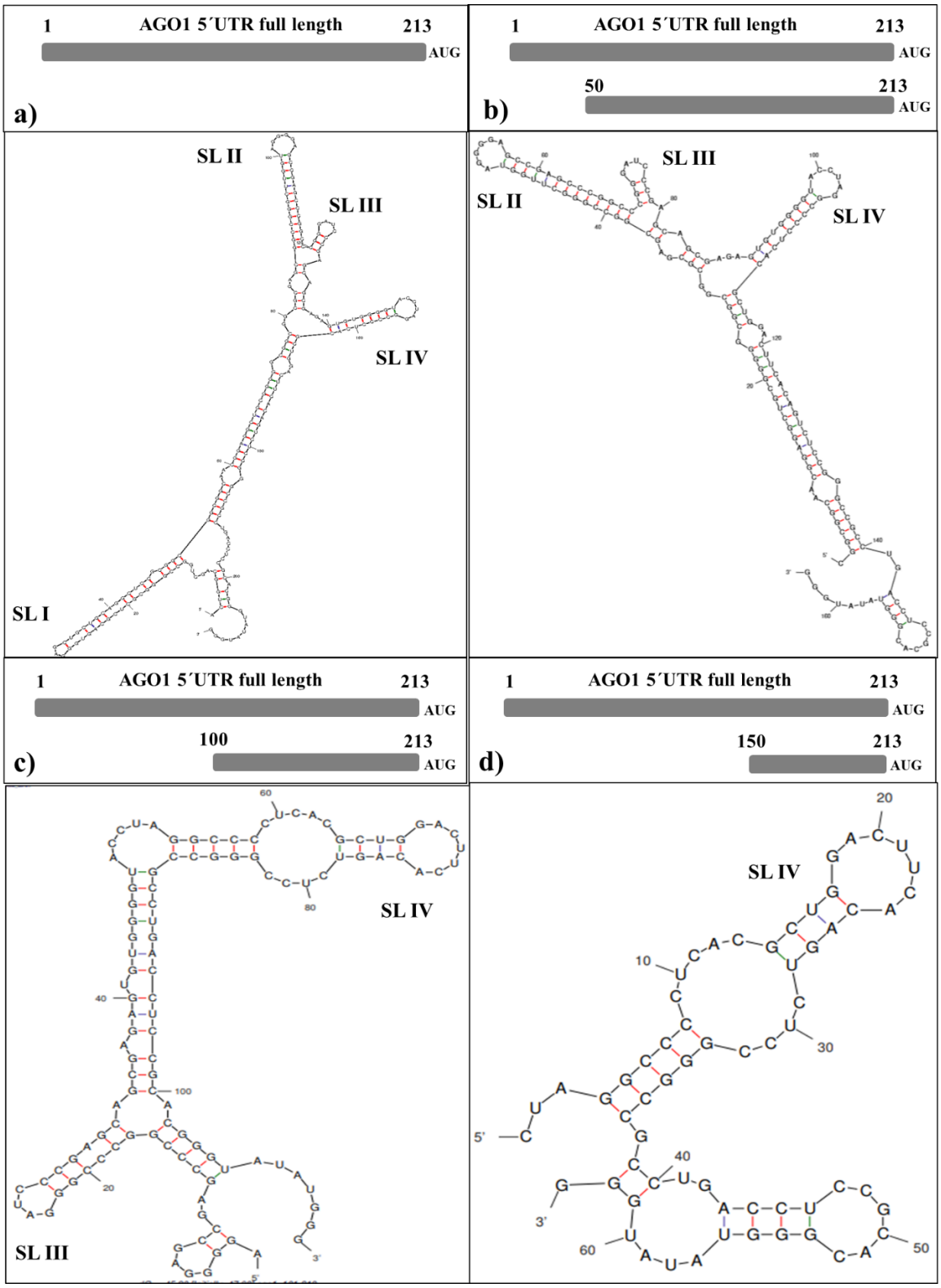

Figure 16. Secondary structures predicted for AGO1 5'UTR with successive deletions in the 5' to 3' direction. a) Full-length AGO1 5'UTR; **b)** nucleotides (nts) 50-213 of AGO1 5'UTR; **c)** nts 100-213 of AGO1 5'UTR; **d)** nts 150-213 of AGO1 5'UTR; Stem loop (SL) I, II, III and IV in each structure represent the predicted formation of stem loops, according to those identified in the full length sequence. Predictions were obtained with mFold software (http://mfold.rna.albany.edu/?q=mfold) using default parameters. Gray lines indicate the length of the sequence compared to full length.

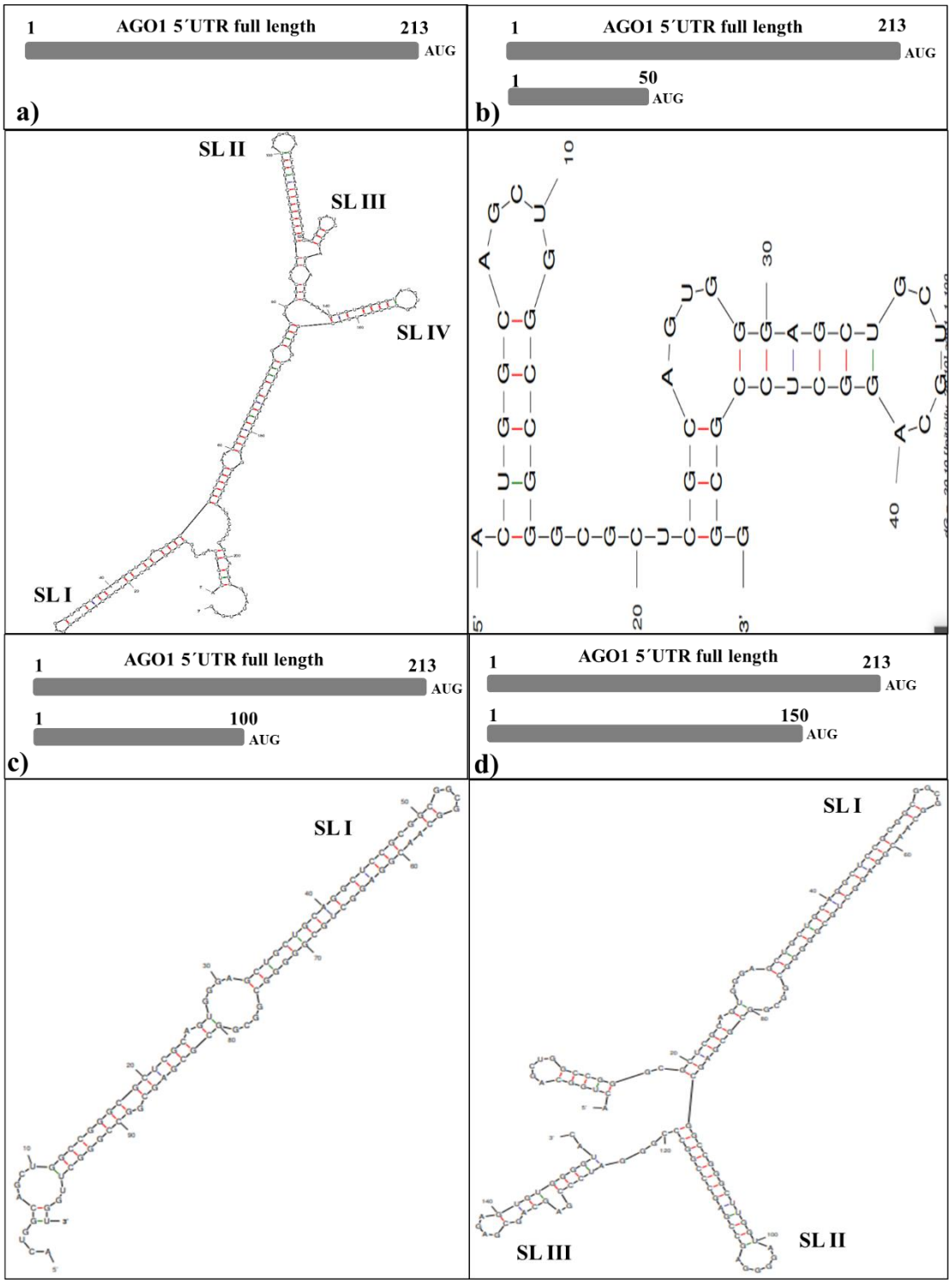

Figure 17. Secondary structures predicted for AGO1 5'UTR with successive deletions in the 3' to 5' direction. a) Full-length AGO1 5'UTR; **b)** nucleotides (nts) 1-50 of AGO1 5'UTR; **c)** nts 1-100 of AGO1 5'UTR; **d)** nts 1-150 of AGO1 5'UTR; Stem loop (SL) I, II, III and IV in each structure represent the predicted formation of stem loops, according to those identified in the full length sequence. Predictions were obtained with mFold software (http://mfold.rna.albany.edu/?q=mfold) using default parameters. Gray lines indicate the length of the sequence compared to full length.

6. Discussion and Future Perspectives

The alternative mechanisms of translation initiation appear as a backup plan, under conditions of inhibition of the canonical translation. Among the alternative mechanisms, one of the best described is the one in which occurs ribosome recruitment to the vicinity of the initiation codon, without prior recognition of the cap m^7G at the 5'end of the mRNA. During cap-independent translation initiation the recruitment of the 40S ribosomal subunit occurs, mostly, through specific cis-elements present on mRNA. These are located on the 5'UTR of the transcript, and either by their sequence or structure, they are able to mediate initiation of translation without the recognition of the cap by the eIF4E. This non-canonical mechanism of translation acts during high energy expenditure or stress conditions, which inhibit cap-dependent translation. Thus, this alternative process maintains protein synthesis required for cell survival and stress response [5]. In humans, several proteins have already been described as able to be translated into stress conditions through cap-independent translation, namely through IRES. Some of these cases are c-Myc [45, 46], VEGF-A under hypoxic conditions [48], XIAP under apoptosis [41], CAT-1 under amino acid deprivation[44], mTOR under hypoxia [55], among others.

Previous work on this thesis done in our laboratory has demonstrated that also the human AGO1 5´UTR can mediate cap-independent translation, which is maintained under stress conditions including hypoxia and endoplasmic reticulum stress. As previously stated, the aim of this work was to continue studying the cap-independent translation of Argonaute 1, and by extension, to understand how the expression levels of the other Argonautes vary under conditions of inhibition of cap-dependent translation.

In order to prove and reinforce the previous results, our first aim was to demonstrate that the human AGO1 5'UTR can promote the cap-independent translation of a reporter gene, using a different bicistronic system. However, until the end of this thesis it was not possible to finish the cloning of the pEGFP_AGO1_mCherry and pEGFP_HBB_mCherry vectors. We think that the problem in cloning the AGO1 5'UTR and HBB 5'UTR in the pEGFP_mCherry vector, is related to the large size difference between the inserts and the vector, which decreases the efficiency of the ligation reaction. In this way we present some protocol solutions that can be conducted in future attempts, to perform this cloning. We first propose that one way of enhancing the binding reaction is to use a larger amount of vector, namely, 100 ng, because when we use this amount of vector, we obtained a larger number of transforming colonies, which may increase the probability of obtaining a positive colony. Another suggestion is to increase the molar ratio vector:insert, so that there is a large amount of insert in relation to the vector, thus trying to potentiate the ligation between them. In addition, we plan to use a different set of primers to perform the colony PCR, in which the forward primer will be located in the insert and the reverse on the vector to avoid false positive results.

In a future perspective, and once the plasmids pEGFP_AGO1_mCherry and pEGFP_HBB_mCherry are obtained, the strategy will be to transfect cells with these plasmids, and then, by confocal microscopy, evaluate the fluorescence of the two reporter genes. Being that in this case we expect that AGO1 5´UTR can drive cap-independent translation of mCherry.

In addition to our previous results that show AGO1 5'UTR is able to mediate the cap-independent translation of a reporter gene under stress conditions, we wanted to understand how the endogenous Argonautes mRNA and protein levels vary under this stress conditions. For this, we used a drug, thapsigargin that induces stress in the endoplasmic reticulum, which consequently leads to the phosphorylation of eIF2α and to the inhibition of canonical translation initiation.

Altogether, our results demonstrate that treatment with 4 μM of thapsigargin induces a decrease in AGO1 and AGO2 mRNA levels, a decrease in Argonaute protein level and a decrease in AGO1 protein level (Fig. 12, 13 and 14). As for the transcript levels it is possible to verify that the treatment with this drug causes a decrease in AGO1 and AGO2 mRNA levels to 54% and 62%, respectively (Fig. 12). Although thapsigargin is described as a translation inhibitor, the impairment consequences of such

important step may affect other cellular processes, in which case it will affect the availability of two transcripts, AGO1 and AGO2. Thus, it is possible to infer that the inhibition of cap-dependent translation induced by thapsigargin causes changes in the turnover or in transcription rate of AGO1 and AGO2 mRNAs. Although there are no described actions of this drug at the mRNA level, this treatment indirectly may be promoting an increase in the degradation rate of mRNA or limiting the synthesis of new mRNA, which leads to a decrease in AGO1 and AGO2 transcript levels. Probably correlated with this, is the decrease in the levels of the Argonaute protein, since with less AGO1 and AGO2 mRNA due to the treatment with thapsigargin, we will have less available transcript and possibly less protein synthesis. Consequently, when there is less AGO1 and AGO2 protein we will have a decrease of the pool of AGOs (Fig. 13).

Finally, we observed that stress induced by thapsigargin leads to decreased levels of Argonaute 1 protein. This result for us was intriguing, since it had previously been observed that the AGO1 5'UTR was able to promote cap-independent translation under stress conditions. So, we had the expectation that even with thapsigargin, the protein synthesis of AGO1 was maintained by an alternative initiation mechanism. However, we believe there are some reasons for this result. First, treatment with 4 μM of thapsigargin decreases the levels of AGO1 mRNA. Thus, even though there may be occurring cap-independent translation of AGO1, there will be less transcript available and less protein produced.

In addition, it should be noted that in these experiments we are measuring the expression of endogenous AGO1, not a reporter gene. Therefore, when we induce stress, the mechanisms of regulation will be more effective in controlling AGO1 than in the reporter gene. For example, it has been described that under hypoxia conditions miRNAs are produced, whose target is 3´UTR of AGO1 mRNA, which cause silencing of the transcript and, consequently, lead to a decrease in the protein levels of AGO1 [132]. Compared with our results, the stress induced by thapsigargin may activate some cellular process that is not directly linked to the translation machinery but will affect the amount of protein synthesized. Therefore, in future experiments it will be convenient to test for other stresses that inhibit cap-dependent translation in HCT116 cells, but that do not cause changes in mRNA levels. This will increase the assurance that we are only measuring alterations in protein synthesis dependent on the translation process. Also, in order to increase the robustness of our results, we also propose to replicate this experience with thapsigargin in other cell lines, in order to see if the same trend is observed.

Furthermore, we believe that there is another factor contributing to the lower amount of AGO1 protein when cells are treated with the drug compared to the control situation. This is because when we inhibit canonical translation, the synthesis of AGO1 is guaranteed by a cap-independent mechanism, and we assume that this alternative pathway will not be as efficient as the canonical process. Our hypothesis is based on the fact that some papers have found that for several cellular mRNAs containing IRES, cap-independent translation is less efficient than cap-dependent translation. This was observed in some research works, that was measured the translational activity of a reporter gene (FLuc) from a monocistronic vector (cap-dependent translation) and a bicistronic vector (cap-independent translation), in which upstream FLuc ORF were c-Myc IRES or Hsp70 IRES. Authors found that FLuc relative activity is much smaller in the second case, suggesting that cap-independent translation is less efficient. Thus, in parallel with our results, AGO1 cap-independent translation may be less efficient when compared to the canonical pathway, which will explain the observed protein levels decrease [133, 134].

Finally, as a future perspective we are going to measure AGO1 cap-independent translation under another type of stress that inhibits translation canonical initiation. In this thesis, we used thapsigargin, which induces the stress of the reticulum and the phosphorylation of eIF2α, however, as previously mentioned, there are IRES whose activity is higher in certain types of stress, since they are present in mRNAs that encode proteins that respond to this stress. For example, in situations of reticulum stress the activity of HIAP2 IRES is increased, or when amino acid starvation occurs there is an increase in the

activity of CAT-1 IRES [86, 90]. Thus, in later experiments it will be convenient to use specific stress conditions that potentiate the cap-independent translation of AGO1.

Since AGO1 5'UTR is able to mediate cap-independent translation, our last task was to try to understand the characteristics of its 5'UTR that enable it to promote this alternative translation initiation and thus, also to try to predict what is the secondary structure adopted by this sequence. As it has been seen, many of these 5'UTRs present complex secondary and tertiary structures that allow an efficient interaction with the translation machinery, thus enabling internal recruitment of the ribosome and/or eIFs in situations of inhibition of cap-dependent translation. The 5'UTRs that contain IRESs are usually long and structured, rich in GC content and may contain upstream initiation codons [36].

Thus, our *in silico* analysis showed that the AGO1 5'UTR has a complex secondary structure, which is organized into four stem loops and which contain on average a GC content of 72.3%, which presumably will confer great stability to this region (Fig. 15). In addition, our deletional *in silico* analysis of the AGO1 5'UTR showed that small deletions in the sequence (Fig. 16 and 17), both in the 5' to 3' and in the 3' to 5 'direction, do not appear to have a large impact on the remaining spatial conformation. In fact, when the SLI (Fig. 16b) or SLIV (Fig.17d) are removed, the structure of the other stem loops does not appear to be significantly affected, which may reflect some structural independence between the four SLs. However with large deletions (more than 100 nts), it is verified that the spatial conformation of the various stem loops changes considerably.

Nevertheless, this *in silico* analysis has its limitations, since it is only a structural prediction, so we intend to do an experimental validation of these results. For this, our strategy would be to synthesize *in vitro* the RNAs corresponding to the AGO1 5'UTR and to the various deleted sequences, and then to evaluate by circular dichroism, a spectroscopic technique, the structure adopted by these fragments, in order to compare to the predictions of the mFold. As well we aim to clone these fragments in our bicistronic vector and measure FLuc activity, in order to identify which portion of AGO1 5’ UTR is able to promote cap-independent translation.

Similar structure/function analyzes have already been done on other sequences, which appear to be able to promote cap-independent translation, namely c-Myc IRES, VEGF IRES and mTOR IRES [51, 54, 55]. For example, c-Myc IRES seems to be organized into two structural domains, the first of which is more complex with a stem loop and an overlapping double pseudoknot motif, while domain 2 contains a single stem loop. In this work, Le Quesne and colleagues did an *in silico* and experimental deletion analysis, in which they removed the structural domains in turn, and verified the effect on the predicted structure and functioning of the IRES. The authors observed computationally that removal of one of the domains does not appear to alter the structure of the other. Experimentally it has been found that the removal of domain 1 or domain 2 leads to a decrease in IRES activity to about 60% whereas the simultaneous deletion of the two leads to a decrease in IRES activity to about 15%.

Overall, our results, indicate that the endoplasmic reticulum stress decreases the level of AGO1 mRNA, thus affecting the amount of protein synthesized. It is therefore necessary to continue studying the cap-independent translation of AGO1, causing that more experiments will be necessary in the future, especially in the sense of perceiving in which stress situations the activity of AGO1 5'UTR is greater and if this alternative form of translation initiation is able to maintain the same level of AGO1 protein synthesis.

7. Bibliography

[1] F. H. Crick, "On protein synthesis.," *Symposia of the Society for Experimental Biology,* vol. 12, pp. 138-163, 1958.

[2] J. D. Watson, T. A. Baker, S. B. Bell, A. Gann, M. Levine, R. Losick and S. C. Harrison, "Molecular Biology," in *SpringerReference*, 2013, p. 872.

[3] D. Ron and H. Harding, "eIF2 Phosphorylation in Cellular Stress Responses and Disease," *Translational Control in Biology and Medicine,* pp. 349-372, 2007.

[4] L. D. Kapp and J. R. Lorsch, "The Molecular Mechanics of Eukaryotic Translation," *Annual Review of Biochemistry,* vol. 73, no. 1, pp. 657-704, 2004.

[5] R. Lacerda, J. Menezes and L. Romão, *More than just scanning: the importance of cap-independent mRNA translation initiation for cellular stress response and cancer,* 2017.

[6] L. D. Kapp and J. R. Lorsch, "GTP-dependent Recognition of the Methionine Moiety on Initiator tRNA by Translation Factor eIF2," *Journal of Molecular Biology,* vol. 335, no. 4, pp. 923-936, 2004.

[7] W. C. Merrick, "Eukaryotic protein synthesis: Still a mystery," *Journal of Biological Chemistry,* vol. 285, no. 28, pp. 21197-21201, 2010.

[8] A. Des Georges, V. Dhote, L. Kuhn, C. U. Hellen, T. V. Pestova, J. Frank and Y. Hashem, "Structure of mammalian eIF3 in the context of the 43S preinitiation complex," *Nature,* vol. 525, no. 7570, pp. 491-495, 2015.

[9] N. Sonenberg and A. G. Hinnebusch, "Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets," *Cell,* vol. 136, no. 4, pp. 731-745, 2009.

[10] A. A. Komar, B. Mazumder and W. C. Merrick, "A new framework for understanding IRES-mediated translation," *Gene,* vol. 502, no. 2, pp. 75-86, 2012.

[11] M. Kozak, "An analysis of vertebrate mRNA sequences: Intimations of translational control," *Journal of Cell Biology,* vol. 115, no. 4, pp. 887-903, 1991.

[12] M. Kozak, "Pushing the limits of the scanning mechanism for initiation of translation," *Gene,* vol. 299, no. 1-2, pp. 1-34, 2002.

[13] A. R. Özeş, K. Feoktistova, B. C. Avanzino and C. S. Fraser, "Duplex unwinding and ATPase activities of the DEAD-box helicase eIF4A are coupled by eIF4G and eIF4B," *Journal of Molecular Biology,* vol. 412, no. 4, pp. 674-687, 2011.

[14] W. C. Merrick, "eIF4F: A retrospective," *Journal of Biological Chemistry,* vol. 290, no. 40, pp. 24091-24099, 2015.

[15] C. Lind and J. Åqvist, "Principles of start codon recognition in eukaryotic translation initiation," *Nucleic Acids Research,* vol. 44, no. 17, pp. 8425-8432, 2016.

[16] T. V. Pestova, V. G. Kolupaeva, I. B. Lomakin, E. V. Pilipenko, I. N. Shatsky, V. I. Agol and C. U. T. Hellen, "Molecular mechanisms of translation initiation in eukaryotes," *Proceedings of the National Academy of Sciences,* vol. 98, no. 13, pp. 7029-7036, 2001.

[17] B. S. Shin, D. Maag, A. Roll-Mecak, M. S. Arefin, S. K. Burley, J. R. Lorsch and T. E. Dever, "Uncoupling of initiation factor eIF5B/IF2 GTPase and translational activities by mutations that lower ribosome affinity," *Cell,* vol. 111, no. 7, pp. 1015-1025, 2002.

[18] T. E. Dever and R. Green, "The elongation, termination, and recycling phases of translation in eukaryotes," *Cold Spring Harbor Perspectives in Biology,* vol. 4, no. 7, pp. 1-16, 2012.

[19] T. M. Schmeing, R. M. Voorhees, A. C. Kelley and V. Ramakrishnan, "How mutations in tRNA distant from the anticodon affect the fidelity of decoding," *Nature Structural and Molecular Biology,* vol. 18, no. 4, pp. 432-437, 2011.

[20] C. U. Hellen, "Translation Termination and Ribosome Recycling in Eukaryotes," *Cold Spring Harbor Perspectives in Biology,* p. a032656, 2018.

[21] A. V. Pisarev, M. A. Skabkin, V. P. Pisareva, O. V. Skabkina, A. M. Rakotondrafara, M. W. Hentze, C. U. Hellen and T. V. Pestova, "The Role of ABCE1 in Eukaryotic Posttermination Ribosomal Recycling," *Molecular Cell,* vol. 37, no. 2, pp. 196-210, 2010.

[22] C. J. Shoemaker and R. Green, "Kinetic analysis reveals the ordered coupling of translation termination and ribosome recycling in yeast," *Proceedings of the National Academy of Sciences,* 2011.

[23] H. Goumans, A. Thomas, A. Verhoeven, H. O. Voorma and R. Benne, "The role of eIF-4C in protein synthesis initiation complex formation," *BBA Section Nucleic Acids And Protein Synthesis,* 1980.

[24] I. N. Shatsky, I. M. Terenin, V. V. Smirnova and D. E. Andreev, "Cap-Independent Translation: What's in a Name?," *Trends in Biochemical Sciences,* vol. xx, pp. 1-14, 2018.

[25] E. Martínez-Salas, D. Piñeiro and N. Fernández, "Alternative mechanisms to initiate translation in eukaryotic mRNAs," *Comparative and Functional Genomics,* vol. 2012, 2012.

[26] J. S. Kieft, "Viral IRES RNA structures and ribosome interactions," *Trends in Biochemical Sciences,* vol. 33, no. 6, pp. 274-283, 2008.

[27] D. E. Andreev, S. E. Dmitriev, R. Zinovkin, I. M. Terenin and I. N. Shatsky, "The 5′ untranslated region of Apaf-1 mRNA directs translation under apoptosis conditions via a 5′ end-dependent scanning mechanism," *FEBS Letters,* vol. 586, no. 23, pp. 4139-4143, 2012.

[28] I. M. Terenin, D. E. Andreev, S. E. Dmitriev and I. N. Shatsky, "A novel mechanism of eukaryotic translation initiation that is neither m7G-cap-, nor IRES-dependent," *Nucleic Acids Research,* vol. 41, no. 3, pp. 1807-1816, 2013.

[29] I. N. Shatsky, S. E. Dmitriev, I. M. Terenin and D. E. Andreev, "Cap- and IRES-independent scanning mechanism of translation initiation as an alternative to the concept of cellular IRESs," *Molecules and Cells,* vol. 30, no. 4, pp. 285-293, 2010.

[30] K. D. Meyer, D. P. Patil, J. Zhou, A. Zinoviev, M. A. Skabkin, O. Elemento, T. V. Pestova, S. B. Qian and S. R. Jaffrey, "5′ UTR m6A Promotes Cap-Independent Translation," *Cell,* 2015.

[31] J. Pelletier and N. Sonenberg, "Internal initiation of translation of eukaryotic mRNA directed by a sequence derived from poliovirus RNA," *Nature,* vol. 334, no. 6180, pp. 320-325, 1988.

[32] T. V. Pestova, S. I. Borukhov and C. U. Hellen, “Eukaryotic ribosomes require initiation factors 1 and 1A to locate initiation codons,” *Nature,* vol. 394, no. 6696, pp. 854-859, 1998.

[33] P. J. Lukavsky, “Structure and function of HCV IRES domains,” *Virus Research,* vol. 139, no. 2, pp. 166-171, 2009.

[34] J. Marcotrigiano, I. B. Lomakin, N. Sonenberg, T. V. Pestova, C. U. Hellen and S. K. Burley, “A conserved HEAT domain within eIF4G Directs Assembly of the Translation Initiation Machinery,” *Molecular Cell,* vol. 7, no. 1, pp. 193-203, 2001.

[35] O. Fernández-Miragall, S. L. d. Quinto and E. Martínez-Salas, “Relevance of RNA structure for the activity of picornavirus IRES elements,” *Virus Research,* vol. 139, no. 2, pp. 172-182, 2009.

[36] A. A. Komar and M. Hatzoglou, “Cellular IRES-mediated translation: The war of ITAFs in pathophysiological states,” *Cell Cycle,* vol. 10, no. 2, pp. 229-240, 2011.

[37] V. I. Agol, “Translational Control of the Picornavirus Phenotype,” *Proteomics,* vol. 35, no. 4, pp. 1191-1199, 2001.

[38] D. G. Macejak and P. Sarnow, “Internal initiation of translation mediated by the 5′ leader of a cellular mRNA,” *Nature,* 1991.

[39] S. Weingarten-Gabbay, S. Elias-Kirma, R. Nir, A. A. Gritsenko, N. Stern-Ginossar, Z. Yakhini, A. Weinberger and E. Segal, “Comparative genetics: Systematic discovery of cap-independent translation sequences in human and viral genomes,” *Science,* vol. 351, no. 6270, 2016.

[40] M. M. B. Hershey John W.B., Sonenberg Nahum, “Principles of Translational Control: An Overview,” *Cold Spring Harb Perspect Biol,* vol. 4, no. 12, 2012.

[41] A. Riley, L. E. Jordan and M. Holcik, “Distinct 5′ UTRs regulate XIAP expression under normal growth conditions and during cellular stress,” *Nucleic Acids Research,* vol. 38, no. 14, pp. 4665-4674, 2010.

[42] K. W. Sherrill, M. P. Byrd, M. E. V. Eden and R. E. Lloyd, “BCL-2 Translation Is Mediated via Internal Ribosome Entry during Cell Stress *,” vol. 279, no. 28, pp. 29066-29074, 2004.

[43] V. Marcel, S. E. Ghayad, S. Belin, G. Therizols, A. P. Morel, E. Solano-Gonzàlez, J. A. Vendrell, S. Hacot, H. C. Mertani, M. A. Albaret, J. C. Bourdon, L. Jordan, A. Thompson, Y. Tafer, R. Cong, P. Bouvet, J. C. Saurin, F. Catez, A. C. Prats, A. Puisieux and J. J. Diaz, “P53 Acts as a Safeguard of Translational Control by Regulating Fibrillarin and rRNA Methylation in Cancer,” *Cancer Cell,* vol. 24, no. 3, pp. 318-330, 2013.

[44] J. Fernandez, I. Yaman, R. Mishra, W. C. Merrick, M. D. Snider, W. H. Lamers and M. Hatzoglou, “Internal Ribosome Entry Site-mediated Translation of a Mammalian mRNA Is Regulated by Amino Acid Availability,” *Journal of Biological Chemistry,* vol. 276, no. 15, pp. 12285-12291, 2001.

[45] M. Stoneley, F. E. Paulin, J. P. Le Quesne, S. A. Chappell and A. E. Willis, “C-Myc 5' untranslated region contains an internal ribosome entry segment,” *Oncogene,* vol. 16, no. 3, pp. 423-428, 1998.

[46] K. A. Spriggs, L. C. Cobbold, C. L. Jopling, R. E. Cooper, L. A. Wilson, M. Stoneley, M. J. Coldwell, D. Poncet, Y.-C. Shen, S. J. Morley, M. Bushell and A. E. Willis, “Canonical Initiation

Factor Requirements of the Myc Family of Internal Ribosome Entry Segments," *Molecular and Cellular Biology*, vol. 29, no. 6, pp. 1565-1574, 2009.

[47] A. A. Komar and M. Hatzoglou, "Internal ribosome entry sites in cellular mRNAs: mystery of their existence," 2005.

[48] S. Braunstein, K. Karpisheva, C. Pola, J. Goldberg, T. Hochman, H. Yee, J. Cangiarella, R. Arju, S. C. Formenti and R. J. Schneider, "A Hypoxia-Controlled Cap-Dependent to Cap-Independent Translation Switch in Breast Cancer," *Molecular Cell*, vol. 28, no. 3, pp. 501-512, 2007.

[49] C. U. T. Hellen and P. Sarnow, "Internal ribosome entry sites in eukaryotic mRNA molecules," *Genes and Development*, vol. 15, no. 13, pp. 1593-1612, 2001.

[50] M. Stoneley and A. E. Willis, "Cellular internal ribosome entry segments: Structures, trans-acting factors and regulation of gene expression," *Oncogene*, vol. 23, no. 18, pp. 3200-3207, 2004.

[51] J. P. Le Quesne, M. Stoneley, G. A. Fraser and A. E. Willis, "Derivation of a structural model for the c-myc IRES," *Journal of Molecular Biology*, vol. 310, no. 1, pp. 111-126, 2001.

[52] C. L. Jopling, K. Spriggs, S. Mitchell, M. Stoneley and A. Willis, "L-Myc protein synthesis is initiated by internal ribosome entry," *Rna*, vol. 10, no. 2, pp. 287-298, 2004.

[53] S. K. Pedersen, J. Christiansen, T. Hansen, M. R. Larsen and F. C. Nielsen, "Human insulin-like growth factor II leader 2 mediates internal initiation of translation," *Biochem J*, vol. 363, no. Pt 1, pp. 37-44, 2002.

[54] I. Huez, L. Créancier, S. Audigier, M. C. Gensac, A. C. Prats and H. Prats, "Two independent internal ribosome entry sites are involved in translation initiation of vascular endothelial growth factor mRNA.," *Molecular and cellular biology*, vol. 18, no. 11, pp. 6178-90, 1998.

[55] A. Marques-Ramos, M. M. Candeias, J. Menezes, R. Lacerda, M. Willcocks, A. Teixeira, N. Locker and L. Romão, "Cap-independent translation ensures mTOR expression and function upon protein synthesis inhibition," *Rna*, vol. 23, no. 11, pp. 1712-1728, 2017.

[56] H. A. King, L. C. Cobbold and A. E. Willis, "The role of IRES trans -acting factors in regulating translation initiation," pp. 1581-1586, 2010.

[57] M. Bushell, M. Stoneley, Y. W. Kong, T. L. Hamilton, K. A. Spriggs, H. C. Dobbyn, X. Qin, P. Sarnow and A. E. Willis, "Polypyrimidine Tract Binding Protein Regulates IRES-Mediated Gene Expression during Apoptosis," *Molecular Cell*, vol. 23, no. 3, pp. 401-412, 2006.

[58] M. Majumder, I. Yaman, F. Gaccioli, V. V. Zeenko, C. Wang, M. G. Caprara, R. C. Venema, A. A. Komar, M. D. Snider and M. Hatzoglou, "The hnRNA-Binding Proteins hnRNP L and PTB Are Required for Efficient Translation of the Cat-1 Arginine/Lysine Transporter mRNA during Amino Acid Starvation," *Molecular and Cellular Biology*, vol. 29, no. 10, pp. 2899-2912, 2009.

[59] D. Rübsamen, J. S. Blees, K. Schulz, C. Döring, M. L. Hansmann, H. Heide, A. Weigert, T. Schmid and B. Brüne, "IRES-dependent translation of egr2 is induced under inflammatory conditions," *Rna*, vol. 18, no. 10, pp. 1910-1920, 2012.

[60] R. G. Fred, C. H. Bang-Berthelsen, T. Mandrup-Poulsen, L. G. Grunnet and N. Welsh, "High Glucose Suppresses Human Islet Insulin Biosynthesis by Inducing Mir-133a Leading to Decreased Polypyrimidine Tract Binding Protein-Expression," *PLoS ONE*, vol. 5, no. 5, 2010.

[61] R. Grover, P. S. Ray and S. Das, “Polypyrimidine tract binding protein regulates IRES-mediated translation of p53 isoforms,” *Cell Cycle,* vol. 7, no. 14, pp. 2189-2198, 2008.

[62] M. G. Romanelli, E. Diani and P. M. J. Lievens, “New insights into functional roles of the polypyrimidine tract-binding protein,” *International Journal of Molecular Sciences,* vol. 14, no. 11, pp. 22906-22932, 2013.

[63] M. Holcik and R. G. Korneluk, “Functional characterization of the X-linked inhibitor of apoptosis (XIAP) internal ribosome entry site element: role of La autoantigen in XIAP translation.,” *Molecular and cellular biology,* vol. 20, no. 13, pp. 4648-57, 2000.

[64] S. a. Tinton, B. Schepens, Y. Bruynooghe, R. Beyaert and S. Cornelis, “Regulation of the cell-cycle-dependent internal ribosome entry site of the PITSLRE protein kinase: roles of Unr (upstream of N-ras) protein and phosphorylated translation initiation factor eIF-2alpha.,” *The Biochemical journal,* vol. 385, pp. 155-163, 2005.

[65] N. Liberman, V. Gandin, Y. V. Svitkin, M. David, G. Virgili, M. Jaramillo, M. Holcik, B. Nagar, A. Kimchi and N. Sonenberg, “DAP5 associates with eIF2β and eIF4AI to promote Internal Ribosome Entry Site driven translation,” *Nucleic Acids Research,* vol. 43, no. 7, pp. 3764-3775, 2015.

[66] Y. K. Kim, B. Hahm and S. K. Jang, “Polypyrimidine tract-binding protein inhibits translation of bip mRNA,” *Journal of Molecular Biology,* vol. 304, no. 2, pp. 119-133, 2000.

[67] S. M. Lewis and M. Holcik, “For IRES trans-acting factors, it is all about location,” *Oncogene,* vol. 27, no. 8, pp. 1033-1035, 2008.

[68] F. Damiano, A. Rochira, R. Tocci, S. Alemanno, A. Gnoni and L. Siculella, “hnRNP A1 mediates the activation of the IRES-dependent SREBP-1a mRNA translation in response to endoplasmic reticulum stress,” *Biochemical Journal,* vol. 449, no. 2, pp. 543-553, 2013.

[69] N. Thakor and M. Holcik, “IRES-mediated translation of cellular messenger RNA operates in eIF2α-independent manner during stress,” *Nucleic Acids Research,* vol. 40, no. 2, pp. 541-552, 2012.

[70] B. Walters and S. R. Thompson, “Cap-Independent Translational Control of Carcinogenesis,” *Frontiers in Oncology,* vol. 6, no. May, pp. 1-8, 2016.

[71] M. Holcik and N. Sonenberg, “Translational control in stress and apoptosis,” *Nature Reviews Molecular Cell Biology,* vol. 6, no. 4, pp. 318-327, 2005.

[72] V. M. Pain, “Initiation of protein synthesis in eukaryotic cells,” *European Journal of Biochemistry,* vol. 236, no. 3, pp. 747-771, 1996.

[73] H.-S. Yang, A. P. Jansen, A. A. Komar, X. Zheng, W. C. Merrick, S. Costes, S. J. Lockett, N. Sonenberg and N. H. Colburn, “The Transformation Suppressor Pdcd4 Is a Novel Eukaryotic Translation Initiation Factor 4A Binding Protein That Inhibits Translation,” *Molecular and Cellular Biology,* vol. 23, no. 1, pp. 26-37, 2003.

[74] M. Kozak, “Downstream secondary structure facilitates recognition of initiator codons by eukaryotic ribosomes.,” *Proceedings of the National Academy of Sciences,* vol. 87, no. 21, pp. 8301-8305, 1990.

[75] M. Kozak, “Regulation of translation via mRNA structure in prokaryotes and eukaryotes,” *Gene,* vol. 361, no. 1-2, pp. 13-37, 2005.

[76] H.-J. Lo, H.-K. Huang and T. F. Donahue, “RNA Polymerase I-Promoted HIS4 Expression Yields Uncapped, Polyadenylated mRNA That Is Unstable and Inefficiently Translated in Saccharomyces cerevisiae,” *Mol. Cell. Biol.,* vol. 18, no. 2, pp. 665-675, 1998.

[77] G. L. Semenza, “Regulation of hypoxia-induced angiogenesis: a chaperone escorts VEGF to the dance,” *The Journal of clinical investigation,* vol. 108, no. 1, pp. 39-40, 2001.

[78] I. Stein, A. Itin and P. Einat, “Translation of vascular endothelial growth factor mRNA by internal ribosome entry: implications for translation under hypoxia,” *Molecular and cellular biology,* vol. 61, no. 3, pp. 161-7, 1998.

[79] F. Morfoisse, A. Kuchnio, C. Frainay, A. Gomez-Brouchet, M. B. Delisle, S. Marzi, A. C. Helfer, F. Hantelys, F. Pujol, J. Guillermet-Guibert, C. Bousquet, M. Dewerchin, S. Pyronnet, A. C. Prats, P. Carmeliet and B. Garmy-Susini, “Hypoxia induces VEGF-C expression in metastatic tumor cells via a HIF-1α-independent translation-mediated mechanism,” *Cell Reports,* vol. 6, no. 1, pp. 155-167, 2014.

[80] K. J. D. Lang, A. Kappel and G. J. Goodall, “Hypoxia-inducible Factor-1alpha mRNA Contains an Internal Ribosome Entry Site That Allows Efficient Translation during Normoxia and Hypoxia,” *Molecular Biology of the Cell,* vol. 13, no. 5, pp. 1792-1801, 2002.

[81] J. J. D. Blais, V. Filipenko, M. Bi, H. H. P. Harding, D. Ron, C. Koumenis, B. G. Wouters and J. C. Bell, “Activating Transcription Factor 4 Is Translationally Regulated by Hypoxic Stress,” *Molecular and Cellular Biology,* vol. 24, no. 17, pp. 7469-7482, 2004.

[82] B. Schepens, S. A. Tinton, Y. Bruynooghe, R. Beyaert and S. Cornelis, “The polypyrimidine tract-binding protein stimulates HIF-1α IRES-mediated translation during hypoxia,” *Nucleic Acids Research,* vol. 33, no. 21, pp. 6884-6894, 2005.

[83] B. Schepens, S. A. Tinton, Y. Bruynooghe, E. Parthoens, M. Haegman, R. Beyaert and S. Cornelis, “A role for hnRNP C1/C2 and Unr in internal initiation of translation during mitosis,” *EMBO Journal,* vol. 26, no. 1, pp. 158-169, 2007.

[84] S. Pyronnet, L. Pradayrol and N. Sonenberg, “A cell cycle-dependent internal ribosome entry site,” *Molecular Cell,* vol. 5, no. 4, pp. 607-616, 2000.

[85] K. A. Spriggs, L. C. Cobbold, S. H. Ridley, M. Coldwell, A. Bottley, M. Bushell, A. E. Willis and K. Siddle, “The human insulin receptor mRNA contains a functional internal ribosome entry segment,” *Nucleic Acids Research,* vol. 37, no. 17, pp. 5881-5893, 2009.

[86] M. Hatzoglou, J. Fernandez, I. Yaman and E. Closs, “Regulation of cationic amino acid transport: the story of the CAT-1 transporter.,” *Annual Review of Nutrition,* vol. 24, no. 1, pp. 377-399, 2004.

[87] S. Palii, M. Thiaville, Y.-X. Pan, C. Zhong and M. Kilberg, “Characterization of the amino acid response element within the human sodium-coupled neutral amino acid transporter 2 (SNAT2) System A transporter gene,” *Biochemical Journal,* vol. 395, no. 3, pp. 517-527, 2006.

[88] F. Gaccioli, C. C. Huang, C. Wang, E. Bevilacqua, R. Franchi-Gazzola, G. C. Gazzola, O. Bussolati, M. D. Snider and M. Hatzoglou, “Amino acid starvation induces the SNAT2 neutral amino acid transporter by a mechanism that involves eukaryotic initiation factor 2α

phosphorylation and cap-independent translation," *Journal of Biological Chemistry,* vol. 281, no. 26, pp. 17929-17940, 2006.

[89] M. Stoneley, S. A. Chappell, C. L. Jopling, M. Dickens, M. Macfarlane, A. E. Willis and M. M. A. C. Farlane, "c-Myc Protein Synthesis Is Initiated from the Internal Ribosome Entry Segment during Apoptosis c-Myc Protein Synthesis Is Initiated from the Internal Ribosome Entry Segment during Apoptosis," *Mol Cell Biol,* vol. 20, no. 4, pp. 1162-1169, 2000.

[90] D. Warnakulasuriyarachchi, S. Cerquozzi, H. H. Cheung and M. Holcík, "Translational Induction of the Inhibitor of Apoptosis Protein HIAP2 during Endoplasmic Reticulum Stress Attenuates Cell Death and Is Mediated via an Inducible Internal Ribosome Entry Site Element," *Journal of Biological Chemistry,* vol. 279, no. 17, pp. 17148-17157, 2004.

[91] T. Subkhankulova, S. A. Mitchell and A. E. Willis, "Synthesis Following Genotoxic Stress," *Cell,* vol. 192, pp. 183-192, 2001.

[92] W. Li, N. Thakor, E. Y. Xu, Y. Huang, C. Chen, R. Yu, M. Holcik and A. N. Kong, "An internal ribosomal entry site mediates redox-sensitive translation of Nrf2," *Nucleic Acids Research,* vol. 38, no. 3, pp. 778-788, 2009.

[93] J. Zhang, T. N. Dinh, K. Kappeler, G. Tsaprailis and Q. M. Chen, "La Autoantigen Mediates Oxidant Induced De Novo Nrf2 Protein Translation," *Molecular & Cellular Proteomics,* vol. 11, no. 6, pp. M111-015032, 2012.

[94] D. Silvera, R. Arju, F. Darvishian, P. H. Levine, L. Zolfaghari, J. Goldberg, T. Hochman, S. C. Formenti and R. J. Schneider, "Essential role for eIF4GI overexpression in the pathogenesis of inflammatory breast cancer," *Nature Cell Biology,* vol. 11, no. 7, pp. 903-908, 2009.

[95] T. Muranen, L. M. Selfors, D. T. Worster, M. P. Iwanicki, L. Song, F. C. Morales, S. Gao, G. B. Mills and J. S. Brugge, "Inhibition of PI3K/mTOR Leads to Adaptive Resistance in Matrix-Attached Cancer Cells," *Cancer Cell,* vol. 21, no. 2, pp. 227-239, 2012.

[96] P. Hainaut, T. Hernandez, A. Robinson, P. Rodriguez-Tome, T. Flores, M. Hollstein, C. C. Harris and R. Montesano, "IARC database of p53 gene mutations in human tumors and cell lines: Updated compilation, revised formats and new visualisation tools," *Nucleic Acids Research,* vol. 26, no. 1, pp. 205-213, 1998.

[97] M. J. Halaby, Y. Li, B. R. Harris, S. Jiang, W. K. Miskimins, M. P. Cleary and D. Q. Yang, "Translational control protein 80 stimulates IRES-mediated translation of p53 mRNA in response to DNA damage," *BioMed Research International,* vol. 2015, 2015.

[98] M.-J. Halaby, B. R. E. Harris, W. K. Miskimins, M. P. Cleary and D.-Q. Yang, "Deregulation of Internal Ribosome Entry Site-Mediated p53 Translation in Cancer Cells with Defective p53 Response to DNA Damage," *Molecular and Cellular Biology,* vol. 35, no. 23, pp. 4006-4017, 2015.

[99] P. W. Vesely, P. B. Staber, G. Hoefler and L. Kenner, "Translational regulation mechanisms of AP-1 proteins," *Mutation Research - Reviews in Mutation Research,* vol. 682, no. 1, pp. 7-12, 2009.

[100] S. Chakraborti, M. Mandal, S. Das, A. Mandal and T. Chakraborti, "Regulation of matrix metalloproteinases. An overview," *Molecular and Cellular Biochemistry,* vol. 253, no. 1-2, pp. 269-285, 2003.

[101] P. Polak, A. Oren, I. Ben-Dror, D. Steinberg, S. Sapoznik, A. Arditi-Duvdevany and L. Vardimon, "The cytoskeletal network controls c-Jun translation in a UTR-dependent manner," *Oncogene,* vol. 25, no. 5, pp. 665-676, 2006.

[102] R. Knirsh, "Loss of E-Cadherin–mediated Cell–Cell Contacts Activates a Novel Mechanism for Up-Regulation of the Proto-Oncogene c-Jun," *Molecular Biology of the Cell,* vol. 65, no. 2, pp. 87-102, 2009.

[103] L. Blau, R. Knirsh, I. Ben-Dror, S. Oren, S. Kuphal, P. Hau, M. Proescholdt, A.-K. Bosserhoff and L. Vardimon, "Aberrant expression of c-Jun in glioblastoma by internal ribosome entry site (IRES)-mediated translational activation," *Proceedings of the National Academy of Sciences,* vol. 109, no. 42, pp. E2875-E2884, 2012.

[104] B. Spangler, L. Vardimon, A. K. Bosserhoff and S. Kuphal, "Post-transcriptional regulation controlled by E-cadherin is important for c-Jun activity in melanoma," *Pigment Cell and Melanoma Research,* vol. 24, no. 1, pp. 148-164, 2011.

[105] J. Höck and G. Meister, "The Argonaute protein family," *Genome Biology,* vol. 9, no. 2, 2008.

[106] M. A. Carmell, Z. Xuan, M. Q. Zhang and G. J. Hannon, "The Argonaute family: Tentacles that reach into RNAi, developmental control, stem cell maintenance, and tumorigenesis," *Genes and Development,* vol. 16, no. 21, pp. 2733-2742, 2002.

[107] K. Bohmert, I. Camus, C. Bellini, D. Bouchez, M. Caboche and C. Banning, "AGO1 defines a novel locus of Arabidopsis controlling leaf development," *EMBO Journal,* vol. 17, no. 1, pp. 170-180, 1998.

[108] G. Meister, "Argonaute proteins: Functional insights and emerging roles," *Nature Reviews Genetics,* vol. 14, no. 7, pp. 447-459, 2013.

[109] Y. R. Yuan, Y. Pei, J. B. Ma, V. Kuryavyi, M. Zhadina, G. Meister, H. Y. Chen, Z. Dauter, T. Tuschl and D. J. Patel, "Crystal structure of A. aeolicus argonaute, a site-specific DNA-guided endoribonuclease, provides insights into RISC-mediated mRNA cleavage," *Molecular Cell,* vol. 19, no. 3, pp. 405-419, 2005.

[110] E. Elkayam, C. D. Kuhn, A. Tocilj, A. D. Haase, E. M. Greene, G. J. Hannon and L. Joshua-Tor, "The structure of human argonaute-2 in complex with miR-20a," *Cell,* vol. 150, no. 1, pp. 100-110, 2012.

[111] M. Jinek and J. A. Doudna, "A three-dimensional view of the molecular machinery of RNA interference," *Nature,* vol. 457, no. 7228, pp. 405-412, 2009.

[112] P. B. Kwak and Y. Tomari, "The N domain of Argonaute drives duplex unwinding during RISC assembly," *Nature Structural and Molecular Biology,* vol. 19, no. 2, pp. 145-151, 2012.

[113] F. Frank, N. Sonenberg and B. Nagar, "Structural basis for 5′-nucleotide base-specific recognition of guide RNA by human AGO2," *Nature,* vol. 465, no. 7299, pp. 818-822, 2010.

[114] S. Werfel, S. Leierseder, B. Ruprecht, B. Kuster and S. Engelhardt, "Preferential MicroRNA targeting revealed by in vivo competitive binding and differential Argonaute immunoprecipitation," *Nucleic Acids Research,* vol. 45, no. 17, pp. 10218-10228, 2017.

[115] J. Winter and S. Diederichs, "Argonaute-3 activates the let-7a passenger strand microRNA," *RNA Biology,* vol. 10, no. 10, pp. 1631-1643, 2013.

[116] D. Zilberman, "Role of Arabidopsis ARGONAUTE4 in RNA-Directed DNA Methylation Triggered by Inverted Repeats," *Indian Veterinary Journal,* vol. 91, no. 9, pp. 20-22, 2014.

[117] M. Alló, V. Buggiano, J. P. Fededa, E. Petrillo, I. Schor, M. De La Mata, E. Agirre, M. Plass, E. Eyras, S. A. Elela, R. Klinck, B. Chabot and A. R. Kornblihtt, "Control of alternative splicing through siRNA-mediated transcriptional gene silencing," *Nature Structural and Molecular Biology,* vol. 16, no. 7, pp. 717-724, 2009.

[118] W. Wei, Z. Ba, M. Gao, Y. Wu, Y. Ma, S. Amiard, C. I. White, J. M. R. Danielsen, Y. G. Yang and Y. Qi, "A role for small RNAs in DNA double-strand break repair," *Cell,* vol. 149, no. 1, pp. 101-112, 2012.

[119] S. Rüdel and G. Meister, "Phosphorylation of Argonaute proteins: regulating gene regulators: Figure 1," *Biochemical Journal,* vol. 413, no. 3, pp. e7-e9, 2008.

[120] A. K. L. Leung, J. M. Calabrese and P. A. Sharp, "Quantitative analysis of Argonaute protein reveals microRNA-dependent localization to stress granules," *Proceedings of the National Academy of Sciences,* vol. 103, no. 48, pp. 18125-18130, 2006.

[121] J. Liu, M. A. Valencia-Sanchez, G. J. Hannon and R. Parker, "MicroRNA-dependent localization of targeted mRNAs to mammalian P-bodies," *Nature Cell Biology,* vol. 7, no. 7, pp. 719-723, 2005.

[122] D. Völler, L. Linck, A. Bruckmann, J. Hauptmann, R. Deutzmann, G. Meister and A. K. Bosserhoff, "Argonaute family protein expression in normal tissue and cancer entities," *PLoS ONE,* vol. 11, no. 8, pp. 1-14, 2016.

[123] Y. Kataoka, M. Takeichi and T. Uemura, "Developmental roles and molecular characterization of a Drosophila homologue of Arabidopsis Argonaute1, the founder of a novel gene superfamily," *Genes to Cells,* vol. 6, no. 4, pp. 313-325, 2001.

[124] D. Wang, Z. Zhang, E. O'Loughlin, T. Lee, S. Houel, D. O'Carroll, A. Tarakhovsky, N. G. Ahn and R. Yi, "Quantitative functions of argonaute proteins in mammalian development," *Genes and Development,* vol. 26, no. 7, pp. 693-704, 2012.

[125] C. R. Faehnle, E. Elkayam, A. D. Haase, G. J. Hannon and L. Joshua-Tor, "The Making of a Slicer: Activation of Human Argonaute-1," *Cell Reports,* 2013.

[126] A. Dueck, C. Ziegler, A. Eichner, E. Berezikov and G. Meister, "MicroRNAs associated with the different human Argonaute proteins," *Nucleic Acids Research,* vol. 40, no. 19, pp. 9850-9862, 2012.

[127] L. Li, C. Yu, H. Gao and Y. Li, "Argonaute proteins: Potential biomarkers for human colon cancer," *BMC Cancer,* vol. 10, pp. 1-8, 2010.

[128] C. Parisi, C. Giorgi, E. M. Batassa, L. Braccini, G. Maresca, I. D'Agnano, V. Caputo, A. Salvatore, F. Pietrolati, C. Cogoni and C. Catalanotto, "Ago1 and Ago2 differentially affect cell proliferation, motility and apoptosis when overexpressed in SH-SY5Y neuroblastoma cells," *FEBS Letters,* vol. 585, no. 19, pp. 2965-2971, 2011.

[129] S. R. Thompson, “So you want to know if your message has an IRES?,” *Wiley Interdisciplinary Reviews: RNA*, vol. 3, no. 5, pp. 697-705, 2012.

[130] M. M. Candeias, D. J. Powell, E. Roubalova, S. Apcher, K. Bourougaa, B. Vojtesek, H. Bruzzoni-Giovanelli and R. Fåhraeus, “Expression of p53 and p53/47 are controlled by alternative mechanisms of messenger RNA translation initiation,” *Oncogene*, vol. 25, no. 52, pp. 6936-6947, 2006.

[131] Y. A. Bochkov and A. C. Palmenberg, “Translational efficiency of EMCV IRES in bicistronic vectors is dependent upon IRES sequence and gene location,” *BioTechniques*, vol. 41, no. 3, pp. 283-292, 2006.

[132] Z. Chen, T. C. Lai, Y. H. Jan, F. M. Lin, W. C. Wang, H. Xiao, Y. T. Wang, W. Sun, X. Cui, Y. S. Li, T. Fang, H. Zhao, C. Padmanabhan, R. Sun, D. L. Wang, H. Jin, G. Y. Chau, H. D. Huang, M. Hsiao and J. Y. Shyy, “Hypoxia-responsive miRNAs target argonaute 1 to promote angiogenesis,” *Journal of Clinical Investigation*, vol. 123, no. 3, pp. 1057-1067, 2013.

[133] W. V. Gilbert, “Alternative ways to think about cellular internal ribosome entry,” *Journal of Biological Chemistry*, vol. 285, no. 38, pp. 29033-29038, 2010.

[134] D. E. Andreev, S. E. Dmitriev, I. M. Terenin, V. S. Prassolov, W. C. Merrick and I. N. Shatsky, “Differential contribution of the m7G-cap to the 5′ end-dependent translation initiation of mammalian mRNAs,” *Nucleic Acids Research*, vol. 37, no. 18, pp. 6135-6147, 2009.

8. Appendix

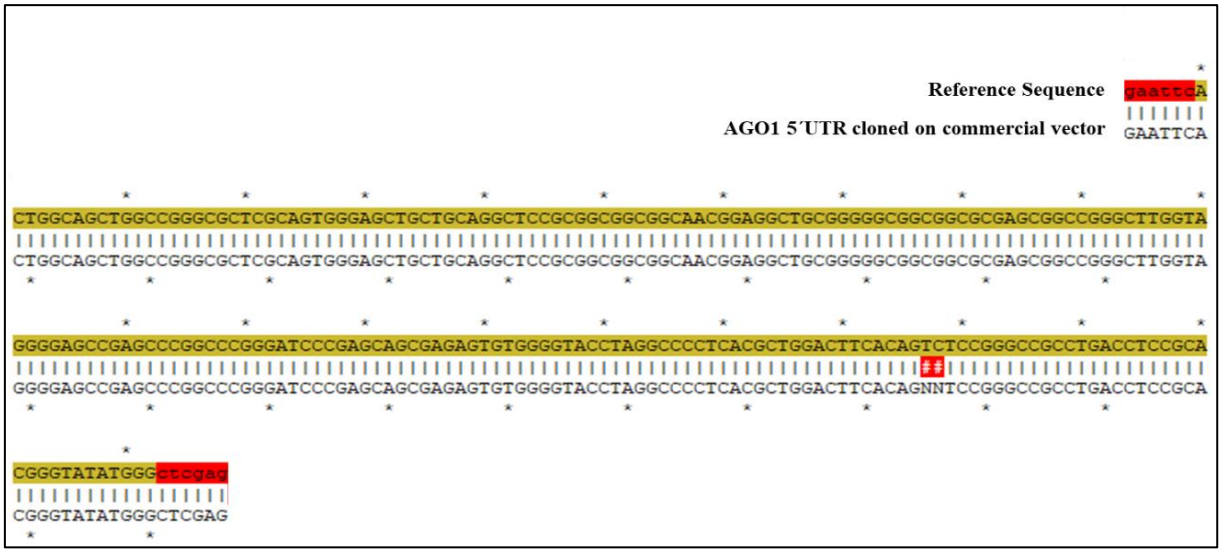

Figure 1. Verification of AGO1 5'UTR cloning into the commercial vector of known sequence (NZY-A PCR cloning kit, NZYTech). Comparison of the reference sequence of the AGO1 5'UTR with the sequencing obtained from the cloning of the insert AGO1 5'UTR into a commercial plasmid of known sequence, made by the program ape (http://en.bio-soft.net/plasmid/ApE.html). On top is described the reference sequence, in which yellow is the human AGO1 5'UTR (NM_012199) and in red the restriction sites of EcoRI (G^AATTC) and XhoI (C^TCGAG), respectively. Below is the sequencing of the commercial plasmid in which the AGO1 5'UTR insert was cloned. Cardinals represented in red mean that the program failed to decode which nucleotides are present. However, by the chromatogram analysis (data not shown) the most relevant peaks appear to correspond to the nucleotides of the reference sequence.

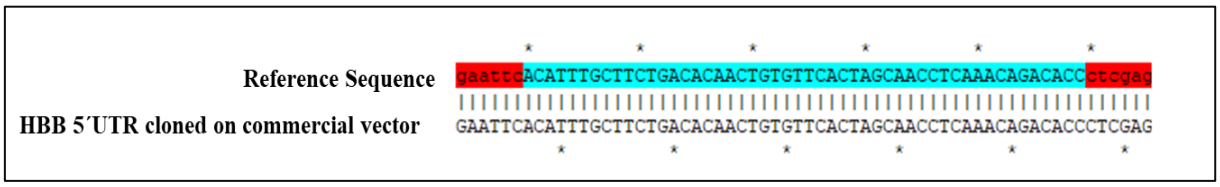

Figure 2. Verification of HBB 5'UTR cloning into the commercial vector of known sequence (NZY-A PCR cloning kit, NZYTech). Comparison of the reference sequence of the HBB 5'UTR with the sequencing obtained from the cloning of the insert HBB 5'UTR into a commercial plasmid of known sequence, made by the program ape (http://en.bio-soft.net/plasmid/ApE.html). On top is described the reference sequence, in which blue is the human HBB 5'UTR (NM_000518) and in red the restriction sites of EcoRI (G^AATTC) and XhoI (C^TCGAG), respectively. Below is the sequencing of the commercial plasmid in which the HBB 5'UTR insert was cloned.

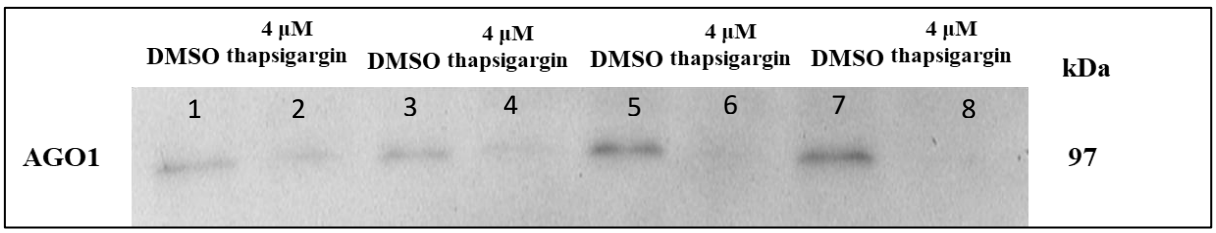

Figure 3. Effect of thapsigargin on Argonaute 1 protein synthesis in HCT116 cells. Western blot analysis of Argonaute 1 protein levels in untreated cells (DMSO, vehicle) – lanes 1, 3, 5 and 7 – and in cells treated with 4 µM of thapsigargin – lanes 2, 4, 6 and 8.