Characterization of Structural and Trans-Acting Elements of PLAC1 5’UTRs That Affect Translation of PLAC1

Citation

Permanent link
https://nrs.harvard.edu/URN-3:HUL.INSTREPOS:37364878

Terms of Use
This article was downloaded from Harvard University’s DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA

Share Your Story
The Harvard community has made this article openly available. Please share how this access benefits you. Submit a story.

Accessibility
Characterization of structural and Trans-acting Elements of PLAC1 5’UTRs that Affect Translation of PLAC1

Maple Gioia

A Thesis in the Field of Biotechnology for the Degree of Master of Liberal Arts in Extension Studies

Harvard University
March 2020
Abstract

The highly conserved developmental protein, Placenta Specific Protein 1 (PLAC1) is widely expressed in placental and fetal tissue during development. Its expression is restricted in normal tissue after birth, though it is reactivated and broadly expressed in many types of cancer.

Expression of PLAC1 is controlled by two promoters, producing alternatively spliced PLAC1 transcripts, each with a unique 5’untranslated region (UTR) and identical protein coding sequences. Patterns of differential transcript expression resulting from alternative promoter usage are observed in both fetal development and cancer. To date, little is known about the post-transcriptional regulation of PLAC1 expression or the significance that 5’UTRs have on the translation of PLAC1. Using a GFP reporter system, the translational efficiency of each 5’UTR was measured. Overall, the 5’UTR sequences of promoter 2 transcripts had slightly lower efficiency than those from promoter 1 and mutation of the upstream open reading frames (uORFs) were minimally effective in restoring %GFP activity. Binding site motifs in PLAC1 were also analyzed using the RPBMapper and ATtRACT database. Binding sites for RBFOX proteins and ELAVL2, important alternative splicing proteins during fetal development, were identified 14 nucleotides apart in exon 4. Intronic binding sites for both of these proteins were also identified. The presence of binding sites for RBFOX and ELAVL2 proteins suggests that PLAC1 undergoes extensive alternative splicing during fetal development, potentially encoding additional isoforms that have not yet been identified.
Dedication

This work is dedicated to my father and mother, who instilled in me the importance of education and supported me as much as they could through these endeavors.
I would like to acknowledge all the members of my team who have helped me out with this project along the way. These special people include Jenn Detellis, Brendan Lang, and Niki Jayachandran. Also, a very special thank you to Purna Venkataraman, who stayed late to help me out multiple times throughout this process. I couldn’t have done this without you. Also, to my Thesis Advisor, Geoff Parsons, I am forever indebted to your help on this project and am so thankful to have you as a manager. Finally, to David, who has always been supportive of me on this journey, even though it was not always an easy road. Thank you.
Table of Contents

Dedication ........................................................................................................................................ iv
Acknowledgments ................................................................................................................................. v
List of Tables ......................................................................................................................................... viii
List of Figures ......................................................................................................................................... ix

Chapter I. Introduction ......................................................................................................................... 1
  Background: PLAC1 Expression in Cancer ......................................................................................... 4
  Background: Post-Transcriptional and Translational Regulation ....................................................... 8
  Background: Upstream Open Reading Frames ................................................................................. 11
  Background: RNA Binding Proteins ................................................................................................. 14

Chapter II Materials and Methods ....................................................................................................... 15
  Plasmid Design and Cloning .............................................................................................................. 15
     Gel purification and Restriction Cloning ......................................................................................... 17
     Transformation ............................................................................................................................ 19
     Sequence Verification of Ligated Plasmids .................................................................................... 19
  Transient Transfection of HEK293T cells ......................................................................................... 19
     Mutation of uORF Start Codon .................................................................................................... 21
     Fluorescence activated cell sorting ............................................................................................ 21
     Quantitatively Measuring GFP with qPCR .................................................................................... 22
     Identification of RNA Binding Motifs within PLAC1 5’ UTR Sequences ................................... 23

Chapter III Results ............................................................................................................................. 27
Effect of PLAC1 5’UTR sequences on GFP in HEK293T cells ...................28
Identification of RBP Binding Motifs in PLAC1 ..................................34
Chapter IV Discussion ..............................................................................38
RNA Binding Proteins in PLAC1 5’UTRs ..............................................39
5’UTR Effects on Translation ...............................................................41
Cancer expression and pathogenicity ....................................................43
Chapter V. Research Limitations ...........................................................45
References ............................................................................................49
List of Tables

Table 1. Features of PLAC1 5’UTR sequences that may affect translation ..................10
Table 2. Reagent Input for Restriction Cloning ..........................................................17
Table 3. High Significance Binding Motifs observed for RBFOX1 .............................37
List of Figures

Figure 1. Overview of PLAC1 Gene and Protein Structure ...........................................2
Figure 2. PLAC1 5’UTR Sequences and Lengths. ..........................................................13
Figure 3. Overview of GFP Reporter Plasmid Construction ...........................................16
Figure 4. Diagnostic Gel of GFP Reporter Plasmids containing PLAC1 .......................18
Figure 5. RBPMapper Position Based Scoring Method .....................................................24
Figure 6. Equation used by AtTRACT ........................................................................25
Figure 7. Translational Efficiency of PLAC1 5’UTR Variants .........................................29
Figure 8. Relative Expression levels of RNA and Protein ..............................................31
Figure 9. Assessing uORF Response to heat stress. .........................................................33
Figure 10. RBFOX1/2 silencer motif in Exon 4 ...............................................................35
Figure 11. RBFOX1 conserved binding motif within PLAC1 introns ..............................36
Figure 12. AUG Codons Present in PLAC1 ..................................................................47
Figure 13. Differences in RNA Folding Structures .........................................................48
Chapter I.

Introduction

The existence of Placenta Specific Protein (PLAC1), a 200kb X-linked gene, was first described in a 2000 publication (Cocchia et. al., 2000). Here PLAC1 was described as a 212 amino acid protein containing a zona pellucida (ZP3)-like domain (Figure 1), whose expression was restricted to the placenta. Further research found it to be expressed in specific cell types of the placenta, most notably those of trophoblastic origin (Cocchia M., et al., 2000). Its function in the development of the fetal-maternal interface occurs through trophoblastic invasion (Fant, M., Farina., A, 2010) of the maternal wall, enabling fetal implantation to occur. More recently, PLAC1 knockout studies in mice revealed that PLAC1 is also critical for fetal brain formation and development. (Kong et. al, 2013). In these studies, a proportion of the PLAC1 null mice developed lethal hydrocephalus after birth. Additional studies exploring the expression of PLAC1 in embryonic development observed that its expression is not restricted to the placenta, but is in fact widely expressed in many developing fetal tissues including the heart, lungs, kidney, and brain (Jackman, S.M, Kong, X.& Fant, M.E, 2012) and is critical for normal neuron and brain structure development (Bourgeois J.B., 2015).
Figure 1. Overview of PLAC1 Gene and Protein Structure.

Overall structure of PLAC1 gene and protein. PLAC1 is expression is driven by two promoters, resulting in four confirmed alternatively spliced sequences. The promoters are differentially activated through either estrogen (P2) or p53 (P1). PLAC1 contains a ZP3 domain from amino acid 40 to amino acid 119 and a signal peptide from amino acid 1 to 23. (Chen, Moradin, Schlessinger, D. Nagaraja, 2011 and Devor, 2014)
While its specific function in these tissues remains unknown, its expression throughout fetal development suggests it plays a critical and tissue-specific role in the development of many fetal tissues. Its significance in fetal development is also evidenced by the high level of evolutionary conservation across placental mammals (Devor, 2014)
Background: PLAC1 Expression in Cancer

PLAC1 is found to be widely expressed in many cancers (Koslowski, et. al, 2007, Silva et al., 2007, Dong et al., 2008, Yin et al., 2017). In a 2008 study of patients with colorectal cancer, researchers found that PLAC1 was localized both in the cytoplasm and at the cell surface (Liu et al., 2008). This observation was discrepant with previously published results in the placenta, which demonstrated on multiple occasions that its localization is restricted to the cell membranes surface of trophoblasts (Fant et al., 2002). Additional studies of placental localization of PLAC1 have also found its expression to be restricted to the cell membranes in the placental compartment (Vandré, Ackerman, Tewari, Kniss & Robinson, 2012). Differences in localization between placental and the cancerous expression may be the result of differences in cancer specific alternative splicing, which may cleave off the PLAC1 signal peptide and result in alternative initiation at a downstream AUG start site. This would yield a PLAC1 protein without a cell membrane localization signal and may be the reason for its localization to the cytoplasm. This alternate localization may result in functional differences in cancer that are divergent from that of the placenta. Further research must be performed to understand the differences between the function in PLAC1 versus in placental and fetal development.

PLAC1 is controlled by a dual promoter system. The promoter predominately responsible for driving PLAC1 expression in cancer was first described by Chen et. al. in 2011 (Chen et. al, 2011). Here, it was demonstrated that PLAC1 transcription is not
solely controlled by the estrogen-dependent P2 promoter, but is also by a second promoter, located 105kb upstream of P2, which is regulated by p53, in addition to other transcription factors (Figure 1) This publication found that PLAC1 transcripts in the placenta are chiefly derived from the proximal P2 promoter, while transcript expression of PLAC1 in cancer is more often associated with P1 activation. Moreover, additional research has shown that PLAC1 transcripts in certain breast cancer lines often express PLAC1 transcripts derived from P2 (Wagner et. al.2014, Koslowski et al. 2007, Koslowski et al. 2009). Alternative promoter usage results in expression of PLAC1 transcripts that contain identical protein coding sequences, but alternatively spliced 5’untranslated regions (5’UTRs) (Devor et al., 2016). Furthermore, exon inclusion of each PLAC1 5’UTR sequence is dependent on promoter usage, as exon 4 is only found in P2 derived transcripts (Chen et al., 2013). While placental proteins are frequently derived from P2, those PLAC1 transcripts expressed in cancer are like those expressed in the fetal brain, heart and kidney, as they are derived from P1 (Figure 1). To date, little is understood about the function and regulation of PLAC1 expression in fetal tissues and whether the well-studied invasive function of PLAC1 in placenta (Koslowski et al. 2007, Koslowski et al. 2009) is also observed in cancers that express PLAC1. Limited functional research of PLAC1 in cancer does not conclusively support a widespread role for PLAC1 in migration and invasion of cancer. Alterations of p53 function by mutation or transformation with T antigen, were demonstrated to be the primary driver of PLAC1 expression from the P1 promoter (Chen et al, 2013). Recent research performed on tumors from Ovarian cancer patients demonstrated that a mutated p53 status was associated with elevated levels of PLAC1 expression and that these patients had a poorer
prognosis (Devor et al., 2017). In a different study of pancreatic tumors, PLAC1 expression was also found to have a significant correlation with poorer survival (Yin et al., 2017) though p53 status of these tumors was not assessed and poorer prognosis may have been the result of p53 mutation. As PLAC1 expression is directly correlated with p53 status, it is difficult to conclude whether the poorer outcomes observed were the result of p53 mutation or PLAC1 expression.

Unlike other cancers, where PLAC1 expression is reactivated by p53 mutation, in some breast cancer lines such as MCF7, it is the proximal P2 promoter that is activated and drives PLAC1 expression (Wagner et al., 2014). P2 activation is described in several other breast cancer cell lines (Koslowski et al., 2007), where it has been demonstrated that P2 is activated through a novel Estrogen receptor-dependent mechanism, which works in coordination with the transcription factors SP1 and nuclear co-factor C/EBPb (Figure 1). Further work demonstrated that NCOA3 was also recruited to the P2 promoter in MCF7 cells (Wagner et al., 2014) and functioned to increase its transcription and translation.

Subsequently, additional research by Chen et al., 2011 demonstrated that the nuclear receptor RXRa activates P1 in conjunction with LXRb (Figure 1) (Chen et al., 2011). The reactivation of PLAC1 expression in cancer has been the subject of several research studies. To understand the function of PLAC1 in cancer, efforts to investigate the functional differences between PLAC1 transcripts originating from either P1 or P2 should be prioritized. Efforts should also be made to understand how it is regulated in fetal development, which may provide insight into its role when expressed in cancer. Overall,
the complexity of PLAC1 reactivation may make it a challenging target to pursue for targeted therapies. To this end, gaining a better understanding of its tissue specific functions would provide further insight into its behavior in cancer and would provide a much more fundamental understanding of how it might be targeted
Background: Post-Transcriptional and Translational Regulation

As the 5’UTR of a PLAC1 transcript is the only feature distinguishing it from other confirmed PLAC1 transcripts, exploring how 5’UTRs affect PLAC1 expression and function may elucidate promoter specific and transcript-specific mechanisms of regulation. Determining their role in PLAC1 expression is also critical to gain further insight into the contribution of PLAC1 to fetal development and disease etiologies.

It is known that structural elements within the 5’UTR can significantly impact translation and function of a downstream ORF, through several mechanisms previously characterized (Sonenberg N. et al., 2009). Differences in 5’UTR GC content, hairpin structure, overall length, and presence of upstream open reading frames are known to affect translation (Kos, Denger, Reid & Gannon, 2002). These 5’UTR characteristics can affect translational efficiencies through mechanisms such as leaky ribosome translation, exon skipping, and translation at alternative initiation sites (Jackson, Hellen, & Pestova, 2010).

The 5’UTR sequences of PLAC1 transcripts are all unique, therefore it is possible that they can differentially affect translational efficiency. Sequence analysis using Geneious software of PLAC1 5’UTR sequences revealed that %GC content between PLAC1 5’UTRs is also different. Transcript variants 1-3 have 45-50% GC content, transcript variant 4 has only 33%. %GC content can control gene expression by affecting the stability of the mRNA molecule (Jackson, Hellen, & Pestova., 2010). Hairpin structure can also affect translation of a transcript by interfering with ribosomal entry to the AUG.
start site (Jackson, Hellen, & Pestova, 2010), therefore, a hairpin upstream of an AUG start site may affect initiation of translation at the primary ORF. Using Geneious software the RNA folding structure of each PLAC1 5’UTR was assessed (Geneious Software 10.2.3). Transcript variant 2 is the only transcript that appeared to have a hairpin upstream of the AUG start site. Finally, the length of each 5’UTR may play a role in its translation, as longer 5’UTRs have been associated with decreased translation (Araujo, P.R, 2012). To assess the effect that each PLAC1 5’UTR has on translation, a GFP reporter system was developed for all four PLAC1 5’UTR sequences. The effect that each 5’UTR has on translation was measured by changes in %GFP reporter expression, as measured by FACS
Table 1. Features of PLAC1 5’UTR sequences that may affect translation

<table>
<thead>
<tr>
<th>Name</th>
<th>Presence of uORF</th>
<th>%GC Content of uORF</th>
<th>Distance of uORF from primary ORF (base pairs)</th>
<th>5’UTR Length (base pairs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PLAC1 Transcript Variant 1</td>
<td>36 base pair uORF</td>
<td>45-50%</td>
<td>249</td>
<td>308</td>
</tr>
<tr>
<td>PLAC1 Transcript Variant 2</td>
<td>No upstream uORF for AUGs.</td>
<td>45-50%</td>
<td>No uORF</td>
<td>326</td>
</tr>
<tr>
<td>PLAC1 Transcript Variant 3</td>
<td>uORF – 51 base pair</td>
<td>45-50%</td>
<td>259</td>
<td>354</td>
</tr>
<tr>
<td>PLAC1 Transcript Variant 4</td>
<td>36bp uORF</td>
<td>33%</td>
<td>170</td>
<td>236</td>
</tr>
</tbody>
</table>

Certain elements of 5’UTR sequences are known to affect the translation of a downstream primary transcript. In the table above, the presence of some of these elements is outlined.
Background: Upstream Open Reading Frames

Normal translation of an mRNA is initiated through a mechanism known as ribosomal “scanning” where a ribosome containing a methionine searches along the mRNA sequence for a corresponding AUG codon (Araujo et al., 2012). Once the ribosome recognizes the AUG start codon in context of a kozak sequence (Kozak, 1989) and transfers its methionine amino acid, strand elongation begins (Babendure, 2006). In conditions where a uORF exists, ribosome scanning can initiate at the incorrect uORF and ablate translation (Morris & Geballe, 2000).

During the process of initiation and elongation, 5’UTR sequences have evolved mechanisms that influence the efficiency of translation. This is accomplished with the presence of a decoy AUG start site upstream of the primary ORF containing the CDS (Sonenberg N. et al., 2009). These decoy sites can result in initiation upstream of the true AUG start site, resulting in reduced translational efficiency. uORFs have been found in approximately 50% of all mRNA sequences and are known to affect translation of the downstream primary protein coding sequence (Eisen & Scannell, 2009). In normal conditions, uORFs have been shown to simply block translation of the downstream transcript (Eisen et al., 2009), while in conditions of cellular stress, leaky translational initiation has been shown to increase the translation of the uORF resulting in short peptides (Young & Wek, 2016). In genes involved in fetal neuron development, uORFs are translated, inhibiting translation of the primary ORF (Rodriguez, Chun, Mills & Tod, 2018).
In one study, upstream unopen reading frames (uORFs) in 5’UTR sequences were found to have a dramatically negative impact on translation, in some cases reducing protein expression of a given transcript by up to 80% (Eisen & Scannell, 2009).

Sequence analysis of the 5’UTRs of PLAC1 revealed the presence of upstream reading frames (uORFs) in three of the four transcripts. The uORF in transcript variant 3 is the largest, at 51 base pairs long, while those in transcript variant 1 and 4 are slightly smaller (both are 36 base pairs). As the uORFs in transcript variant 1 and 4 are different distances from the primary ORF, they may have a different effect on translation, as previous research has shown may be the case (Koš, Denger, Reid & Gannon, 2002). As PLAC1 5’UTR sequences have uORFs, it is possible that they are translated in response to stress or may have an impact on stress response (Harding et. al, 2003). To date, the effect which uORFs have on translation of PLAC1 transcripts has not been explored, both in normal or stress conditions.

To measure the effect of the uORFs present in the 5’UTRs of PLAC1, plasmid GFP reporters with mutated uORF sites were designed for each of the transcript variants of PLAC1 containing a uORF. This was accomplished by mutating the AUG start codon (AUG>UUG). %GFP of cells transfected with wild-type 5’UTR GFP plasmids were then compared to cells transfected with 5’UTR GFP plasmids containing mutated uORFs. Differences in relative expression levels of %GFP expression will determine whether uORFs of PLAC1 may function in translational efficiency.
This figure shows the sequence for each of the PLAC1 5'UTR sequences. The sequences here were inserted into plasmids, downstream of a promoter and upstream of a GFP reporter gene. These constructs will enable the visualization and quantification of any differences in translation caused by the 5'UTR.
RNA binding proteins (RBPs) have the capacity to bind to specific motifs on mRNA sequences, eliciting regulatory functions at the posttranscriptional level. These proteins have a diverse array of responses upon binding a recognized RNA motif, including, alternative splicing of transcripts in a tissue specific manner (Minovitsky, 2005). RNA binding proteins recognize highly conserved motifs within exons or flanking intronic regions to elicit their effect on expression. (Clayton, 2013). The RBPs often exert their effect, either alone or with protein co-factors, by binding to conserved motifs found in the 5’UTR and 3’UTR regions of an mRNA transcript, and within the intronic regions flanking an exon. (Kwon et al., 2018). Using mapping algorithms, an mRNA or intron of interest can be analyzed for the presence of these RBP binding sites to gain insight into how the gene is regulated both globally and, in a tissue specific manner. The expression of these proteins may be regulated through the course of development or restricted to specific tissues, enabling tissue and developmental specific regulation of mRNA (Minovitsky, 2005, Conboy 2016). As the 5’UTR sequences are the only distinguishing difference between PLAC1 transcripts, it is possible that their unique 5’UTR sequences may allow for differential regulation by RNA binding proteins. Identifying RBPs recognition motifs in both the 5’UTRs and their flanking intronic regions of PLAC1 would provide valuable insight into the regulation and function of PLAC1 expression. Furthermore, identifying motifs that regulate PLAC1 may also provide more insight into its putative function in fetal development and cancer.
Chapter II

Materials and Methods

To determine whether PLAC1 5’UTR sequences alter translation of their transcripts, GFP reporter system was developed for each confirmed 5’UTR sequence of PLAC1. In Geneious, MluI restriction sites were added to both the 5’ and 3’ ends of each PLAC1 5’UTR sequence. To measure the effect of the uORF, three additional plasmids were made for each variant containing a uORF. In these three plasmids the AUG start codon of the uORF sequence was mutated (A>T). These seven sequences were then synthesized as linear DNA at Integrated DNA Technologies as gblocks. PLAC1 5’UTR gblocks were cloned into an expression plasmid using an MLUI restriction site located between the promoter and a GFP reporter gene. This location, outlined in Figure 3, is located downstream of a strong promoter.

Plasmid Design and Cloning

A total of 8 GFP reporter plasmids were created to assess the function of PLAC1 5’UTRs. The first four plasmids contain one of the four confirmed wild-type 5’UTR sequences of PLAC1 (Figure 2). For the next three, the uORFs of Transcript Variant 1, Transcript Variant 3, and Transcript Variant 4, were deleted. This was achieved by mutating the first base of the AUG start codon (A>T). The eighth plasmid does not
contain a 5’UTR sequence between the promoter and GFP gene, serving as the control to
which expression is normalized. All confirmed wild-type sequences of PLAC1 were
obtained from the NCBI nucleotide database.

Figure 3. Overview of GFP Reporter Plasmid Construction

*GFP Reporter Construction includes a strong promoter (MNDU3) followed by the
insertion of a PLAC1 5’UTR sequence (transcript variants 1-4). These two elements
drive the translation of a GFP located downstream of the PLAC1 5’UTR sequence.
MLUI and PSTI restriction sites are shown on the plasmid above.*
Gel purification and Restriction Cloning

Digests were run on a .5% gel (100 Volts, 40 minutes) with 13.5uL digest added in each lane. Plasmid backbone and insert were excised from the gel and the Zymo Gel Clean up kit was used for extraction. Plasmids were eluted in 20uL of buffer, while inserts were eluted in 10uL. Furthermore, in an effort to determine whether inserts for each plasmid were present, a gel was run for constructed plasmids using mini prep DNA (Figure 4). Mini prep DNA was isolated and MLU I was used as the restriction enzyme to cut inserted DNA. Bands on the gel were consistent with the size expected for each insert (~250-400 base pairs). Those that didn’t show the correct band size were re-constructed (Figure 4, lane5). All plasmids were also sequence verified for proper insert and orientation.

Table 2. Reagent Input for Restriction Cloning and diagnostic gel.

<table>
<thead>
<tr>
<th>Reagent</th>
<th>Backbone vector (1-2 ug)</th>
<th>Insert (1-2 ug): gBlock</th>
</tr>
</thead>
<tbody>
<tr>
<td>DNA</td>
<td>1 uL</td>
<td>1 uL</td>
</tr>
<tr>
<td>10 x Buffer</td>
<td>4 uL</td>
<td>4 uL</td>
</tr>
<tr>
<td>Restriction Enzyme (MLUI)</td>
<td>2 uL</td>
<td>1 uL</td>
</tr>
<tr>
<td>FAST AP</td>
<td>1 uL</td>
<td>--</td>
</tr>
<tr>
<td>Water</td>
<td>32 uL</td>
<td>32 uL</td>
</tr>
</tbody>
</table>

*Reagents used for the restriction cloning of backbone to each PLAC1 5’UTR insert. MLUI was used to clone each insert into the vector.*
A diagnostic gel with MLU1 was run on mini prep DNA after ligation to confirm presence of gblock insert. Plasmids not showing a band were re-cloned. All final mini-prep plasmids were sequence verified. (Lane 1 – 1kb ladder: Lane 2 – Transcript variant1 5’UTR plasmid, expected insert size, 308: Lane 3 – Transcript variant 2 5’UTR plasmid, expected insert size 326: Lane 4 – Transcript variant 3 5’UTR plasmid, expected insert size 354: Lane 5 – Transcript variant 4 5’UTR plasmid, expected insert size 236: Lane 6 – Transcript variant 4 mutated uORF 5’UTR plasmid, expected insert size, 236: Lane 7 – Transcript variant 3 mutated uORF 5’UTR plasmid, expected insert size 354: Lane 8 – Transcript variant 1 5’UTR, expected insert size 308). Lack of band in Lane 5 was observed. This plasmid was sent out for sequence verification and found to have an inverted insert. The plasmid was re-cloned and sent out for sequence verification. The re-cloned plasmid had the correct insert and orientation. All 7 plasmids used in the experiment matched reference plasmids.
Transformation

Competent E.coli cells (Z-Competent™ E. Coli Cells - Strain Zymo 5α, T3009) were thawed on wet ice and 5uL of ligation reaction was added to each. Transformed cells were plated on Ampicillin treated plates from Teknova (L1004) and placed at 37C overnight. Colonies were picked after 16 hours at 37C. Each colony was placed into 2mL liquid culture containing ampicillin and placed in a shaking incubator overnight.

Sequence Verification of Ligated Plasmids

A diagnostic gel was run on DNA extracted from liquid cultures to verify the correct sequence and orientation of each plasmid insert. A restriction digest using MLUI was used for verification of the insert, while PSTI was used to verify both presence of the insert and sequence orientation (data not shown). DNA sequences were verified by Sanger sequencing. Only those with fully verified sequences were used to make maxi prep

Transient Transfection of HEK293T cells

Three days prior to transfection, cells were seeded at 1.0E6 cells per 10cm plate, in a total of 10mL of D10 (DMEM/10% FBS). On the day of transfection, confluency was 80%. Transfection master mix was prepared using .1ug of plasmid. Plasmid DNA was normalized to the appropriate concentration with NaCl/HEPES transfection buffer. 250uL of pre-made transfection master mix containing NaCl/HEPES transfection buffer
plus PEI Pro was added to the plasmid DNA and allowed to incubate for five minutes before cells were transfected.

500uL of transfection mixture was added to each 10cm flask in a drop-wise manner and placed back in the incubator immediately. Transfected cells were incubated for 24 hours before harvest. On the day of harvest, cells were harvested by aspirating out D10 media and adding 1mL of TrypLE to lift the cells off the bottom of the cell culture flasks. 5mL of D10 was added to each culture dish. Cells were then spun down at 300g for five minutes and resuspended in 1000uL of PBS.

For the heat shock experiment, those cells that were subjected to heat shock were incubated at the same conditions described above for 21 hours, and then incubated at 42C for 3 hours before harvest. Heat shock was chosen because stresses have been known to cause leaky ribosomal translation, which can occur quickly after subjected to heat shock, causing the uORF to become translated (Babendure, 2006). The three hours at 42C was chosen as a point at which the cells would most likely show signs of heat stress, but would still survive.
Mutation of uORF Start Codon

HEK293T cells were seeded at 1e6 cells/mL in DMEM and placed at 37°C and were transfected 72 hours after seeding. Cells were transfected with plasmid DNA using the PEI Pro transfection reagent. Cells were harvested 24 hours after transfection for qRT-PCR and FACS analysis and stained for 10 minutes with a viability dye. Flow Cytometry was performed with the MACS QUANT Studio. %GFP+ cell counts taken from the singlet GFP positive gated populations

Fluorescence activated cell sorting

Harvested samples were placed in 1000mL of PBS were split equally into two 1.5mL Eppendorf tubes. 500uL of sample was stained with a viability dye, and the rest of the cells were left unstained. Both stained and unstained cells were run on the flow cytometer. Mock HEK293T sample was used to set the gating parameters on the MAQS Quant flow cytometer. 200uL of sample was used to capture 10,000 events. For assessing GFP expression, %GFP positive cells FITC gates were used and populations were gated from singlet cell populations
GFP expression from both the mutated uORF-GFP plasmid and the wild-type-GFP plasmid were normalized against the control GFP vector to obtain percent expression levels for each plasmid relative to the control vector. mRNA was quantitated using a quantitative PCR (qPCR) method. To confirm that any differences observed in GFP expression were due to changes in translation and not because of differences in RNA levels. Primers were developed for GFP and ACTB. ACTB was used as the endogenous normalizer for gene expression data. For each sample, data were normalized by subtracting the Ct values of ACTB from those of GFP Ct (GFP Ct – ACTB Ct) in the same sample.

The ddCp calculation was obtained by subtracting the test sample dCp from that of the control plasmid, to get Log2 Fold values. Percent relative quantification was determined with the following calculation: $2^{\Delta \Delta Ct}$. Relative expression levels for %GFP were calculated by dividing the %GFP positive cells by the %GFP positive cells of the control plasmid, then multiplying by 100 to get the relative amount compared to control. These two relative calculations were then graphed with each other.
Identification of RNA Binding Motifs within PLAC1 5’ UTR Sequences

Four PLAC1 5’UTR sequences and all PLAC1 intronic regions were analyzed for the presence of protein binding motifs using two computational methods which scan RNA sequences for conserved RNA binding site motifs with RBPMapper (Figure 5) and ATtTract (Figure 6) (Paz, Kosti, Ares, Cline, Mandel-Gutfreund, 2014) and (Giudice, Sanchez-Cabo, Torroja & Lara-Pezzi, 2016).

The RBPMapper uses a weighted-rank function and position specific scoring method to analyze an input sequence for the presence of a pre-defined list of 114 human RNA protein binding motifs (Figure 5).
The calculations above are those used by the RBPMapper software program to determine whether an input sequence contains any confirmed binding motifs hits. A&B) these are used to initially calculate a score to determine if a given sequence is a match. C) This calculation assesses the similarity of a match score calculated in A&B with the match score of the conserved pre-defined motif database. D) The motifs are calculated as a weighted rank, comparing the hit with the similarity of the database sequence. This provides the z-score and p.value output. (Paz, Kosti, Ares, Cline, Mandel-Gutfreund, 2014)
The ATtRACT database uses experimentally confirmed data to identify binding motifs present in a sequence input into the sequence scanner. A) The algorithm uses a position-specific probability matrix (PPM) derived from the experimentally determined sequences. B) Represents the calculation used to determine the sequence score for each scanned hit. (Giudice, Sanchez-Cabo, Torroja & Lara-Pezzi, 2016)

For all sequences put into the RBPMapper, the highest stringency level was used for identifying statistically significant hits (P < .001) and the conservation filter was also used. While the algorithm has a built-in function for reducing the number of false positive results, an mRNA with validated and experimentally determined RNA binding motifs was analyzed using RBPmap (Paz et al., 2014), to determine whether this method would positively identify sequences with validated affinity with RBPs. To accomplish this, a study measuring the binding affinity of Quaking gene 1 (QK1) to motifs within the 3’UTR of Myelin Binding Protein (MBP) was used as a test case (Rodriguez, Chun, Mills & Tod 2018) In this study, electrophoretic mobility shift assays and florescence polarization assays the binding affinity of QK1 to each of the 5 sites was validated, with
one motif (CACUAA) showing the highest binding affinity (Rodriguez, Chun, Mills & Tod 2018). To test whether the RPBmap would identify these motifs, the 3’UTR of MBP was analyzed with RBmap, first using the lowest stringency filter, and finally using the highest. With the lowest stringency filter, 20 putative QK1 binding motifs were identified within the MBP 3’UTR, including all five of the motifs identified in the paper. Using the highest stringency filter, only one putative binding site was identified. This single binding site matched the binding motif experimentally determined to have the highest binding affinity with QK1.

To provide further confirmation of the RNA Binding motifs identified by RPBMapper, the AtTract sequence analysis tool was also used to assess the binding motifs present on PLAC1 5’UTR sequences (Giudice, Sanchez-Cabo, Torroja & Lara-Pezzi, 2016)
Chapter III

Results

The widespread expression throughout fetal development indicates a critical role for this gene in development. The primary goal of this study was to explore the regulation of PLAC1 at the post-transcriptional level, to gain a better understanding of its function, and understand whether dysregulation may contribute to disease etiology. Given the difference in PLAC1 transcripts expressed in placenta or cancer, this may be indicative of a difference in translation or function of PLAC1 in cancer. To explore whether PLAC1 transcripts have different translational efficiencies due to the 5’UTR sequences present in each PLAC1 transcript, a GFP reporter system was developed to assess this question. If translation is different for each given PLAC1 transcript, then those transcripts with lower translational efficiency may not contribute to disease etiology.

Using a GFP reporter system the 5’UTR sequences, were assessed for their potential to affect translational efficiency, a major mechanism of 5’UTR regulation. These GFP plasmids were assessed in both normal and heat stress conditions. Subsequently, the 5’UTR sequences of each transcript and their corresponding intronic regions were analyzed for the presence of highly conserved RNA binding site motifs using RBPMapper, and ATtTRACT databases. Identifying highly conserved protein binding subsequences present within the exons and their flanking intronic regions would
provide a more comprehensive understanding of the global regulation of PLAC1 expression. Differences in the function of PLAC1 transcripts or in binding protein composition present within a cancer may alter its function in different cancer subtype.

Effect of PLAC1 5’UTR sequences on GFP in HEK293T cells

To assess the proposed mechanism of regulated PLAC1 translation, a GFP reporter plasmid was created for each of the four confirmed PLAC1 5’UTR sequences. Upstream open reading frames are known to regulate translation of downstream mRNA. uORFs of various lengths are found in three of the four PLAC1 5’UTR sequences. To assess whether the uORF sequences could function to regulate PLAC1 translation, three additional GFP reporter plasmids were made, one for each of the uORF containing PLAC1 5’UTR sequences. The AUG start codon of each uORF was altered by mutation of the AUG start codon. The single A>T base pair change should effectively eliminate the presence of the uORF function.

Translational control mechanisms driven by the presence of a uORF sequence were quantitatively assessed by comparing each of the seven plasmids containing a 5’UTR to a control plasmid containing no 5’UTR sequence. (Figure 7). To determine whether any differences observed in translation were merely associated with differences in RNA expression, qPCR was performed and compared side-by-side for each vector (Figure 8). Finally, as it is known that uORFs may change the translational efficiency of downstream primary ORFs during heat stress (Araujo et. al, 2012), plasmids containing the wild-type vector (does not contain any PLAC1 5’UTR), transcript variant 3 5’UTR
vectors, and transcript variant 3 mutated uORF 5’UTR plasmids, were transfected into HEK293T cells, incubated for 21 hours at 37°C and then subjected to heat shock (42°C) for 3 hours. GFP expression from these plasmids subjected to heat shock was compared to the same plasmids not subjected to heat shock (Figure 9).

![Translational Efficiency of PLAC1 5’UTR Sequences](image)

**Figure 7.** Translational Efficiency of PLAC1 5’UTR Variants.

The HEK293T cells transfected with plasmid containing the PLAC1 5’UTR sequence from transcript variant three had the highest percentage of %GFP positive cells of all plasmids tested. Transcript variant 4 5’UTR had the lowest expression of %GFP and expression was increased with mutation of the uORF.
When assessing translational efficiency alone (Figure 7) some surprising results were observed. First, HEK293T cells transfected with plasmid containing the PLAC1 5’UTR from transcript variant three had the highest percentage of %GFP positive cells of all plasmids tested (Figure 9). This result was largely unexpected, as uORFs, which are well-known inhibitors of translation, are located within this sequence and the 5’UTR of transcript variant 3 has the largest uORF. The transcript variant 3 5’UTR contains a 51 base pair uORF sequence. %GFP expression in cells transfected with 5’UTR sequence of PLAC1 transcript variant 2, the only PLAC1 transcript lacking a 5’UTR uORF sequence was similar to the %GFP in cells transfected with PLAC1 5’UTR transcript variant 3 plasmids. Therefore, these results indicate that differences in translational efficiency of PLAC1 do not appear to be influenced by the presence of uORF sequences. To further explore the reason for differences observed in the translational efficiencies in Figure 7, RNA expression analysis was performed for all samples assessed in Figure 7. Differences in RNA expression between each plasmid appear to explain the differences in translation between the plasmids (Figure 8), as mRNA and protein expression appear to match identically for each vector. This further promotes the finding that uORFs or 5’UTRs in general don’t contribute to differences in translation in PLAC1 under the conditions tested.
Figure 8. Relative Expression levels of RNA and Protein.

The percent relative expression for RNA and Protein was obtained to determine whether the expression levels of RNA were contributing to the differences in %GFP expression observed. The RNA expression in PLAC1 5’UTR transcript variant 3 is 350% higher than that of the control. The additional RNA measurements generally reflect the %GFP
On the other hand, when PLAC1 5’UTR transcript variant 3 wild-type and mutated uORF plasmids were subjected to heat stress significant differences were observed in the wild-type plasmids, and those with a transcript variant 3 mutated uORF (Figure 9), but differences were not observed between the heat shock and normal condition of those plasmids containing the unmutated uORF of PLAC1 transcript variant 3 (Figure 9). Overall, in cells transduced with PLAC1 transcript variant three 5’UTRs where the A in the AUG had been mutated, a significant 10% decrease in GFP expression was observed in those cells subjected to heat stress versus those that were also transduced with PLAC1 transcript variant three 5’UTR with a wild-type uORF sequence (Figure 9). The reasons for the differences observed between the unmutated and mutated uORF sequences must be further explored, and may indicate a role for the uORF in heat stress conditions.
Figure 9. Assessing uORF Response to heat stress.

HEK293T cells transfected with control vector (313), wild-type PLAC1 5’UTR transcript variant 3 (548), or PLAC1 5’UTR transcript variant 3 mut uORF (564) were subjected to heat shock conditions for 3 hours and then compared to cells in normal conditions. %GFP in control and wild-type PLAC1 5’UTR transcript variant 3 cells was reduced, while no change was observed in the cells transfected with wild-type PLAC15’UTR TV3.
Identification of RBP Binding Motifs in PLAC1

RNA Binding Proteins (RBPs) are known to have diverse regulatory functions at the post-transcriptional level. Their function is exerted by binding to evolutionarily conserved motifs found in mRNA or intronic sequences. The aim of this study is to identify post-transcriptional regulatory elements of PLAC1 which may affect its expression. Analysis of PLAC1 exons and their flanking intronic regions by RBP Mapper, identified RBFOX1 as a binding protein with a very high binding probability in both the AtTRACT database and with RBPMapper.

In the ATrTract database, an experimentally determined binding motif which RBFOX1 and RBFOX2 bind to (Zhou et. al, 2016) is observed in exon 4, (Figure 10) where this binding sequence UGACUG functioned as an exonic silencer, promoting exon 4 exclusion outside of neural tissue (Zhou, Baraniak, & Lou, 2006). Interestingly, this exon is only found in transcripts derived from promoter 2, which are predominantly expressed in placental tissue (Figure 1). Another binding site was identified 14 nucleotides away and is known to be a motif for ELAVL binding proteins (Figure 10).

Interestingly, both of these proteins are found together in the brain during fetal development (Berto et. al, 2016). The RBPMapper did not identify the experimentally determined UGACUG binding site in exon 4, as it only uses the highly conserved UGCAUG motif to recognize RBFOX binding sites. Using RBPMapper, many RBFOX binding sites were determined, using the conserved motif UGCAUG, with high significance (p<1.0^-6 to p<1.0^-6) from region 186492 to 186576 [Table 3], these sites were also found within the AtTRACT sequence analysis. The intronic region between
exon 5 and exon 6. ELAVL2 sites were not found with RBPMapper, as ELAVL2 is not one of the binding motifs which RBPMapper is able to scan for, though they were identified in the intronic regions when using the ATtRACT database. Additionally, sequence analysis for the RBFOX1/2 “silencer” UGACUG site revealed that it was also widely expressed throughout PLAC1 introns, near the conserved RBFOX1/2 binding site UGCAUG (Figure 11).

Figure 10. RBFOX1/2 silencer motif in Exon 4

Silencer region of RBFOX1/2 site was identified in exon 4. This sequence UGACUG was also identified 51 additional times throughout the PLAC1 gene, (Geneious Software 10.2.3)
Figure 11. RBFOX1 conserved binding motif within PLAC1 introns

Intronic regions between the 6 PLAC1 exons were assessed for the presence of the highly conservative RBFOX1 binding site motif \( U/A(GCAUG)C/A \) using RBPMapper. A total of 5 sequences were assessed, each containing the leading exon sequence and its corresponding intronic region. As shown above, the sequence sizes were 30kb, 19kb, 55kb, 56kb, and 36kb, respectively. The last sequence assessed for RBFOX1 binding site motifs also contained the CDS exon (colored orange above). A total of 105 RBFOX1 binding sites \( (p<.001) \) were identified within the intronic regions of PLAC1.
Table 3. High Significance Binding Motifs observed for RBFOX1

<table>
<thead>
<tr>
<th>Sequence Position</th>
<th>Motif</th>
<th>K-mer</th>
<th>Z-score</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>186461</td>
<td>wgcaugm</td>
<td>ugcauga</td>
<td>3.987</td>
<td>3.35e-05</td>
</tr>
<tr>
<td>186468</td>
<td>wgcaugm</td>
<td>uguauug</td>
<td>4.184</td>
<td>1.43e-05</td>
</tr>
<tr>
<td>186472</td>
<td>wgcaugm</td>
<td>uguauug</td>
<td>4.263</td>
<td>1.01e-05</td>
</tr>
<tr>
<td>186488</td>
<td>wgcaugm</td>
<td>cccaugc</td>
<td>3.763</td>
<td>8.39e-05</td>
</tr>
<tr>
<td>186492</td>
<td>wgcaugm</td>
<td>ugcaugc</td>
<td>4.553</td>
<td>2.64e-06</td>
</tr>
<tr>
<td>186496</td>
<td>wgcaugm</td>
<td>ugcauuu</td>
<td>3.763</td>
<td>8.39e-05</td>
</tr>
<tr>
<td>186511</td>
<td>wgcaugm</td>
<td>uacauug</td>
<td>3.487</td>
<td>2.44e-04</td>
</tr>
<tr>
<td>186572</td>
<td>wgcaugm</td>
<td>cgcaugc</td>
<td>4.211</td>
<td>1.27e-05</td>
</tr>
<tr>
<td>186576</td>
<td>wgcaugm</td>
<td>ugcaugc</td>
<td>4.461</td>
<td>4.08e-06</td>
</tr>
<tr>
<td>186591</td>
<td>wgcaugm</td>
<td>ggcauuc</td>
<td>3.513</td>
<td>2.22e-04</td>
</tr>
<tr>
<td>194334</td>
<td>wgcaugm</td>
<td>ggcauggg</td>
<td>3.434</td>
<td>2.97e-04</td>
</tr>
<tr>
<td>194348</td>
<td>wgcaugm</td>
<td>ugcccugc</td>
<td>3.395</td>
<td>3.43e-04</td>
</tr>
<tr>
<td>194352</td>
<td>wgcaugm</td>
<td>ugcccugu</td>
<td>3.132</td>
<td>8.68e-04</td>
</tr>
</tbody>
</table>

RNA Binding sites for the RBFOX family were identified with high significance throughout the PLAC1 intronic regions, most notably within the location between exon 5 and exon 6. This table is obtained from RBPMappe sequence analysis tool (Paz, Kosti, Ares, Cline, Mandel-Gutfreund, 2014)
Chapter IV

Discussion

PLAC1 has widespread expression in cancers with diverse tissue origins. Expression of PLAC1 in cancer is driven by two separate promoters, each controlled by different regulatory elements. The resultant transcripts all comprise an identical protein coding sequence, differing only in their 5’UTR sequences. The goal of this study was to assess the regulatory elements of PLAC1, specifically exploring 5’UTRs, exons and introns for regulatory elements which may affect its expression. Using a GFP reporter system, translational efficiency of the 5’UTR sequences for each confirmed variant was measured. Using this system, the goal was to assess functional elements of PLAC1 5’UTRs that may be contributing to a reduced translational efficiency. As 5’UTRs uORFs are regulatory elements of 5’UTRs that are known to significantly alter translation, their function in PLAC1 5’UTR sequences was explored with the GFP reporter system. Three of the four confirmed PLAC1 variants express uORFs of varying lengths. To measure their effect on translational efficiency, the AUG start codon was mutated by substituting the first A for a T, therefore eliminating the presence of a uORF. Furthermore, sequence analysis of PLAC1 exons and intronic regions identified several splice factors that may affect PLAC1 expression in a tissue specific manner. Results from this GFP reporter system demonstrated that the 5’UTR doesn’t affect translation in this system.
RNA Binding Proteins in PLAC1 5’UTRs

Placenta Specific Protein-1 (PLAC1) is widely expressed in cancers of diverse tissue origins, yet the role that PLAC1 expression plays in cancer pathogenesis is not well understood. To map locations of motifs of binding proteins within PLAC1, two different RNA binding site mapping algorithms, RBPMapper and AtTRACT databases were used to identify conserved protein binding sequences. Sequence analysis of the entire 100kb PLAC1 gene revealed 51 ELAVL2 binding site motifs and 151 RBFOX binding sites. RBFOX protein binding family functions in alternative splicing, most notably in highly specialized tissues which require a highly specialized protein diversity for development (Pedrotti et al., 2015), binding to a highly conserved binding motif (U/A)GCAU/CG.

The presence of binding sites for ELAVL, RBFOX1 and RBFOX2 proteins suggest that Fox proteins regulate the expression of PLAC1 in a tissue specific manner during fetal development, and that it is subjected to extensive alternative splicing events. The presence of the UGACUG binding motif previously determined to be a RBFOX1 and RBFOX2 driven silencer of alternative splicing, suggests a possible mechanism for tissue specific alternative splicing and exon inclusion and exclusion.

As PLAC1 expression is derived from two separate promoters, each driven by different regulatory elements, there is a strong potential for pre-mRNA to be subjected to different alternative splicing proteins present in each fetal tissue during development. PLAC1 is a large gene, at 100kb in length, and P1 transcription initiates 97kb upstream
of P2. Both intronic regions of P1 and P2 exons have an extensive number of RBFOX binding sites present, and therefore have the potential to undergo extensive alternative splicing. This also gives rise to the possibility that alternative proteins derived from the PLAC1 gene locus exist and have yet to be identified. These putative proteins would most likely only exist in fetal development. If there are alternative proteins which exist, they may also be present during cancer pathogenesis, and may be important targets of novel cancer therapies. These bioinformatic findings must be tested experimentally, to assess the impact RBFOX proteins have on alternative splicing of PLAC1, and confirm whether alternative protein isoforms of PLAC1 exist as a result of alternative splicing patterns.

Furthermore, RBFOX1 proteins are known to bind with the greatest affinity to the sequence UGACUG to cause a silencing effect in exons (Zhou, Baraniak, & Lou, 2006). The UGACUG motif was identified 51 times throughout all PLAC1 intronic regions, near the conserved RBFOX binding sites. The presence of these putative silencing regions may suggest a previously unknown PLAC1 regulatory element used for exon inclusion and exclusion across fetal and adult tissues based on tissue expression. The highly regulated expression of PLAC1 throughout fetal tissues during development indicates that it is most likely extensively alternatively spliced. The identification of many RBFOX binding sites provides further insight into how this regulation is accomplished. Furthermore, given the potential for such an extensive quantity of alternative splicing in a tissue specific manner, the co-expression of these splicing factors in cancer may alter the expression or function of PLAC1 in cancer. To this point, the expression of PLAC1 may be heavily dependent on the presence of certain splice factors.
and these may not be expressed by all cancer types. Further understanding of the regulatory elements controlling PLAC1 expression fetal development is needed to gain insight into its possible function within cancer.

Furthermore, given the number of binding sites for RBFOX proteins observed throughout the PLAC1 gene, it is likely that additional RNA binding proteins exist and function to alternatively regulate the expression of PLAC1. It is of interest to further explore what these binding proteins are, to gain a better understanding of how PLAC1 is regulated in cancer.

5’UTR Effects on Translation

uORFs are known to regulate diverse post-transcriptional functions, predominantly at the point of translational initiation. uORF sequences often repress or diminish translation of a downstream CDS (Eisen & Scannell, 2009). Using Geneious software, uORFs were identified in three of the four 5’UTR sequences of PLAC1. To address the impact that PLAC1 uORFs have on translation of the downstream CDS, 8 GFP reporter plasmids were constructed.

The first four contained one of the four 5’UTR sequences of PLAC1 directly ahead of a GFP reporter gene. In the last three, the AUG start codon of the upstream open reading frame was mutated (A>U) to eliminate the uORF function. Each was transfected into HEK293T cells and harvested 24 hours later. uORF function was assessed by obtaining the number of %GFP positive cells present.
The results showed that in normal conditions, little difference is observed between plasmids containing the mutated uORFs compared to those with wild-type 5’UTR sequences. Additionally, 5’UTR sequences derived from p2 showed approximately 50% less expression than those from p1. It is unclear what is driving this difference, as the RNA expression from these transfected cells also generally correlates with this observation. It may be possible that RNA stability of these transcripts is reduced compared to those transcripts with higher %GC content, such as transcript variant 3. Interestingly, the RNA expression of plasmids with 5’UTRs of transcript variant 3 was 350% higher than control and was higher than the protein expression observed.

Two possible explanations could be that RNA from these 5’UTRs is not translated as efficiently or that RNA from these plasmids is more stable than those containing other 5’UTR sequences. Further investigation must be performed to understand the reason for this observation.

Next, uORFs may also function during cellular stress to initiate translation at alternative start codons downstream of the primary ORF (Banasiak & Haddad, 1998). Of particular interest is the AUG start codon located 75 base pairs downstream of the primary CDS initiation site, and directly after the signal peptide sequence. If the uORFs of PLAC1 function to initiate translation at alternative start codons, it is possible that alternative initiation may result in loss of the PLAC1 signal peptide and would no longer be shuttled to the cell surface. Unfortunately, the experiments here are not able to determine whether cellular stress results in alternative translational initiation, as the GFP
plasmid tested here does not have a downstream ATG start site, such as those found in PLAC1, downstream of the signal peptide.

The translation of uORFs has previously been demonstrated in the context of cellular stress events (Young & Wek, 2016). The translation of uORFs during cellular stress serves as a global mechanism for regulating translation (Harding et al., 2003). It was of interest to determine whether uORFs from PLAC1 5’UTR sequences exist as a mechanism of translational regulation. In the cellular stress experiment tested here, there was a statistically significant reduction in %GFP positive cells in the cells containing the control vector and mutated 5’UTR of transcript variant 3, but this was not observed in the vector containing a wild-type 5’UTR sequence from transcript variant 3. This observation may suggest that the presence of a uORF in transcript variant 3 may have an effect on PLAC1 expression in response to cellular stress. Further testing should be completed with the additional 5’UTR sequences of PLAC1 to understand whether this is also observed for other uORF sequences. In future studies it would be worthwhile to evaluate whether PLAC1 uORFs initiate translation at alternative start codons downstream of the primary ORF. Alternative initiation of translation would lead to alternative protein sequences and may be indicative of additional functions of PLAC1.

Cancer expression and pathogenicity

The contribution of PLAC1 to the pathogenesis of cancer is poorly understood. As PLAC1 transcripts are expressed in different ratios across cancer types, differential
regulation by RBPs may alter its function in cancer cells and change its ability to
contribute to cancer pathogenicity. The first promoter of PLAC1, regulated by p53, lies
more than 100kb upstream of the protein coding sequence. Therefore, 5’UTR sequences
of PLAC1 are derived from almost 197kb of sequence in between. Given the extensive
function of PLAC1 in fetal development, combined with the identification of binding
sites for proteins involved in alternative splicing, is likely that additional alternative
exons of PLAC1 exist. As PLAC1 has widespread expression in the developing fetus,
most notably within highly specialized tissues, it also seems likely that additional protein
isoforms derived from this 200kb gene exist, as a result of tissue specific alternative
splicing. Further experimental investigation would need to be done to understand the
functional outcome of alternative splicing of PLAC1 by ELAVL and RBFOX proteins,
and how they may contribute to cancer pathogenicity of PLAC1.
Chapter V.
Research Limitations

The research outlined in this project attempts to identify putative differences in expression in PLAC1, to further understand how it may contribute to cancer and its progression. As the pattern of expression of PLAC1 transcripts expressed in cancer is different from those expressed in placenta, understanding the differences between transcripts is essential. As the transcripts of PLAC1 each contain different 5’UTR sequences, but identical protein coding regions, assessing the effect of the 5’UTR was an essential first step to elucidating any differences. To do this, a GFP reporter system was made to assess translational efficiency of the 5’UTR sequences of PLAC1. Using this method, there are a number of limitations to accurately assessing the translation of each 5’UTR in translational control of PLAC1. First, this system uses a very strong constitutively active promoter, which may result in a high level of transcript, and mitigate the ability to observe any translational differences between the constructs. To limit this possibility, 20% of the normal input of plasmid DNA was used for the transfection experiments.

Though, it is unclear whether the minimal differences observed in %GFP between most of the plasmids could be associated with an abundance of transcript in transfected cells. To avoid this in future experiments, the endogenous promoters P1 and P2 could be used as inducible promoters, and provide a more suitable test system, and more comparable to endogenous expression levels of PLAC1. Another additional constraint
of this research: PLAC1 5’UTRs may function differently in different cell types, especially those derived from cells which normally express PLAC1 in the fetal and placental tissues, as well as cancer cells. HEK293T cells were the only cell type used for making the assessment of PLAC1 5’UTR function in translational efficiency and therefore does not represent the possibility of different translational effects of PLAC1 5’UTRs in different cell types. Further work should be performed to understand how PLAC1 translation is differentially affected in other cell types. Furthermore, the GFP reporter system is an artificial system constructed to determine the possible effects on translation of each 5’UTR in PLAC1.

However, adding a PLAC1 5’UTR to GFP eliminates some of the potential interactions of each 5’UTR with the downstream PLAC1 ORF, which may contribute to differential translation between transcripts. For instance, the 5’UTR sequences may cause alternative initiation in downstream PLAC1 AUG initiation sites within the primary ORF. There are 18 AUG codons within the PLAC1 primary ORF (Figure 12), with 2 located within the signal peptide region of the first 12 amino acids, and 5 after the signal peptide (Figure 12).

The potential for alternative initiation to yield a PLAC1 product without a signal peptide could alter its localization within the cell, leading to cytoplasmic expression, which has previously been observed (Liu et. al., 2008), as opposed to cell surface expression normally observed in the placenta (Chang et al., 2014). The GFP reporter does not contain the same downstream AUG start sites as the PLAC1 gene, and therefore cannot assess this function. Finally, the GFP reporter system does not replicate the RNA secondary structure that would be observed in endogenous expression of a 5’UTR with
PLAC1 (Figure 13). As hairpin structure can affect the translation of an mRNA, the specific interactions of each 5’UTR with the downstream PLAC1 mRNA are lost in the GFP reporter system (Figure 13). Therefore, developing a GFP reporter system that may replicate this interaction would be more ideal for assessing the effect each 5’UTR has on translation of PLAC1.

Figure 12. AUG Codons Present in PLAC1

Downstream of the primary ORF in PLAC1 (PLAC1 TV1 is shown here), there are 18 AUG codons present within the CDS region. Each of these codons represents a potential location for alternative translation initiation. This is important for AUG codons directly 3’ of the signal peptide, as alternative initiation may change localization. The GFP Reporter system used for this research cannot make localization assessments (Geneious Software 10.2.3).
Figure 13. Differences in RNA Folding Structures

Folding Structures were measuring using default parameters in Geneious, to assess differences in the folding structure between the mRNA produced in A) GFP plasmid with a PLAC1 TV1 5’UTR compared to B) endogenous PLAC1 transcript variant 1 C) Linear sequences of A and B. (Geneious Software 10.2.3).

Bourgeois, J. R. (2015). The Characterization of the Neuropathological Consequences of Plac1 Ablation in a Mutant Mouse Model. *University of South Florida Scholar Commons.*


Geneious Software 10.2.3


PLAC1, a trophoblast-specific cell surface protein, is expressed in a range of human tumors and elicits spontaneous antibody responses.


