The Role of the QA Repeat Domain of TCERG1 in the Inhibition of C/EBPα Activity

A Thesis Submitted to the
College of Graduate Studies and Research
In Partial Fulfillment of the Requirements
For the Degree of Master of Science
In the Department of Biochemistry
University of Saskatchewan
Saskatoon

By Nicholas James Miller

Permission to Use

In presenting this thesis in partial fulfillment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use, which may be made of any material in my thesis.

Requests for permission to copy or to make use of material in this thesis in whole or part should be addressed to:

Head of the Department of Biochemistry
University of Saskatchewan
Saskatoon, Saskatchewan, S7N 5E5

Abstract

Transcription elongation regulator 1 (TCERG1) has previously been demonstrated to be an inhibitor of the transactivation and growth arrest activities of CCAAT/Enhancer binding protein alpha (C/EBPα). Furthermore, TCERG1 had been demonstrated to become relocalized from nuclear speckles to the pericentromeric regions where C/EBPα resides when both proteins are co-expressed in the cell. This thesis demonstrates that the deletion of a unique, imperfect series of 38 glutamine-alanine (QA) repeats near the amino terminus of TCERG1 is able to abrogate the ability of TCERG1 to inhibit C/EBPα-mediated growth arrest, the physical interaction between TCERG1 and C/EBPα, and the relocalization of TCERG1 from nuclear speckles when C/EBPα is co-expressed in the cell. The deletion of the QA domain demonstrated that there was a threshold amount of QA repeats required in TCERG1 for the relocalization and growth arrest inhibitory activities between TCERG1 and C/EBPα. It was demonstrated that between 11 and 20 QA repeats were required in TCERG1 to produce the relocalization from nuclear speckles or to be able to inhibit C/EBPα-mediated growth arrest. The physical interaction of TCERG1 and C/EBPα as examined by co-immunoprecipitation was also found to be QA dependent, with a diminishing interaction observed as the number of QA repeats in TCERG1 were reduced. However, experiments examining the isolated QA domain indicated that it was insufficient to relocalize an mCherry fluorescent protein fusion to either the nucleus or to pericentromeric regions where C/EBPα is concentrated. This inability to produce relocalization suggests that the QA domain requires another domain or domains from TCERG1 to mediate the relocalization activity. When expressed with the WT TCERG1, ΔQA TCERG1 was able to act in a dominant negative manner, preventing the relocalization of the WT TCERG1 protein to pericentromeric domains. Interestingly, the transactivation inhibitory activities of TCERG1 on C/EBPα do not appear to require the QA domain, but rather are localized to the carboxy half of TCERG1, somewhere within amino acids 612-1098. The data obtained provides the first report of a role for this unique QA repeat domain.

Acknowledgements

Firstly I would like to thank my supervisor, Dr. William Roesler for the support he has shown me in this project. His drive and direction has been invaluable to ensure I don't stray too far from the path in this project.

I would also like to thank my committee members for their guidance, Dr. Scot Stone, and Dr. Erique Lukong. I would also like to thank Dr. Ramji Khandelwal for chairing my defense as well as Dr. Chris Eskew for providing feedback as my external examiner.

Lastly I would like to thank Drs. Kenneth Gagnon and Sheng-Pin Hsiao for their friendship, guidance, troubleshooting and general sounding board advice.

"The definition of insanity is doing something over and over again and expecting a different result."
-Albert Einstein

To my wife, Natasha.

Without you, this wouldn’t have been possible.

Table of Contents

Permission to Use ..... **i**

Abstract ..... **ii**

Acknowledgements ..... **iii**

Table of Contents ..... **v**

List of Tables ..... **ix**

List of Figures ..... **x**

List of Abbreviations ..... **xii**

1. Introduction and Overview **1**

2. Review of the Literature **2**

2.1 Transcription and Splicing 2

2.2 Subnuclear Compartmentalization 3

2.2.1 Pericentromeric Region *4*

2.2.2 Nuclear Speckles *5*

2.3 CCAAT/Enhancer Binding Protein Family (C/EBP) 8

2.3.1 The Growth Arrest Properties of C/EBPα *10*

2.3.1.1 The E2F-Rb Model of C/EBPα-Mediated Growth Arrest 11

2.3.1.2 The Cdk2/4 Model of C/EBPα-Mediated Growth Arrest 13

2.3.1.3 The SWI/SNF model of C/EBPα-mediated growth arrest 14

2.3.1.4 The Combined Model of C/EBPα-Mediated Growth Arrest 15

2.3.2 C/EBPα Functionality is Not Location Dependent *17*

2.3.3 C/EBPα and its Role in Disease *18*

2.4 Transcription Elongation Regulator 1 (TCERG1) 19

2.4.1 TCERG1 and its Role in Transcription and Splicing *20*

2.4.2 TCERG1 Modulates the Rate of Transcription of RNAP II *21*

2.4.3 TCERG1 as an RS Domain Protein *22*

2.4.4 TCERG1 QA_{38} Domain *23*

2.4.4.1 Evolution of the QA Domain of TCERG1 24

2.5 TCERG1 is an Inhibitor of C/EBPα 25
2.6 Hypothesis 27
2.7 Rationale and Objectives 27

3. Materials and Methods 28

3.1 Reagents 28
3.2 Bacterial Strains and Media Preparations 30
3.3 Molecular Cloning 30
3.3.1 Transformation of Competent Bacterial Cells 30
3.3.2 Plasmid DNA Preparations 31
3.3.3 Plasmids 31
3.3.3.1 Restriction Digest of Plasmid DNA 32
3.3.3.2 DNA Fragment Isolation and Verification 33
3.3.3.3 DNA Sequencing 33
3.3.4 DNA Cloning Techniques 33
3.3.4.1 Gibson Assembly Plasmid Engineering 33
3.3.4.2 Restriction Site Plasmid Engineering 34
3.3.4.3 cDNA Amplification and Cloning from HEK293T RNA 34
3.4 Mammalian Cell Culture 35
3.4.1 Transient Cell Transfection 35
3.4.1.1 PEI MAX Transfection 35
3.4.1.2 LipoD293 Transfection 36
3.4.2 Growth Arrest Assay 36
3.5 Mammalian Cell Culture Protein Expression Determinations 36
3.5.1 Preparation of Protein Extracts From Cultured Cells 36
3.5.2 Protein Quantification of Cellular Extracts 37
3.5.3 SDS Polyacrylamide Gel Electrophoresis 37
3.5.4 Western Blotting 38
3.5.5 Co-Immunoprecipitation 39
3.5.6 Luciferase Assay 40
3.6 Laser Scanning Confocal Microscopy 41
3.6.1 Transient Transfection and Microscope Slide Preparation 41

3.6.2 Preparation of Cells for Immunostaining *41*
3.6.3 Image Acquisition and Manipulation *42*

4. Results **43**
4.1 Role of the QA domain of TCERG1 in inhibiting C/EBPα-induced growth arrest 43
4.2 The Length of the QA Domain in TCERG1 Influences the C/EBPα-Mediated Relocalization of TCERG1 45
4.2.1 The QA Domain is Required for the Relocalization of TCERG1 from Nuclear Speckles to Pericentromeric Regions *47*
4.2.2 The Relocalization of TCERG1 by C/EBPα is Not Fluorophore Specific *51*
4.2.3 ΔQA TCERG1 Acts as a Dominant Negative With Respect to Relocalization of TCERG1 ..... *54*
4.2.4 Differential Localization Patterns of C/EBPα and TCERG1 are Demonstrated Through Amino or Carboxy-End Epitope Tagging of C/EBPα *59*
4.2.4.1 Differential Epitope Tagging of C/EBPα Does Not Change the Localization Pattern of C/EBPα Except for Large Carboxy-Terminal Tags 64
4.3 The QA Domain of TCERG1 is Important For Some of The Interactions Between TCERG1 and C/EBPα 66
4.3.1 The QA Domain of TCERG1 is Involved in the Physical Interaction Between TCERG1 and C/EBPα *66*
4.3.2 The Isolated QA Domain is Unable to Mediate Relocalization *68*
4.4 The QA Domain Does Not Play a Role in the Transactivation Inhibition of C/EBPα by TCERG1 75
4.4.1 The C/EBPα Epitope Tag Does Not Affect the Ability of TCERG1 to Inhibit C/EBPα Transactivation Potential. *76*

5. Discussion **81**
5.1 The Growth Arrest and Transactivation Inhibition of C/EBPα by TCERG1 are Mediated by Different Domains in TCERG1 81
5.2 The TCERG1 QA Repeat Domain is Important for Inhibitory Activity Toward C/EBPα 82
5.2.1 The Relocalization of TCERG1 by C/EBPα is QA Length Dependent *82*

5.2.2 The Isolated QA Domain is Unable to Target Proteins to the Nucleus, Nuclear Speckles or Mediate Relocalization by C/EBPα.....84
5.3 The Carboxy Terminus of TCERG1 Mediates the Transactivation Inhibition of C/EBPα 86
5.4 The Epitope Tagging of C/EBPα Has Implications For its Function.....87
5.5 Complex Domain Structure Provides Opportunities to Coordinate Different Nuclear Events 89
5.6 Summary of Findings Presented in This Thesis.....92
6. Future Directions.....95
6.1 Further QA Domain Exploration.....95
6.2 C/EBPα Interaction Site 96
6.3 Whole Animal Testing 97
7. References 98

List of Tables

Table 1: List of Reagents and Suppliers....28
Table 2: List of Names and Addresses of Reagent Suppliers....29
Table 3: Plasmids Created for This Project....31
Table 4: Plasmid Backbones Used....32
Table 5: Pre-existing Plasmids Used....32
Table 6: List of Antibodies Used....39

List of Figures

Figure 1 – Functional domains of C/EBPα. 10
Figure 2 – The Growth Arrest Domain Specific Sites of C/EBPα....13
Figure 3 – Domain structure of TCERG1. 20
Figure 4 – The RS amino acid composition of TCERG1 suggests that FF4 and FF5 are acting as RS domains....23
Figure 5 – The evolution of the TCERG1 QA domain. 25
Figure 6 – The QA domain from TCERG1 affects the growth arrest potential of C/EBPα. 44
Figure 7 – TCERG1 deletion mutants. 46
Figure 8 – C/EBPα mutant schematics. 47
Figure 9 – Shortening or deletion of the QA domain of TCERG1 does not alter the localization of TCERG1 to nuclear speckles. 48
Figure 10 – Deletion of the TCERG1 QA domain abrogates the relocalization of TCERG1 from nuclear speckles to pericentromeric domains in the presence of C/EBPα. 50
Figure 11 – The nuclear localization pattern of TCERG1 does not change dependent upon the fluorophore fused to TCERG1. 53
Figure 12 – TCERG1 ΔQA prevents TCERG1 WT from being relocalized by C/EBPα. 57
Figure 13 – Endogenous TCERG1 along with the ΔQA mutant of TCERG1 shows overlap of expression patterns. 58
Figure 14 – Amino tagged eGFP-C/EBPα relocalizes TCERG1 in a QA dependent manner whereas carboxy tagged eGFP-C/EBPα does not. 60
Figure 15 –AcGFP-IRES-C/EBPα and eGFP-P2A-C/EBPα produce a single band of the correct molecular weight. 63
Figure 16 – Untagged C/EBPα displays similar TCERG1 relocalization characteristics to eGFP-C1-C/EBPα. 63
Figure 17 – Differential epitope tagging of C/EBPα produces patterns of C/EBPα expression in COS7 cells except for fluorescent markers fused to the carboxy terminus. 65
Figure 18 – The QA domain of TCERG1 is important for the physical interaction between TCERG1 and C/EBPα. 67
Figure 19 – Schematic of the mCherry-QA fusion constructs used. 68

Figure 20 – The QA domain of TCERG1 when fused to the carboxy terminus of mCherry does not act as an NLS nor does it co-localize with C/EBPα. 70
Figure 21 – Fusing the QA domain to the amino terminus of mCherry does not change the localization of the isolated QA domain. 72
Figure 22 – Even when localized to the nucleus, the QA domain alone does not co-localize with C/EBPα. 74
Figure 23 – The QA domain of TCERG1 is not required for the transactivation inhibition of C/EBPα by TCERG1. 76
Figure 24 – Differential epitope tagging of TCERG1 does not affect the transactivation inhibition of C/EBPα using TCERG1 and amino and carboxy terminal deletion mutants. 80
Figure 25 – Representation of the Interactions Between TCERG1 and C/EBPα Explored in This Thesis 94

List of Abbreviations

19R	pTZ19R plasmid
AML	Acute Myeloid Leukemia
bZIP	Basic Region-Leucine Zipper
C/EBPα	CCAAT/Enhancer-Binding Protein Alpha
CBP	CREB-Binding Protein
CDK	Cyclin Dependent Kinase
CMV	Cytomegalovirus
Co-IP	Co-Immunoprecipitation
COS7	African Green Monkey Kidney Cells
CTD	Carboxyl Terminal Domain of RNA Polymerase II
DBD	DNA Binding Domain
dH2O	Distilled Water
ddH2O	Deionized Distilled Water
DMEM	Dulbecco's Modified Eagle Medium
DNA	Deoxyribonucleic Acid
E2F	Early Gene 2 Factor
E. coli	*Escherichia coli*
EDTA	Ethylene-Diamine Tetraacetic Acid Disodium Salt
eGFP	Enhanced Green Fluorescent Protein
FLAG	Protein Sequence of DYKXXD
FRET	Forster Resonance Energy Transfer
HA	Hemagglutinin
HEK293T	Human Embryonic Kidney 293T
HEPES	2-[4-(2-hydroxyethyl)piperazin-1-yl]ethanesulfonic acid
HD	Huntington’s Disease
HRP	Horseradish Peroxidase
IB	Immunoblot
IGC	Interchromatin Granule Cluster
IP	Immunoprecipitation

IRES	Internal Ribosome Entry Site
kDa	Kilodalton
KE	Lysine-Glutamate Rich Domain
LB	Luria-Bertani
MIG	Mitotic Interchomatin Granule Cluster
NT	Nucleotide
NLS	Nuclear Localization Signal
P2A	2A Peptide Derived from Porcine Teschovirus-1
PAGE	Polyacrylamide Gel Electrophoresis
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PFA	Paraformaldehyde
PP2A	Protein Phosphatase 2A
QA	Glutamine-Alanine Repeat
RB	Retinoblastoma Protein
RNAP	RNA Polymerase
RT	Room Temperature
SDS	Sodium Dodecyl Sulfate
STP	Serine, Threonine, Proline Rich Domain
SWI/SNF	SWItch/Sucrose Non-Fermentable Complex
SV40	Simian Virus 40
TAE	Tris-Acetate EDTA Buffer
TBP	TATA-Binding Protein
TBS	Tris-Buffered Saline
TBST	Tris-Buffered Saline + 0.05% Tween-20
TCERG1	Transcription Elongation Regulator 1
Tris	Tris-[hydroxymethyl]-aminomethane

1. Introduction and Overview

Some years ago, the Roesler lab sought to identify interactors of the transcription and differentiation factor, C/EBPα, using a two-hybrid screen that might shed light on its mechanism of action and regulation (McFie *et al.*, 2006). This screen identified TCERG1 (formerly CA150) as a potential candidate, and through several publications the lab has established that TCERG1 interacts with and inhibits C/EBPα. The mechanism of C/EBPα inhibition by TCERG1 is unusual in that it involves the regulation of the nuclear compartmentation of TCERG1 (Banman *et al.*, 2010; McFie *et al.*, 2006; Moazed *et al.*, 2011). C/EBPα concentrates at pericentromeric regions through its ability to bind to the α-satellite repeats situated near the centrosomes of chromosomes, and is thought to be recruited from these sites when needed for transcriptional regulation or growth arrest. On the other hand, TCERG1 localizes to nuclear speckles, which are sites of storage for splicing and transcription factors in the nucleus. Intriguingly, in Banman *et al.*, (2010) they demonstrated that when C/EBPα and TCERG1 are co-expressed in cells, TCERG1 undergoes relocalization from nuclear speckles to the pericentromeric regions where C/EBPα resides. Moreover, Moazed *et al.*, (2011) demonstrated that C/EBPα activity is inhibited when this relocalization occurs, and through various TCERG1 mutants it was shown that inhibition of C/EBPα by TCERG1 requires relocalization of TCERG1. Based on these findings, it is hypothesized that TCERG1 inhibits C/EBPα by preventing the release of C/EBPα from its intracellular depots. TCERG1 was discovered to contain a unique domain only found in one other human protein; an extended, imperfect repeat of 38 QA residues with a few of the alanine residues substituted by valines. If denoted properly the QA domain would be $QAQV(QA)_{28}(QAQV)_3(QA)_2$ but for brevity is usually termed QA_{38}.

This thesis is focused on characterizing the role of the QA domain in mediating these activities between C/EBPα and TCERG1: the relocalization activity, C/EBPα-inhibitory activity and the interaction between C/EBPα and TCERG1 with respect to TCERG1. Previous work to this thesis demonstrated that all of these activities lie in the N-terminal half of the protein (Banman *et al.*, 2010; Moazed *et al.*, 2011), and that none of the WW domains play a role (unpublished observations).

2. Review of the Literature

2.1 Transcription and Splicing

The processes of transcription and splicing are highly complicated and interconnected. In humans, transcription is achieved using three different polymerases: RNA Pol I, II or III (RNAP I, II, or III). Initially thought to be separate processes, in recent years there has been a surge in research tying transcription and splicing processes together in an elaborate way. RNAP II is responsible for the synthesis of mRNA in the cell and as such is intricately tied to the splicing machinery. As transcription of the gene is occurring, the polymerase appears to be able to signal and recruit the splicing machinery to begin splicing the elongating mRNA transcript even before it is finished being transcribed. RNAP II contains a conserved C-terminal domain (CTD) which contains repeats of YSPTSPS which vary in number across eukaryotes from 26 to 52 repeats although the consensus sequence varies slightly from species to species (Hsin and Manley, 2012; Yogesha *et al.*, 2014). This CTD can be modified using post-translational modifications such as phosphorylation and in turn these post-translational modifications serve as signals to other proteins to perform a variety of functions. Phosphorylation of the CTD appears to be highly important for signaling members of the transcription and splicing machineries (Hsin and Manley, 2012).

Two different models have been proposed to describe the mechanism whereby transcription and splicing are linked: the recruitment model and the kinetic model (Montes *et al.*, 2012). In the recruitment model, the CTD of RNAP II functions as a hub where the various proteins involved in the splicing process are recruited to and are assembled according to the CTD "code" of phosphorylation or other post-translational modifications. In the kinetic model of transcription and splicing, the transcription elongation rate controls splicing and alternative splicing. The kinetic model relies on various transcription and splicing factors acting in such a way to effectively speed up or slow down splicing reactions to affect alternative splicing of mRNA. By changing the rate of transcription, various splicing factors may be able to bind which would not have been able to do so previously. The binding of these alternative splicing

factors would change the splicing characteristics for the transcript. TCERG1, which is examined in this thesis, has been suggested to be one of the transcription and splicing factors which are able to affect the rate of transcription and therefore the splicing reactions taking place (Montes *et al.*, 2012). There is also evidence that TCERG1 is able to mediate the phosphorylation status of the CTD as well (Coiras *et al.*, 2013).

2.2 Subnuclear Compartmentalization

The common misconception of the nucleus is that it has no higher order structure since it is not compartmentalized like the rest of the cellular organelles. However, we are discovering that this is not the case. The nucleus is a very protein-rich environment; the protein levels inside the nucleus are higher than outside at an estimated 400 mg/mL compared to the estimated cellular values of ~200 mg/mL (Milo, 2013; Misteli, 2007). For the proper functioning of the nucleus it needs to be able to efficiently organize the different proteins involved in varying aspects of nuclear function so that they can be effectively recruited at appropriate levels at the appropriate times in the required places in the nucleus (Misteli, 2007; Sleeman and Trinkle-Mulcahy, 2014). To accomplish this, the nucleus has evolved many different suborganelles to perform specific functions, similar to organelles found in the rest of the cell. A membrane does not demarcate suborganelles, but they are able to mediate very specific functions inside the nucleus. This is similar to the rest of the cell, which contains different organelles such as mitochondria, endoplasmic reticulum, golgi apparatus, etc.

The organization of the nucleus has three separate levels: the organization of the chromatin, the arrangement of the genes and chromosomes in the nuclear space, and the basic organization of nuclear processes such as splicing and transcription (Misteli, 2007). Each of these levels are independently regulated inside the nucleus. For example, some of these compartments are used for storage of transcription factors and mRNA processing proteins whereas others are sites of active transcription (Dundr and Misteli, 2010; Misteli, 2005, 2007; Schor *et al.*, 2012). It is still unknown how many of these suborganelles are formed, whether they form due to specific interactions inside these sites, or if there are structural proteins that have yet to be discovered which mediate their formation. What we do know is that these sites are dynamic in nature, forming fairly predictably, although the size and number of sites of each

of these features can vary from cell to cell (Dundr and Misteli, 2010). It has been demonstrated that the proteins inside of these sites can shuttle rapidly from site-to-site as well as through the surrounding nucleoplasm which can be monitored using experiments such as Förster resonance energy transfer (FRET) (Dundr and Misteli, 2010). The ability of proteins to shuttle between these compartments effectively and efficiently at the correct time is crucial for the proper functioning of the cell and cellular survival (Dundr and Misteli, 2010). Nuclear body biogenesis is also very important to proper functioning as many of these organelles break down during mitosis and need to be regenerated upon completion of mitosis for the proper functioning of the nucleus. Two such subnuclear compartments that are relevant to this project are pericentromeric regions and nuclear speckles.

2.2.1 Pericentromeric Region

While not traditionally considered subnuclear compartments since they do not typically catalyze processes inside the nucleus, pericentromeric domains have some characteristics of subnuclear compartments. These characteristics become important especially in terms of the proteins related to this thesis. Pericentromeric regions are heterochromatic sites near the centromeres of chromosomes that contain α-satellite repeat sequences. Alpha-satellite repeat sequences are 171 base pair repeated sections of DNA located at the centromeres of chromosomes (Horvath *et al.*, 2000). The repeats are organized in these regions into larger subunits consisting of α-satellite DNA which comprises much of the heterochromatic DNA in the centromeric regions. Together, the α-satellite repeat regions are estimated to comprise 3-5% of all genomic DNA in the cell. There are still gaps in our knowledge about α-satellite repeats since due to their highly repetitive nature they are extremely difficult to research. In fact, although most of the human genome has been sequenced, the centromeres of chromosomes continue to be problematic to sequence due to their repeating nature. Certain DNA binding proteins can use these large repeat domains for a variety of purposes but mostly it is suspected that proteins which bind to these domains do so as a means of protein segregation and/or storage (Liu *et al.*, 2007). There is evidence which suggests that these α-satellite sequences sequester and concentrate various proteins such as transcription factors which are currently not needed at that moment in the cell, thus acting as a reservoir (Liu *et al.*, 2007).

Pericentromeric regions are also typically associated with transcriptional silencing of genes along with the promotion of genome stability (Lejeune *et al.*, 2010). Pericentromeric regions can be visualized using DNA staining methods such as staining the nucleus with Hoescht 33342 which binds to the minor groove of double stranded DNA but preferentially to sequences rich in adenine and thymine (Portugal and Waring, 1988). Since the pericentromeric regions are rich in highly compact heterochromatic DNA, this means that a DNA binding dye such as Hoescht 33342 binds in greater quantities in these areas and therefore is seen as areas of intense staining when observed under a fluorescent microscope. Several researchers who study the dynamics of the pericentromeric region binding protein, C/EBPα, have effectively used this dye to demarcate the pericentromeric regions inside the nucleus (Liu *et al.*, 2002; Zhang *et al.*, 2001).

2.2.2 Nuclear Speckles

Nuclear speckles are sites of storage of transcription and splicing factors in the interchromatin regions of the nucleoplasm and as such are rich in transcription factors and pre-messenger RNA splicing factors (Lamond and Spector, 2003). Nuclear speckles are often found in numbers ranging from 20-50 and 2-3 μm in size, although the size of nuclear speckles can expand under certain conditions such as the inhibition of transcription using α-amantin (Dundr and Misteli, 2010). Nuclear speckles are usually found next to sites of active transcription but they do not appear to participate in transcription directly, reinforcing the notion that nuclear speckles are simply sites of storage for splicing and transcription factors (Rieder *et al.*, 2014; Schor *et al.*, 2012). The conservation of nuclear speckles across species suggests that they play an important role in the nucleus. Nuclear speckles have been discovered in all of the more complex eukaryotes such as humans, mice, clawed frogs, and fruit flies but are absent in zebrafish, nematodes, and yeast (Morimoto and Boerkoel, 2013).

Two separate groups discovered nuclear speckles only a few years apart. Initially, Hewson Swift described a region in the nucleus in 1959 when he observed through electron microscopy a part of the nucleus that would form patterns. These patterns were termed interchromatin granule clusters (IGCs) which were comprised of small granules, each measuring 20-25 nm in diameter (Misteli and Spector, 1997). In 1961, J. Swanson Beck noted speckles of what was discovered to be splicing and transcription factors inside the nucleus. At

this time, these “speckles” were thought to be sites of active transcription; later, these sites were officially termed “nuclear speckles”. It was subsequently discovered that by imaging the nucleus using the same protein markers it was determined that ICGs and nuclear speckles were indeed one and the same (Spector and Lamond, 2011).

The purpose of the storage of splicing factors is thus far not fully known. It has been suspected that the nucleus may use these sites for a variety of reasons including the potential assembly of splicing and/or transcription factors prior to recruitment, and/or regulating the concentration of factors in the nucleoplasm and therefore their solubility. As well, nuclear speckles likely affect the ability of the nucleus to rapidly respond to changing gene expression patterns (Misteli and Spector, 1998).

The biogenesis of nuclear speckles is currently unknown although they have been shown to break down during mitosis and then rapidly regenerate upon completion of mitosis, suggesting there may be some seeding mechanism which helps the rapid re-formation (Spector and Lamond, 2011). There is some evidence to suggest that nuclear localized, nascent poly(A)-RNA could potentially form the basis of the scaffolding for nuclear speckles. Poly(A)-RNA has been detected in nuclear speckles as well as sites where nuclear speckles form. When cells were treated with an inhibitor of speckle formation, tubercidin, it was discovered that the poly(A)-RNA becomes dispersed and the nuclear speckles break down (Kurogi *et al.*, 2014). It has also been recently observed that there is a pool of poly(A)-RNA in the cell which is not degraded when the nuclear envelope breaks down during mitosis. This poly(A)-RNA is subsequently transported back through the nuclear envelope upon the re-formation of the nucleus, potentially acting as a seeding scaffold for the formation of nuclear speckles (Huang and Spector, 1996).

The formation and inheritance of nuclear speckles has been highly researched, although it is still not certain how these processes occur. In early mitosis, nuclear speckles disassociate and their components become dispersed in the cytoplasm. The dispersion allows the sharing of components when the cell undergoes division. These dispersed components then re-form in cytoplasmic clusters referred to as mitotic interchromatin granule clusters (MIGs) during late telophase and are subsequently recruited into the nucleus once re-formed (Prasanth *et al.*, 2003). During early G1 phase, the pre-mRNA splicing machinery is recruited into the nucleus from MIGs and as the levels of RNA polymerase II transcription increase the nuclear speckles

form in an RNAP II activity-dependent manner (Dundr and Misteli, 2010; Prasanth *et al.*, 2003). This activity-dependent formation mirrors the determination of nuclear speckle size in the cell, which can vary depending upon the cellular levels of RNAP II transcription. Therefore, if there is a large amount of transcription occurring in the cell, the nuclear speckles form slowly as the transcription factors are recruited to the sites of transcription, whereas if transcription is low, the nuclear speckles form quickly.

Although there have been no definitive answers to determine how nuclear speckles form and are regulated, it has been demonstrated using photobleaching FRET experiments that the flow of individual proteins into and out of nuclear speckles is very rapid (Lamond and Spector, 2003). Upon bleaching of the fluorophores of proteins inside of the nuclear speckles, it was demonstrated that the nuclear speckles are able to half recover their pre-bleaching fluorescence levels after only approximately 3-5 seconds; this would suggest that the exchange rates of the proteins in nuclear speckles is quite substantial (Lamond and Spector, 2003). While it is not fully understood why or how proteins shuttle so rapidly into and out of nuclear speckles, it has been suggested to be due to the change in phosphorylation rates of the individual proteins during transcription and splicing reactions (Misteli and Spector, 1997). This phosphorylation change could be a major player in the turnover of proteins inside nuclear speckles (Misteli and Spector, 1997).

It has been demonstrated that it is possible to target proteins to nuclear speckle compartments using a basic amino acid region next to a Ser-Arg (SR) rich domain, although other researchers have suggested that all that is required is the SR rich region (Misteli and Spector, 1997; Spector and Lamond, 2011). The leading theories on the factors controlling the flow into and out of speckles of various splicing factors suggest that phosphorylation levels of these SR rich regions are a major factor in controlling whether individual proteins which can be directed to nuclear speckles are stored in nuclear speckles (Carrero *et al.*, 2006; Mao *et al.*, 2011). The transcription and splicing machinery is actively phosphorylated and dephosphorylated using several kinases and phosphatases; several of these are also localized to nuclear speckles which suggest a direct regulatory role in the post-translational modification of splicing factors (Mao *et al.*, 2011; Sleeman and Lamond, 1999). Hypo-phosphorylation appears to keep proteins inside nuclear speckles, ready for activation (Sacco-Bubulya and Spector, 2002). Upon hyper-phosphorylation the proteins become active and participate in

splicing reactions. This hyper-phosphorylation lowers the protein's affinity for nuclear speckles. During splicing, various phosphatases lower the phosphorylation states of the splicing proteins, causing the proteins to re-establish their affinity for speckles, thereby preparing them for another round of splicing. It was also noticed that nuclear speckles become larger and more round when transcription is inhibited (Spector and Lamond, 2011). This enlargement would suggest that the transcription factors inside of the nuclear speckles are no longer recruited away from the nuclear speckles and therefore are "stored" longer inside nuclear speckles.

2.3 CCAAT/Enhancer Binding Protein Family (C/EBP)

The C/EBPs (CCAAT-Enhancer binding protein) are a family of six different transcription factor isoforms, named based upon the order in which they were discovered (alpha, beta, epsilon, gamma, delta and zeta) (Nerlov, 2008). The C/EBP family has been shown to play pivotal roles in several cells types such as haematopoietic cells, adipocytes and hepatocytes by regulating numerous responses including inflammation, metabolism, cellular proliferation, and differentiation. Moreover, they also act as accessory factors in hormone response units by interacting with other transcription factors to enhance specific hormone responsiveness (Ramji and Foka, 2002; Roesler, 2001).

C/EBP proteins function as dimers, due to a highly conserved leucine zipper domain that is found in all of the C/EBPs. The carboxy terminus contains a basic region of amino acids along with a leucine zipper (bZIP). This bZIP domain allows two monomers of C/EBPs to bind each side of the DNA strand and create a dimer that is bound to DNA. This DNA bound dimer allows the C/EBPs to then mediate various functions such as transcriptional activation. This feature allows for both homo-and heterodimers of C/EBPs in cells as all of the C/EBP family members contain the highly conserved bZIP domain and therefore are able to dimerize with other members of the C/EBP family.

The first C/EBP isoform discovered, C/EBPα, was coined by McKnight and coworkers as a "central regulator of energy metabolism" as it was discovered to play pivotal roles in several metabolic pathways (McKnight *et al.*, 1989). C/EBPα is expressed in high levels in adipose tissues, liver, lung, intestine, adrenal glands, peripheral-blood mononuclear cells as

well as placenta. C/EBPα-knockout mice die soon after birth due to severe hypoglycemia, a result of reduced expression of glycogen synthase, which leads to undetectable levels of glycogen in the liver, as well as three gluconeogenic enzymes (glucose-6-phosphatase, phosphoenolpyruvate carboxykinase and tyrosine aminotransferase) (Wang *et al.*, 1995). In addition to being a transcription factor, C/EBPα also has a separate function in regulating the differentiated state of cells by mediating growth arrest at the G1/S boundary. A domain distinct from the transactivation domains mediates this growth arrest function of C/EBPα and the mechanism varies somewhat depending on the cell type.

C/EBPα is expressed from a single intronless gene that contains two alternative start codons (see figure 1). These two codons allow for two different forms of C/EBPα to be produced, which are designated as p42 and p30. The p42 isoform of C/EBPα is the full-length product shown in figure 1 whereas the p30 isoform contains an amino terminus truncation that excludes both activation domains. The N-terminal part of C/EBPα which is deleted contains an E2F binding domain, therefore without this part of the protein in the p30 form C/EBPα is unable to mediate G1/S growth arrest in certain cell types (Roe and Vakoc, 2014).

As shown in figure 1, p42 C/EBPα contains several activation domains along with a repression domain, a basic region and a leucine zipper. C/EBPα can act as a transcription factor that is able to mediate the activation or repression of different target genes using the three activation domains and the single repression domain found in the amino terminus of the protein. Along with these domains there is also a growth arrest domain around S193. The growth arrest domain, as explained previously allows C/EBPα to put the cell into a state of growth arrest but C/EBPα is not required to be bound to DNA for this to occur (Liu *et al.*, 2002). In the carboxy terminal half of C/EBPα, two domains are found which mediate the DNA binding properties of C/EBPα. By using the basic amino acid region, C/EBPα is able to bind DNA sequences and the leucine zipper allows for the dimerization of two members of the C/EBP family to come together to mediate a variety of nuclear activities.

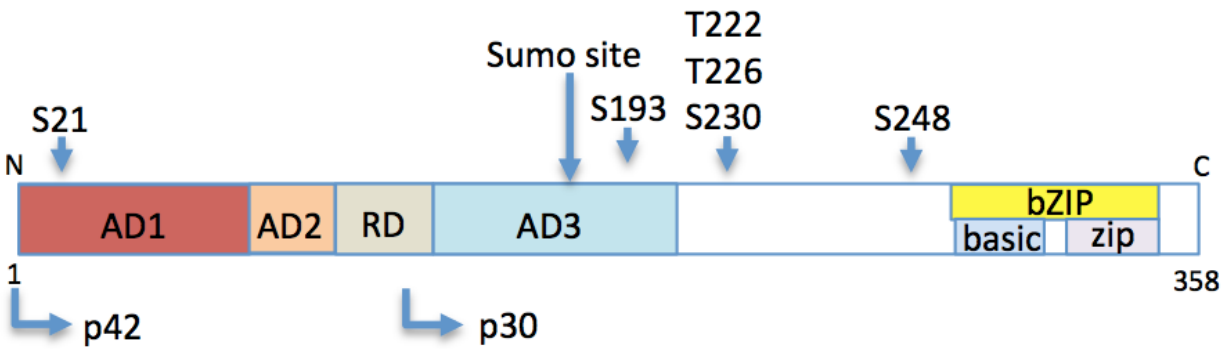

Figure 1 – Functional domains of C/EBPα.

AD denotes activation domains, RD denotes the negative regulatory domain, basic denotes the basic region, zip is the leucine zipper region, bZIP denotes the combined basic and leucine zipper regions. The p42 and p30 arrows denote the full length and amino truncated isoforms, respectively. The top arrows depict sites of phosphorylation and sumoylation. Around S193 there is also a growth arrest domain.

2.3.1 The Growth Arrest Properties of C/EBPα

Although it is well established that C/EBPα is able to mediate the growth arrest of cells at the G1/S boundary, there has yet to be consensus as to the exact mechanism mediating this. There are several proposed mechanisms of C/EBPα-mediated growth arrest but the prevailing hypothesis implicates the early gene 2 factor (E2F) family, a family of transcription factors which activate the genes controlling G1/S cell cycle progression. In this model the E2F acts as the controlling factor with retinoblastoma protein (RB) as a secondary factor which modulates the function of E2F. Although the E2F-Rb model is regarded as the most likely mechanism driving C/EBPα-mediated growth arrest, there are two other models, one which implicates the SWI/SNF complex as a potential mediator of growth arrest and the other which proposes the interactions of cyclin dependent kinase 2 (CDK
+) and cdk4 with C/EBPα are able to mediate C/EBPα-mediated growth arrest. Since C/EBPα is only expressed in certain cell types, at specific times and at fairly low levels it is difficult to study the exact mechanism of growth arrest. By over-expressing proteins or using cell lines derived from tissues without natural C/EBPα expression the mechanisms observed could provide artifactual data (Johnson, 2005). It will be up to future groups to comb through the currently available data to build a cohesive view of what is occurring in this complex system. Currently in the literature the three models are presented separately but there are links between

the models. Using the knowledge obtained thus far we can begin to piece together a cohesive picture of the growth arrest pathway as one signaling pathway. In the next three sections each current model will be presented and then in the fourth section a discussion of how each model could tie into each other will be presented.

2.3.1.1 The E2F-Rb Model of C/EBPα-Mediated Growth Arrest

The E2F family of cell cycle transcription factors function to modulate the various stages of growth a cell undertakes. By binding to different promoters and/or activating or repressing various proteins in the cell, the various E2Fs are able to provide the cell with the various checks which are required for the proper division and proliferation of cells. When C/EBPα is not expressed in cells, unphosphorylated retinoblastoma protein (Rb) is able to bind to E2F, thereby inactivating it (Khanna-Gupta, 2008). Retinoblastoma protein is part of a family of proteins which are commonly referred to as pocket proteins. Pocket proteins are able to bind to the various E2Fs and mediate differential binding and repression characteristics. It is this E2F-Rb complex that C/EBPα is able to eventually bind to and stabilize, allowing for C/EBPα-mediated growth arrest. Although Rb binding to E2F is mentioned in this model, it is worthwhile to note that other pocket proteins such as p107 or p130 are also able to bind and modulate the function of the E2F complex (Harbour and Dean, 2000). Therefore these proteins may play a role in the C/EBPα-mediated growth arrest as well. The binding of the other pocket proteins is not as extensively studied as E2F-Rb binding and may be a source of insight into the inconsistencies with this model.

The various E2F proteins in a cell do not function as a single unit; rather, there are a family of E2F-dimerization partner (DP) transcription factors that are able to bind to E2F to form a variety of complexes, presumably with different properties and targets. There is also evidence that the DP proteins are able to bind to C/EBPα and repress specific genes such as E2F-target genes as well as binding the bZIP and inhibiting the transcriptional activity of C/EBPα (Zaragoza *et al.*, 2010).

When C/EBPα is expressed in a cell it strengthens the E2F-DP-Rb complex, although the exact mechanism by which this occurs has been subject to some debate. The domain around S193 encompassing amino acids 175-217 has been termed the minimal growth arrest domain

and appears to be sufficient to induce growth arrest (Wang *et al.*, 2001). Phosphorylation of the S193 residue in C/EBPα is able to promote C/EBPα-mediated growth arrest while dephosphorylation of this residue inhibits it (Khanna-Gupta, 2008). When S193 is phosphorylated, C/EBPα is able to bind to cyclin-dependent kinase-2 as well as the SWI/SNF ATPase subunit, thereby preventing proliferation of the cell (Khanna-Gupta, 2008). When S193 is dephosphorylated it promotes C/EBPα to bind to Rb (Slomiany *et al.*, 2000). This binding of C/EBPα to Rb is disruptive to the interaction of E2F-DP-Rb and therefore the E2F is able to bind to and promote the advancement of the cell cycle. This binding and subsequent disruption of the E2F-DP-Rb complex is C/EBPα dose-dependent (Slomiany *et al.*, 2000).

Although C/EBPα is able to bind to E2F target promoter sites, the growth arrest activities of C/EBPα have been demonstrated to be separate from the DNA binding activities of C/EBPα; therefore C/EBPα does not need to be able to bind DNA to mediate G1/S growth arrest. Furthermore, mutants of C/EBPα that did not contain the DNA binding domain of C/EBPα (bZIP region, see figure 1 and figure 2) were still capable of producing growth arrest (Liu *et al.*, 2002).

The amino terminus of C/EBPα has been demonstrated to be able to bind to E2F and inhibit its function (Nerlov, 2004). It is currently unknown how this E2F-binding domain is linked to C/EBPα-mediated growth arrest. In Wang *et al.*, (2001) they claim that amino acids 175-217 from C/EBPα is sufficient to mediate growth arrest. There are several reports in literature which claim that the p30 isoform of C/EBPα is unable to mediate growth arrest, even though the p30 isoform of C/EBPα contains amino acids 175-217 (Muller *et al.*, 2004; Nerlov, 2008). By this logic, the p30 isoform of C/EBPα should be able to meditate growth arrest if the 175-217 domain was the controlling factor, which it cannot. Since the p30 isoform is unable to mediate growth arrest, this suggests that there is something else which helps the p42 isoform differentially provide growth arrest from p30 that have not been discovered yet. Although, these conflicting reports could be due to differential cell lines being used or artifacts from over-expression of proteins, more research will have to be performed to resolve these inconsistencies.

In experiments which overexpress proteins which inhibit the function of Rb it was demonstrated that C/EBPα-mediated growth arrest was still possible (Harbour and Dean, 2000). In cell lines that do not express Rb it has been observed that C/EBPα is still able to mediate

growth arrest (Muller *et al.*, 2004). Although not directly addressed in the papers that observe this, other pocket proteins could be still expressed in these cell lines and these pocket proteins would still able to bind to E2F and/or C/EBPα. Triple negative cell lines for all three pocket proteins have been demonstrated to be unable to undergo C/EBPα-mediated growth arrest (Johnson, 2005). Potentially these other proteins bound to E2F could also be mediating the growth arrest state since C/EBPα is able to bind to E2F directly and therefore should be able to mediate the inhibition. These findings highlight that there is more to this mechanism than is currently known.

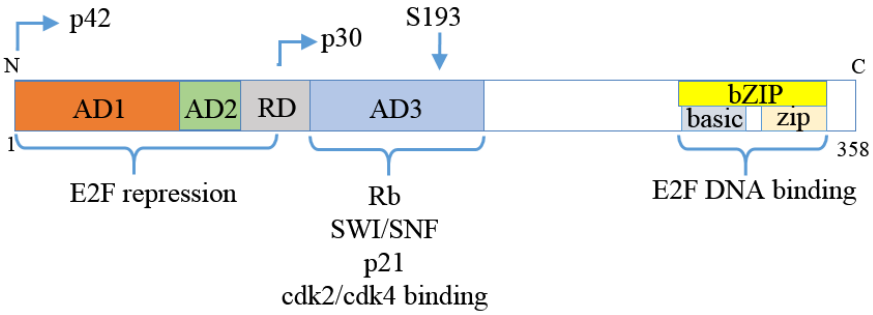

Figure 2 – The Growth Arrest Domain Specific Sites of C/EBPα

C/EBPα contains several important protein interaction domains that are involved in growth arrest. The amino terminus of C/EBPα is able to bind to various E2F proteins and repress their functions. The domain around S193, comprising most of activation domain 3 is able to directly bind to Rb, SWI/SNF, p21, cdk2 and cdk4. As well, differential phosphorylation of the S193 residue is able to influence the binding properties of these proteins to this region. The carboxy bZIP region is able to bind to cognate E2F binding sites and influence the transcription of E2F binding promoters.

2.3.1.2 The Cdk2/4 Model of C/EBPα-Mediated Growth Arrest

Cdk2 and cdk4 are cyclin dependent kinases (cdk) that are able to modulate the progression of the cell cycle by mediating the signals presented to them through cyclins in the cell. The cell cycle is partially driven by cdks which are able to phosphorylate Rb (Harbour and Dean, 2000). Different cyclin/cdk combinations phosphorylate Rb depending on the stage of the cell cycle, with early G1 engaging cdk4/cyclin D, late G1 engaging cdk2-cyclin E and S phase engaging cdk2-cyclin A (Harbour and Dean, 2000). It has been demonstrated that C/EBPα is able to bind to both of these cdks via the S193 region (see figure 2) and through this binding differentially modulate the activities of the proteins. Through the binding of the cdk proteins by C/EBPα it has been suggested that C/EBPα is able to prevent them from forming

cdk/cyclin complexes, which blocks cell cycle progression (Wang *et al.*, 2001). Furthermore, it has been suggested that by binding p21 as well, C/EBPα is able to further enhance the inhibition of cdk2 by p21 (Harris *et al.*, 2001; Johnson, 2005). Thus, C/EBPα is able to act as an inhibitor of cdk function. Furthermore, it has been demonstrated that by increasing the levels of cdk2 in the cell such that C/EBPα is unable to bind all the free cdk2, the cell is able to overcome C/EBPα-mediated growth arrest (Wang *et al.*, 2001).

As illustrated in figure 2, C/EBPα has been shown to be able to mediate growth arrest via a domain surrounding S193 (Harris *et al.*, 2001; Wang *et al.*, 2001). Work by Wang *et al.*, (2001) demonstrated that a C/EBPα minimal growth arrest domain consisted of amino acids 175-217 of rat C/EBPα. This minimal growth arrest domain was demonstrated to bind to cdk2 as well as cdk4, which was suggested to be driving the C/EBPα-mediated cellular growth arrest state. Alternatively, Harris *et al.*, (2001) argued that the binding site for cdk2 is located between amino acids 119 and 160 of human C/EBPα. While the Wang paper tested their results directly, the Harris paper made their conclusions via exclusionary methods wherein they found that a 1-119 fragment of C/EBPα did not interact with cdk2 and cause growth arrest whereas a fragment containing 1-160 of C/EBPα did. Since the Wang paper used rat C/EBPα and the Harris paper used human C/EBPα there will be some slight differences in the amino acids expressed in these areas and these two findings cannot be directly compared. They both demonstrate cdk2 binding in a region that is close to the S193 region that has been implicated in the E2F-Rb binding model of growth arrest. This intersection with the E2F pathway could provide some clues as to the interrelationship between these pathways.

2.3.1.3 The SWI/SNF model of C/EBPα-mediated growth arrest

The SWI/SNF complex is a group of proteins that interact with the chromatin remodeling complex inside the nucleus. Through this complex, nucleosomes are able to be remodeled through the interplay between the SWI/SNF, HDAC and HAT complexes, therefore upregulating or downregulating the expression of target genes (Harbour and Dean, 2000). Experiments have demonstrated that members of the SWI/SNF complex such as Brm are able to bind to C/EBPα and the binding of these members modulate the progression of the cell cycle. As illustrated in figure 2, around the S193 region of C/EBPα an area has been mapped in which

Brm binding has been demonstrated. The SWI/SNF complex has been shown to be required for C/EBPα-mediated growth arrest. Cells that are deficient in members of the SWI/SNF core complex are unable to undergo C/EBPα-mediated growth arrest (Muller *et al.*, 2004). Members of the SWI/SNF complex such as Brg1 and Brm are able to bind to Rb. This Rb binding could provide a link between the SWI/SNF model of C/EBPα-mediated growth arrest and the E2F-Rb-C/EBPα model (Harbour and Dean, 2000). As well, Brg1 and cyclin E have been demonstrated to interact, providing a link with the cdk-mediated pathway of growth arrest (Harbour and Dean, 2000). Taken together, the SWI/SNF complex's ability to interact with both members of the E2F-Rb pathway and the cdk-mediated pathway of C/EBPα-mediated growth arrest may provide the link between these three separate models.

2.3.1.4 The Combined Model of C/EBPα-Mediated Growth Arrest

There are several pieces of evidence presented in the previous three models that indicate that they may be linked. The p30 isoform of C/EBPα gives several insights into the linkage between the three presented models. The E2F-Rb-C/EBPα model suggests that the amino terminal binding of C/EBPα to E2F is important for the functioning of the growth inhibitory activity. Since the p42 and the p30 isoforms of C/EBPα both contain the Rb-binding domain, with the only difference being that p42 has an intact E2F-binding domain in the amino terminus, there must be another domain in C/EBPα which helps in this interaction. Both the SWI/SNF and the cdk2/4 models they are unable to directly address this issue since they both suggest that the C/EBPα-mediated growth inhibitory region is solely in the S193 region. Therefore, if the growth inhibitory region is solely in the S193 region, how is it that the p30 isoform of C/EBPα, which still contains this region is unable to mediate growth arrest? Although the E2F model relies directly upon the involvement of Rb protein, how is it that cell lines that do not contain Rb protein are still able to mediate growth arrest? These are unanswered questions in the literature, but by not considering each system to be separate but interacting with each other a clearer picture of C/EBPα-mediated growth arrest can be pieced together.

Experiments that looked at the growth arrest protein complex containing C/EBPα observed that this complex migrated in complexes with molecular weights ranging from 158-

300 kDa (Wang *et al.*, 2001). This suggests that there may be more proteins that have yet to be identified since only specific proteins were probed out of this complex.

The phosphorylation/dephosphorylation of Rb occurs during the cell cycle with hypophosphorylated Rb more common in differentiating cells and the hyperphosphorylated version being more active in proliferating cells (Harbour and Dean, 2000). One of the kinases driving the phosphorylation of Rb is cdk4 which is able to interact with and phosphorylate Rb protein (Wang *et al.*, 2001). The differential phosphorylation/dephosphorylation of this protein is able to interrupt or promote differential protein interactions of Rb. This function provides a link between the cdk model and the E2F model since increased phosphorylation of Rb has been demonstrated to interrupt the E2F-Rb complex and drive the release of free E2F in the cell.

The phosphorylation of the S193 residue also provides some linkages that integrate the cdk model and the SWI/SNF model (Wang *et al.*, 2001). When the C/EBPα S193 is dephosphorylated by protein phosphatase 2A (PP2A) it becomes unable to bind cdk2 as well as Brm and is unable to drive cellular growth arrest (Wang and Timchenko, 2005). The dephosphorylation of S193 by PP2A appears to be insulin dependent, further linking the role of C/EBPα-mediated growth arrest into a physiological state condition-dependent pathway. In this pathway, cells stimulated by insulin, instead of being driven into growth arrest are actually observed to accelerate the rate of growth, although this acceleration of growth appears to be Rb dependent (Wang and Timchenko, 2005).

While generally termed E2F-Rb complexes, there is evidence that other proteins can be bound to these complexes as well. While it has been established that C/EBPα is able to bind to the E2F-Rb complex, it has also been demonstrated that Brm, which is a member of the SWI/SNF complex, is able to bind to the E2F-Rb-C/EBPα complex (Harbour and Dean, 2000). Thus far it is unknown if C/EBPα mediates the binding of Brm to this complex or if E2F-Rb-Brm is found in non-C/EBPα expressing cells. It should be possible to discover the actual relationships between these three currently disjointed models and this information will be key to fully understanding how this important pathway functions inside the cell.

2.3.2 C/EBPα Functionality is Not Location Dependent

Through its bZIP domain and the V296 residue in particular, which resides near the amino terminus of the basic amino acid region of the bZIP (see figure 1), C/EBPα has been shown to bind to the α-satellite regions in the pericentromeric, heterochromatic regions near the centromeres of chromosomes (Liu *et al.*, 2007). By using a V296 mutant it was discovered that it was possible to disrupt the subnuclear localization of C/EBPα from the pericentromeric regions to a nuclear dispersed pattern. Through the testing of this V296A mutant it was possible to determine many of the localization-specific properties of C/EBPα.

Liu *et al.,* (2007) demonstrated that when C/EBPα bound to the α-satellite repeats it was less available for transcriptional activation. It was later demonstrated that Pit-1 was able to release C/EBPα from these pericentromeric regions and therefore allow C/EBPα to become more available to the cell. Therefore it was suggested that by binding to the α-satellite regions in pericentromeric domains, C/EBPα was being segregated from the rest of the nucleus and unable to participate in the various functions it is able to mediate. It was also demonstrated that a V296A mutant of C/EBPα that was dispersed throughout the nucleoplasm, was still able to mediate relocalization of proteins as well as activation or repression of target genes (Moazed *et al.,* 2011).

It has been shown that C/EBPα is able to mediate the relocalization of different proteins in the cell from their normal subnuclear compartments to compartments where C/EBPα resides. This relocalization activity was initially demonstrated for CREB-binding protein and TATA binding protein but since has also been demonstrated for TCERG1 (Banman *et al.,* 2010; Schaufele *et al.,* 2001). Since the pericentromeric regions are considered transcriptionally inactive, this represents another way for the nucleus to regulate various processes occurring inside the nucleus. By recruiting various proteins to the pericentromeric regions and regulating the subsequent release of these proteins, C/EBPα is able to affect the regulation of many other processes in the nucleus. It was demonstrated that when TCERG1 was recruited to the pericentromeric regions where C/EBPα resides it was then able to inhibit the transactivation potential of C/EBPα (Moazed *et al.,* 2011). This TCERG1 recruitment was not specific to the pericentromeric region as once again, the V296A mutant of C/EBPα was able to have its transactivation properties repressed by TCERG1, even though the C/EBPα was dispersed across the nucleus.

2.3.3 C/EBPα and its Role in Disease

One of the major diseases that C/EBPα has been implicated in is acute myeloid leukemia (AML). AML is a leukemia or blood cancer that begins in the bone marrow when myeloid cells are blocked from normal maturation into macrophages or granulocytes. This blockage in turn causes an uncontrolled proliferation of myeloid progenitor cells (Roe and Vakoc, 2014). This form of leukemia is the most common form of leukemia among adults, although it is rare among people under the age of 40. The prognosis of this disease is poor among adults with a 5 year survival rate of ~30% with better rates among younger individuals due to lower occurrences of chemotherapy-resistant strains (Roe and Vakoc, 2014).

Since C/EBPα is important in the development and differentiation of the progenitor myeloid cells, dysregulation of C/EBPα can drive a disease state transformation. It has been suggested that the two different isoforms of C/EBPα could play a role in the development of this disease. In mice studies it was discovered that knock out of the p42 isoform of C/EBPα leaving only expression of the p30 isoform, AML formation occurs with full penetrance (Kirstetter *et al.*, 2008). This suggests that the p42 isoform acts as a myeloid tumor suppressor. More specifically, there is likely a role of the two activation domains not included in the p30 isoform which are acting to prevent the onset of this disease. Interestingly, it was observed in most AML cases that one of the C/EBPα alleles had been mutated, mostly to a variant which causes a problem with the expression of the p42 variant, leaving only expression of p30 from that allele. That the mutation of one allele to only allow the expression of p30 correlated to a higher penetrance of AML suggests that the p30 isoform could potentially act as a dominant negative (Nerlov, 2004).

2.4 Transcription Elongation Regulator 1 (TCERG1)

Since its discovery as a coactivator of HIV-1 tat transcription in a TATA-box dependent manner, transcription elongation regulator 1 (TCERG1) has been found to be involved with splicing and transcription elongation in eukaryotic cells (Sune and Garcia-Blanco, 1999; Sune *et al.*, 1997). TCERG1 appears to be ubiquitously expressed across most, if not all human adult tissues, as would be expected for a gene which is highly involved in transcription and splicing regulation (Bohne *et al.*, 2000). TCERG1 is expressed as two isoforms with one being 1098 amino acids and the other being slightly shorter at 1077 amino acids. The shorter isoform contains a deletion of 21 amino acids between the STP domains and WW2 domains; currently it is unknown what effect this deletion has upon the protein as no known activities have been mapped to this region of TCERG1. As shown in figure 3, TCERG1 contains several important domains including three WW domains, six FF domains, a unique QA_{38} repeat domain, and a nuclear localization signal (NLS). Even though there have been six FF domains described in literature for TCERG1, according to the standards and consensus sequences published about FF domains, only the first five FF domains are actual FF domains (Bedford and Leder, 1999). FF domains are typically characterized as containing to highly conserved phenylalanine residues on each end of the domain. There are several other interesting but functionally-undescribed domains in TCERG1 such as the N-terminal polyproline region, the STP domain, and the KE domain. The STP domain is a domain that is rich in serine, threonine and proline residues whereas the KE domain is a lysine and glutamate rich region. The FF domains appear to mediate the targeting of TCERG1 to the periphery of nuclear speckles where it interacts with splicing and other transcription factors as well as targeting to the C-terminal domain of RNA polymerase II in a phospho-dependent manner (Carty *et al.*, 2000; Sanchez-Hernandez *et al.*, 2012; Smith *et al.*, 2004). The FF domains of TCERG1 have also been demonstrated to be the site of interaction with various other splicing factors inside the nucleus (Smith *et al.*, 2004). The WW domains function in protein recognition and interaction, similar to FF domains, and proline-rich proteins (of which C/EBPα is one) are common targets. Although not confirmed in humans, TCERG1 has been shown to be involved in the pathway controlling lifespan in C. *elegans* (Ghazi, 2013; Ghazi *et al.*, 2009; Keith and Ghazi, 2014). TCERG1 has also been implicated as a regulator of the apoptotic pathway (Wang *et al.*, 2000). Many of the functions

of TCERG1 are poorly described in literature and there is still much to be learned from this protein that seems to affect many important processes in the cell.

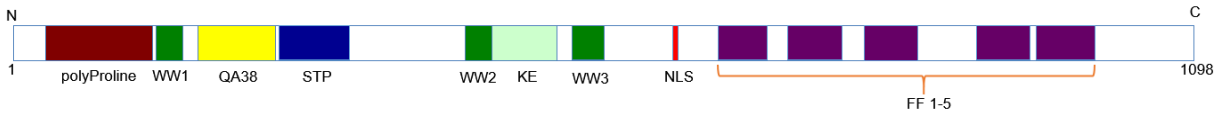

Figure 3 – Domain structure of TCERG1.

Schematic representation of the domain structure of TCERG1. The regions of TCERG are as follows (amino acids are indicated): Polyproline: 32-132, WW1: 137-164, QA38: 180-255, STP: 259-322, WW2: 433-457, KE: 458-516, WW3: 532-561, NLS: 626-630, FF1: 661-709, FF2: 727-776, FF3: 794-843, FF4: 898-949, FF5: 956-1007.

2.4.1 TCERG1 and its Role in Transcription and Splicing

TCERG1 is capable of binding residues of the RNA Pol II C-terminal domain (CTD) when phosphorylated (pCTD). Through the use of its FF4 and FF5 domains it has been demonstrated that TCERG1 specifically binds to the phosphorylated Ser2, Ser5, and Ser7 moieties of the RNA Pol II CTD (Liu *et al.*, 2013). Interestingly, for the high affinity binding of TCERG1 to the pCTD it requires that all three of these residues need to be phosphorylated. The phosphorylation of the CTD acts as indicators of the presence of active transcription and splicing occurring in individual RNA Pol II complexes. Since TCERG1 is characterized as a linker protein between transcription and splicing this finding would suggest there is a time in the elongation of transcription (see figure 1) in which all three serine residues of the CTD are phosphorylated. This would most likely be when the cell is recruiting the initial splicing factors to begin splicing of the nascent mRNA transcript or further into transcriptional elongation but likely before termination of transcription, unless for a short transcript. TCERG1 depletion has been demonstrated to inhibit proper splicing of mRNA transcripts, which would support this observation (Pearson *et al.*, 2008). Thus far it is known that TCERG1 is able to modify the pre-mRNA slicing for β-globin, β-tropomysin, Bcl-x, CD44 and fibronectin (Montes *et al.*, 2012). When TCERG1 is knocked down in the cell, it has been observed that the alternative splicing characteristics of many genes is affected (Pearson *et al.*, 2008).

The FF2 domain of TCERG1 is known to bind to the transcription factor DACH1 which helps DACH1 to repress gene function by binding to promoter regions of specific genes. Deleting FF2 has been shown to abolish this interaction and repression thus demonstrating

TCERG1's ability to act as a co-repressor (Zhou *et al.*, 2010). Thus TCERG1 has gene-specific regulatory roles in addition to its more general role in regulating transcriptional elongation.

2.4.2 TCERG1 Modulates the Rate of Transcription of RNAP II

Transcription by RNAP II does not occur at a constant rate. During several stages of transcription the rate of transcription may slow down or "stall" or it may increase. These events offer the cell and the emerging mRNA transcript the ability for differing regulatory events to occur such as alternative splicing.

TCERG1 has been linked to alternative splicing of mRNA. It has been demonstrated that overexpressing TCERG1 in a cell was able to mediate an increase in the elongation rate of RNAP II which would lead to the production of higher than normal levels of alternative splice products of the Bcl-x gene (Montes *et al.*, 2012). Conversely, through the binding of SF1 to the WW1 and WW2 domains of TCERG1, transcriptional repression of certain genes can be accomplished by the inhibition of RNAP II transcriptional elongation (Goldstrohm *et al.*, 2001). As well, TCERG1 has been demonstrated to decrease the rate of RNAP II elongation from the HIV-1 promoter (Sune *et al.*, 1997). Taken together, it appears that TCERG1 is both able to upregulate or downregulate transcription from different genes. Sumoylation of TCERG1 has also been implicated in its ability to affect transcriptional regulation of TCERG1 target genes (Sanchez-Alvarez *et al.*, 2010). Furthermore, it was discovered that TCERG1 was able to increase the phosphorylation rate of the Ser2 residues in the CTD of RNAP II. The overall increase of phosphorylated Ser2 residues in turn increases the transcription rate of RNAP II (Coiras *et al.*, 2013). Coiras *et al.*, (2013) demonstrated that when TCERG1 was knocked down that the phosphorylation state of the RNAP II CTD remained unaltered. However, when TCERG1 was overexpressed there was an accumulation of phosphorylated Ser2 on the CTD as well as a small increase of the total RNAP II within the nucleus. This increase in turn facilitated the increase of RNAP II transcription rate.

2.4.3 TCERG1 as an RS Domain Protein

It is well established in the literature that TCERG1 is a nuclear speckle protein. Sanchez-Alvarez *et al.,* (2006) determined initially that TCERG1 was resident in nuclear speckles. Furthermore, Sanchez-Hernandez *et al.,* (2012) refined this observation, demonstrating that TCERG1 is actually localized to the periphery of nuclear speckles via the FF4/5 domains of TCERG1. Interestingly, the FF4/5 domains of TCERG1 also appear to be the domains which bind to the phosphorylated CTD of RNAP II. It was also observed that the recruitment of TCERG1 to nuclear speckles was not dependent upon active transcription in the cell; when transcription was globally inhibited, TCERG1 was still localized to nuclear speckles (Sanchez-Alvarez *et al.*, 2006). One of the major characteristics of a nuclear speckle protein is that they contain a domain which is rich in arginine and serine (RS) amino acids. This RS-rich region serves to act as a nuclear speckle localization signal. Thus far nobody has specifically characterized TCERG1 as an RS domain protein even though it localizes to nuclear speckles.

Figure 4 is an illustration of TCERG1 RS amino acids across the FF domains as well as N and C terminal fragments. TCERG1 full length contains ~6% of each R and S amino acids with a slight enrichment being in the carboxy half of the protein which contains the FF domains. Examining the FF domains found in this area and the regions between each FF domain it is possible to see a pattern arise. For most of the domains a differential pattern of the R and S residues is observed, in that if one amino acid is enriched, the other is found at a lower level. The only two domains in which this pattern is not the case is FF4 and FF5; in these regions both the R and S residues are both enriched, suggesting that both FF4 and FF5 combine to form a nuclear speckle targeting RS domain. This conclusion would agree with the results found by Sanchez-Hernandez *et al.,* 2012 in which they discovered that the FF4 and FF5 domains are required to effectively target TCERG1 to the nuclear speckle compartment.

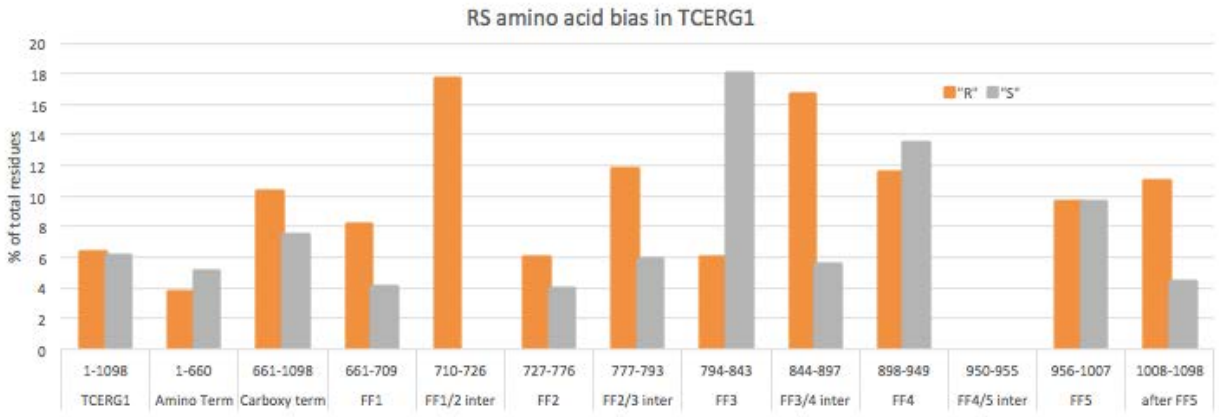

Figure 4 – The RS amino acid composition of TCERG1 suggests that FF4 and FF5 are acting as RS domains.

Graph demonstrating the % of total amino acids containing either arginine or serine residues for each of the domains examined. Each region is denoted by the specific amino acids tested along with the name of the region.

2.4.4 TCERG1 QA_{38} Domain

Very little is known about the role of the QA repeat domain in TCERG1, or QA domains in general; there is only one other described protein in humans, ZNF384/CIZ, which contains an identical albeit shorter QA_{14} repeat. In other described organisms there are two proteins in *Saccharomyces cerevisiae*, GAL11 and SSN6, and Zeste in *Drosophilia melanogaster* which contain significant stretches of QA repeats (Sune *et al.*, 1997). While currently there are no observed direct links between the QA domain and the polyQ repeat region found in several proteins, these proteins could potentially provide some important insights into the potential functions and implications of the QA domain of TCERG1 as they contain a much more developed literature base than TCERG1. Whereas for the QA domain of TCERG1 it has only been mentioned as a possible functional domain prior to this work, the polyQ domain has been extensively studied due to its intimate association with various diseases. Most of these polyQ containing proteins generate a disease state when their polyQ regions expand to a certain length such as huntingtin where Huntington's disease begins to manifest at Q_{39}. It is interesting that Huntington's disease manifests at 39 repeats and TCERG1 contains 38 QA repeats. Furthermore it has been demonstrated that if the TCERG1 QA repeat domain gets slightly larger it can contribute to disease state progression (see section 5.4)

It was been discovered that stretches of polyQ or polyA peptides have been known to form β-sheets (Bauer *et al.*, 2011). The formation of β-sheets has been linked to aggregation disorder diseases such as Alzheimer's, prion disease, Huntington's disease (HD) and similar. Aggregations of β-sheets in the cell are commonly termed amyloid plaques (Ramirez-Alvarado *et al.*, 2000). Since it has been shown that both stretches of polyQ peptides or polyA peptides are able to form β-sheets it can be hypothesized that the QA domain of TCERG1 could as well form β-sheets if the number of QA repeats is large enough. Interestingly for such a highly repetitive domain there are no known diseases directly associated with an expansion of QA repeats in TCERG1. The only polymorphisms of TCERG1 have been reported to be small expansions of ~9-12 base pairs (3-4 amino acids) (Andresen *et al.*, 2007; Chattopadhyay *et al.*, 2003). Even these very small changes to the QA repeat domain of TCERG1 have significant effects on the age of onset of HD; going from 38 repeats to 39.5 repeats decreased the age of onset of symptoms by about 15.6 years (Andresen *et al.*, 2007). This suggests that even minor gains in the size of the QA repeat length can bring about clinical consequences. Since the QA domain is highly conserved it might be subject to polymerase slippage and expansion, similar to other DNA sections which contain highly repeated segments. There have been no known cases of the QA domain of TCERG1 extending past a very small expansion as mentioned above. This suggests that even modest expansions in the QA repeat length of TCERG1 are destabilizing to the cell enough that any cells which contain this QA expansion are unable to survive; this suggests that the QA domain of TCERG1 is playing a much more integral role in the cell than we are currently aware of.

2.4.4.1 Evolution of the QA Domain of TCERG1

The QA domain of TCERG1 appears to be a recently evolved functional motif. As shown in figure 5, the QA domain is essentially absent in lower life forms with nematodes (C. *elegans*) having no repeats, and zebrafish (D. *rerio*), and fruit flies (D. *melanogaster*) both having a single QA. Even in more recent life forms such as chickens (G. *gallus*) the QA domain is almost absent with only two QA repeats. It isn't until bovines (B. *taurus*) that a significant repeat QA domain appears, with 20 copies. From here, the QA domain becomes longer in length with dogs (C. *lupus*) and rats (R. *norvegicus*) both having 29 copies, macaques

(M. *mulatta*) 37 copies, humans (H. *sapiens*) 38 copies, mouse (M. *musculus*) 39 copies, and chimpanzees (P. *troglodytes*) having 41 copies.

The implications of the expanded TCERG1 QA repeat are not immediately clear from examination of the data. Potentially, the expanded QA repeat is required in higher eukaryotes to mediate more complex splicing reactions. Of the data in figure 5 there is one interesting deviation from what would be the expected pattern if the QA length evolution was driven from mediating complex splicing reactions. In chickens, there is a distinct lack of QA repeats although this organism would be performing complex splicing reactions similar to the other higher eukaryotes on this list. The lack of TCERG1 containing a QA domain in chickens could be subject to further investigation to determine what, if any, implications the lack of a QA domain would have in this organism.

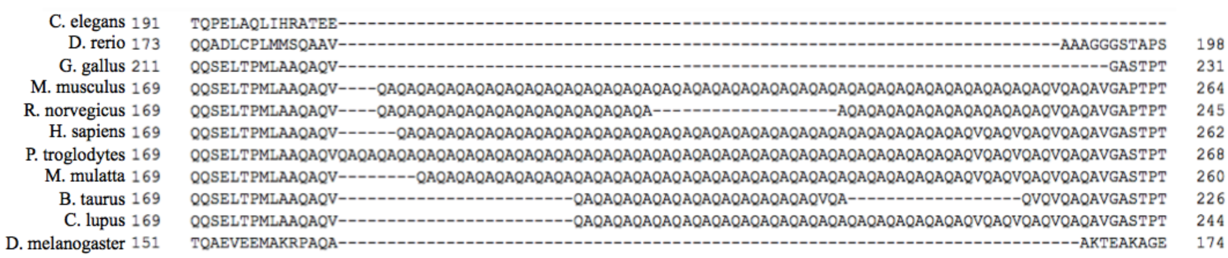

Figure 5 – The evolution of the TCERG1 QA domain.
Multiple sequence alignment of the TCERG1 QA domain across various species.

2.5 TCERG1 is an Inhibitor of C/EBPα

C/EBPα and TCERG1 are both important transcription factors on their own but in recent years the link between their properties has been explored. The initial link between TCERG1 and C/EBPα was reported in 2006 by McFie *et al.* in which they performed a two-hybrid bait and prey experiment wherein they looked for interactors of C/EBPα using a human liver cDNA library. The strongest interactor pulled out coded for an amino terminal fragment of TCERG1 which contained amino acids 89-480. Although TCERG1 was not demonstrated to bind to the promoter directly with C/EBPα it was shown to inhibit C/EBPα-mediated transactivation. Furthermore, TCERG1 was able to inhibit C/EBPα-mediated G1/S growth arrest using full length C/EBPα. However, when using the minimal 175-217 C/EBPα growth arrest domain, TCERG1 was unable to prevent growth arrest. This suggested that there is a

second site in C/EBPα with which TCERG1 is able to interact. This first paper was able to establish the interaction between C/EBPα and TCERG1 but there were still many unanswered questions.

The second paper, published in 2010 by Banman *et al.* attempted to further refine several of the unanswered points between the interaction between C/EBPα and TCERG1. Prior to this paper, it had been known that C/EBPα was able to mediate the relocalization of several proteins inside the nucleus. This paper demonstrated that although TCERG1 normally resided inside nuclear speckles and C/EBPα resided in pericentromeric regions, that when both proteins were expressed in a cell, TCERG1 was recruited to pericentromeric regions. Furthermore, the interaction domain of TCERG1 with C/EBPα which mediates TCERG1 relocalization as well as growth arrest inhibition was determined to reside in the amino terminus of TCERG1, within amino acids 32-668.

Previously, it was suspected that the movement of TCERG1 to pericentromeric regions was a mechanism to inhibit C/EBPα by keeping it sequestered inside the pericentromeric regions. The paper by Moazed *et al.,* (2011) demonstrated that this was not the case. Using a V296A mutant of C/EBPα, which was unable to bind the α-satellite repeats and therefore was dispersed throughout the nucleus, they were able to demonstrate that TCERG1 was able to effectively inhibit this mutant C/EBPα's growth arrest properties as well as inhibit C/EBPα-mediated transactivation. This paper also established that the relocalization of TCERG1 from nuclear speckles to pericentromeric regions in the presence of C/EBPα was not an artifact of protein overexpression as endogenous TCERG1 was also overexpressed.

This thesis aims to build upon the findings presented in all three of the foundational papers presented here. In initial experiments, our collaborator, Dr. Nick Timchenko demonstrated that by deleting the QA domain in TCERG1 that TCERG1 was no longer able to inhibit C/EBPα-mediated growth arrest. This thesis builds upon these findings to determine if the QA domain is involved in other interactions between TCERG1 and C/EBPα.

2.6 Hypothesis

The QA repeat domain of TCERG1 is required for relocalization by C/EBPα and inhibitory activity against C/EBPα.

2.7 Rationale and Objectives

1. To investigate the role of the QA domain in the context of TCERG1, and the length requirement, in the relocalization and transcriptional inhibitory activities of TCERG1.

Data from Dr. Nick Timchenko's lab suggested that the QA domain of TCERG1 is involved in some of the activities involved in the interaction between TCERG1 and C/EBPα. Therefore, it should be possible to test the various interaction properties of TCERG1 and C/EBPα such as relocalization, transcriptional inhibition of C/EBPα and physical interactions using QA deletion mutants of TCERG1 to attempt to determine the role of the QA domain on these activities.

2. To investigate whether the QA domain, isolated or fused to a marker protein, is sufficient for the relocalization and inhibitory activity of TCERG1, and interaction with C/EBPα.

By using the QA domain fused to a fluorescent marker it should be possible to examine if the minimal QA domain has any of the similar activities demonstrated from previous experiments in objective 1.

3. Materials and Methods

3.1 Reagents

All reagents used were of analytical grade or higher. Names of reagents and suppliers are listed in Table 1. Addresses for each supplier are subsequently listed in Table 2.

Table 1: List of Reagents and Suppliers

Reagent	Supplier
Common Reagents	
100 Base Pair (N3231), 1 Kilobase Pair (N3232) DNA Ladder	NEB
Agarose	Sigma-Aldrich
Q5 High Fidelity Polymerase	NEB
Taq DNA Polymerase with ThermoPol Buffer	NEB
Gel Loading Dye Blue (6X)	NEB
Polyethylenimine "MAX" MW 40,000	Polysciences
Taq DNA Ligase	NEB
GelRed Nucleic Acid Gel Stain	VWR
Restriction Enzymes, Various	NEB
LipoD293 Transfection Reagent	FroggaBio
Bacterial Cell Culture Reagents	
Ampicillin	BioShop
Kanamycin	Fisher
NEB5α Ultracompetent *Escherichia coli*	NEB
DH5α Subcloning Efficiency *Escherichia coli*	Invitrogen
Tryptone	Sigma-Aldrich
NaCl	EMD
Bacto Yeast	BD Biosciences
Commercial Kits	
GeneJET Plasmid Miniprep Kit	Fisher
GeneJET Gel Extraction kit	Fisher
Qiagen Plasmid Plus Midiprep/Maxiprep Kit	Qiagen
Dual Luciferase Assay System	Promega

Cell Culture Reagents	
Dulbecco's Modified Eagles Medium – High Glucose with L-Glutamine and Sodium Pyruvate	Fisher
Fetal Bovine Serum (Canadian Origin)	Sigma-Aldrich
Trypsin 0.25%	Fisher

Reagents for Confocal Imaging	
ProLong Diamond Antifade Mountant	Invitrogen
Paraformaldehyde	Sigma-Aldrich

Protein Analysis Reagents – Western and Co-IP	
30% Acrylamide/Bis Solution (29:1)	Bio-Rad
Nitrocellulose Membrane	Bio-Rad
HyBlot CL Film (5X7)	Denville Scientific
PageRuler Prestained Protein ladder (PI26616)	Thermo Scientific
IGEPAL CA-630	Sigma-Aldrich
Glycine	Sigma-Aldrich
cOmplete Ultra, EDTA-free Protease Inhibitor Tablets	Roche

Table 2: List of Names and Addresses of Reagent Suppliers

Supplier	Address
BD Biosciences	Mississauga, Ontario, Canada
BioShop	Burlington, Ontario, Canada
Bio-Rad	Hercules, California, USA
Clontech	Mountain View, California, USA
Denville Scientific	Metuchen, New Jersey, USA
EMD	Madison, Wisconsin, USA
Fisher Scientific	Ottawa, Ontario, Canada
FroggaBio	North York, Ontario, Canada
Invitrogen	Burlington, Ontario, Canada
NEB	Mississauga, Ontario, Canada
Polysciences	Warrington, Pennsylvania, USA
Promega	Madison, Wisconsin, USA
Qiagen	Mississauga, Ontario, Canada
Roche	Mississauga, Ontario, Canada
Santa Cruz	Santa Cruz, California, USA
VWR	Mississauga, Ontario, Canada

3.2 Bacterial Strains and Media Preparations

E. coli DH5α (Invitrogen) and NEB5α (NEB) ultracompetant cells were used for the amplification of plasmid DNA. SOC (NEB) medium containing appropriate amounts of antibiotic was used to transform DNA into NEB5α cells, as per manufacturer's directions for cloning applications. For routine subcloning, DH5α competent cells were used with Luria-Bertani (LB) broth (10 g Bacto-tryptone, 5 g Bacto-yeast and 10 g NaCl to 1 L total volume in ddH_2O), as per manufacturer's directions.

LB agar plates were prepared using prepared LB broth with 15 g bacto-agar. Once mixed, the broth was subsequently autoclaved for 20 minutes at 15 lb/sq. in. Once cooled to ~50°C kanamycin or ampicillin was added to a final concentration of 50 µg/mL and 100 µg/mL, respectively. Twenty mL of broth was added to each 100 mm plate and then allowed to cool and solidify. The plates were subsequently stored at 4°C until required.

Prior to amplification, a single *E. coli* colony was picked from the transformation plate and grown in liquid LB broth. LB broth containing appropriate amounts of antibiotic was used to further propagate *E. coli* bacterial cultures. LB agar plates were grown at 37°C. Liquid LB cultures were grown overnight at 37°C with 215 RPM agitation.

3.3 Molecular Cloning

3.3.1 Transformation of Competent Bacterial Cells

Transformation of NEB5α and DH5α *E. coli* cells were performed according to manufacturer's directions. One ng of plasmid DNA was introduced to 50 µL of the thawed cells on ice. After a 30 minute incubation on ice the cells were incubated at 42°C for 30 seconds, then further incubated on ice for five minutes. Nine hundred fifty µL of growth media was added to the mixture and then further incubated with 225 RPM agitation at 37°C for 1 hour prior to plating 100-200 µL of transformation mixture onto LB agar plates containing appropriate antibiotics.

3.3.2 Plasmid DNA Preparations

Plasmid DNA was isolated from confluent overnight LB broth bacterial cultures using GeneJET plasmid mini prep kits (Fisher Scientific) according to manufacturer's directions using 5 mL of overnight cultures. When larger quantities of plasmid DNA were required, a Qiagen plasmid plus midi or maxi prep kit (Qiagen) was used according to manufacturer's directions. The volume of bacteria grown overnight in LB broth was 25 mL for a midi prep and 100 mL for a maxi prep using appropriate antibiotic.

3.3.3 Plasmids

Table 3: Plasmids Created for This Project

Plasmid	Source of backbone	Source of Insert/Change
TCERG1		
mCherry-C1-TCERG1-WT	mCherry-C1	mCherry-TCERG1
mCherry-C1-TCERG1-QA_{20}	mCherry-C1	mCherry-TCERG1
mCherry-C1-TCERG1-QA_{10}	mCherry-C1	mCherry-TCERG1
mCherry-C1-TCERG1-ΔQA	mCherry-C1	mCherry-TCERG1
mOrange2-C1-TCERG1-WT	pmOrange2-C1	BOST7-TCERG1-WT
mOrange2-C1-TCERG1-ΔQA	pmOrange2-C1	BOST7-TCERG1-ΔQA
mNeptune2-C1-TCERG1-WT	pmNeptune2-C1	BOST7-TCERG1-WT
mNeptune2-C1-TCERG1-ΔQA	pmNeptune2-C1	BOST7-TCERG1-ΔQA
BOST7-TCERG1-QA_{17}	BOST7-TCERG1-WT	SDM
BOST7-TCERG1-QA_{11}	BOST7-TCERG1-WT	SDM
BOST7-TCERG1-ΔQA	BOST7-TCERG1-WT	SDM
BOST7-TCERG1-Δ1-611	BOST7-MCS	BOST7-TCERG1-WT
C/EBPα		
FLAG-C1-C/EBPα	eGFP-C1	FLAG-N1-C/EBPα
HA-C1-C/EBPα	HA3-C1	FLAG-N1-C/EBPα
eGFP-C1-C/EBPα	eGFP-C1	FLAG-N1-C/EBPα
QA		
mCherry-C1-QA	mCherry-C1	BOST7-TCERG1-WT
mCherry-C1-3XNLS-QA	mCherry-C1-QA	Synthetic SV40 NLS
mCherry-C1-3XNLS-FLAG-QA	mCherry-C1-3XNLS-QA	FLAG-C1
mCherry-N3-QA	mCherry-N3	BOST7-TCERG1-WT
mCherry-N3-QA-NLS	mCherry-N3-QA	Synthetic SV40 NLS

Expression vectors		
FLAG-C1	eGFP-C1	FLAG-rCEBPa-pcDNA3
T7-C1	eGFP-C1	pEF-BOST7
mCherry-N1	eGFp-N1	mCherry-C1
mCherry-N3	mCherry-N1	mCherry-N1
HA3-C1	eGFP-C1	pHA3
pGL4.71-CMV	pGL4.71	eGFP-C1

Table 4: Plasmid Backbones Used

Plasmid	**Source**
mCherry-C1	Clontech
eGFP-C1	Clontech
mOrange2-C1	Clontech
mNeptune2-C1	Addgene
pEF-BOST7	Banman *et al.*, 2010
pGL4.71	Promega

Table 5: Pre-existing Plasmids Used

Plasmid	**Source**
pEF-BOST7-TCERG1-WT	Banman *et al.*, 2010
eGFP-N1-C/EBPα	Banman *et al.*, 2010
FLAG-N1-C/EBPα	Banman *et al.*, 2010
BOST7-TCERG1-21-680	Sheng-Pin Hsiao (Roesler lab PDF)
pEF-BOST7-MCS	Sheng-Pin Hsiao (Roesler lab PDF)
pGL4.11-68Fx4-luc	Banman *et al.*, 2010

3.3.3.1 Restriction Digest of Plasmid DNA

All restriction enzymes were purchased from New England Biolabs (NEB). In a total volume of 15 µL, 1 µg of plasmid DNA, 1 µL of each restriction enzyme, and 1.5 µL of appropriate 10X buffer were mixed and incubated at 37°C for 1-2 h using ddH2O to make up the volume difference.

3.3.3.2 DNA Fragment Isolation and Verification

To facilitate visualization and isolation of plasmid DNA and PCR amplified DNA, DNA was subjected to TAE agarose electrophoresis. The DNA of interest was mixed with an appropriate amount of 6X gel loading dye (NEB) prior to loading into the wells of a 1% agarose gel containing 1X TAE (40 mM Tris, 20 mM acetic acid, 1 mM EDTA, pH 8.0) and 2 µL/50 mL GelRed (VWR) suspended in 1X TAE buffer. As required, a 1 kb or 100 bp DNA ladder (NEB) was also electrophoresed to facilitate identification of lengths of DNA fragements. The gel was electrophoresed at 90V until the dye front migrated an appropriate distance. The gel was visualized using a 305 nm UV source and imaged using a UVP Gel documentation system. When required, DNA fragments were excised using a GeneJET gel extraction kit (Fisher).

3.3.3.3 DNA Sequencing

DNA sequencing was performed using automated Sanger sequencing by commercial sequencing service providers NRC plant biotechnology institute (Saskatoon, SK) or Operon-Eurofins genomics (Huntsville, Alabama).

3.3.4 DNA Cloning Techniques

Depending upon the initial source of DNA, several methods were used to amplify and clone DNA fragments into subsequent plasmid backbones.

3.3.4.1 Gibson Assembly Plasmid Engineering

Gibson assembly cloning was performed based upon the original protocol described by Gibson *et al.,* (2009). Single strand DNA oligonucleotide primers were designed using Snapgene software (Chicago, IL) to amplify the DNA of interest from the source DNA. Each primer was designed to contain an overhang on the 5’ end of the primer which is complementary to the 3’ end of linearized plasmid backbone to be inserted. The overhang was

engineered to be either 25 base pairs or have a 60°C anneal temperature, whichever being greater. The insert fragment was PCR amplified using 12.5 µL of 2X master mix Q5 polymerase (NEB) using 1.25 µL of each 10 µM primer, 1 ng of template DNA and ddH2O to a final volume of 25 µL. In order to provide linear backbone plasmid DNA, the template was either PCR amplified using appropriate primers to create complementary overhangs for the insert or was linearized using restriction digestion (see section 3.3.3.1). The linearized backbone plasmid and amplified insert DNA were then electrophoresed on a 1% TAE agarose gel and appropriate fragments were isolated using a GeneJET gel extraction kit (Fisher Scientific). One µL of extracted vector and 4 µL of extracted insert were combined with 15 µL 1.33X Gibson master mix (0.08 U T5 exonuclease, 0.5 U Phusion DNA polymerase, 80 U TAQ DNA ligase, 5% PEG-8000, 100 mM Tris-HCl pH 7.5, 10 mM MgCl2, 10 mM DTT, 0.2 mM dNTP mixture, 1 mM NAD) and incubated at 50°C for 1 hour in a PCR thermocycler. Two µL of incubated mix were transformed into NEB5α (NEB) ultracompetent cells. Two hundred µL of bacteria was plated onto LB agar plates containing the appropriate antibiotic for the vector of interest and subsequently grown overnight. Isolated colonies were then picked from the LB agar plates and grown overnight in LB broth for mini prep to further test via restriction digestion and/or DNA sequencing.

3.3.4.2 Restriction Site Plasmid Engineering

Purified DNA obtained by restriction digestion (see section 3.3.3.1) and purified by gel extraction (see section 3.3.3.2) was incubated together with T4 DNA ligase (NEB) in a 6:1 insert to backbone ratio for 2 hours at room tempertaure with appropriate T4 DNA ligase buffer. Two µL of the mix was subsequently transformed into NEB5α cells. Transformed cells were plated onto LB agarose plates containing appropriate antibiotics.

3.3.4.3 cDNA Amplification and Cloning from HEK293T RNA

A confluent plate of HEK293T cells was processed using a Qiagen RNeasy mini kit (Qiagen) to obtain RNA. The obtained RNA was then processed into cDNA using Thermo-

Script RT-PCR system (Invitrogen). The fragment of interest was amplified using primers designed specific to the gene of interest from the cDNA. The insert was cloned into a vector backbone using the techniques described above in sections 3.3.4.1 and 3.3.4.2.

3.4 Mammalian Cell Culture

Mammalian tissue culture was performed inside a biosafety level 2 cabinet. Human embryonic kidney cells (HEK293T) and green African monkey kidney cells (COS7) were passaged by washing the cell layer with 1X PBS followed by addition of 1 mL of 0.25% trypsin derived from porcine pancreas. The cells were incubated at 37°C, 5% CO_2 for 5-15 minutes to allow the cell layer to detach from the plate. Appropriate amounts of pre-warmed 37°C complete medium (high glucose DMEM with sodium pyruvate, L-glutamine and 10% FBS) were added to the trypsinized cells to obtain the required confluency for the downstream application of the cells as well as to inactivate the trypsin. The cells were resuspended in various amounts of medium depending upon the vessel downstream used to culture them. One hundred mm plates use 10 mL of medium, 6 well plates use 2 mL/well, and 12 well plates use 1 mL/well.

3.4.1 Transient Cell Transfection

Transient protein overexpression was obtained through the transient transfection of HEK293T or COS7 cells. Polyethylenimine "MAX" (PEI, Polysciences) and LipoD293 (FroggaBio) were used for the transfection of mammalian cell culture.

3.4.1.1 PEI MAX Transfection

A 1:3 ratio of DNA to PEI was used for the transfection of HEK293T and COS7 and the amount of DNA used varied depending on the size of the plate/well being used. For 100 mm plates 7 µg of DNA was used, whereas 2.5 and 1 µg of DNA was used for 6 well and 12 well plates, respectively. The total amount of DNA used was kept constant. For transfections that required lower amounts of experimental plasmid DNA to be used, pTZ19R (19R) plasmid

DNA was used to make up the remaining amounts. The DNA and PEI were mixed together in a 1:3 ratio; for example, if 1 µg of DNA was used, 3 µL of PEI was added. The DNA/PEI mixture was allowed to equilibrate at RT for 10 minutes for the complexes to be formed. Subsequently, the mixture was added drop wise into the media of plates/wells of ~70-80% confluent cells. The plates were then incubated for 4-5 hours at 37°C, 5% CO_2 after which the medium was aspirated off and fresh pre-warmed medium was added to the cells. The cells were further incubated for a further 24-48 hours depending on the experiment.

3.4.1.2 LipoD293 Transfection

Cells were transfected according to manufacturer's directions using pTZ19R plasmid as filler DNA to ensure DNA amounts stay the same across transfection conditions.

3.4.2 Growth Arrest Assay

Growth arrest experiments were performed using COS7 and HEK293T cells transiently transfected (see section 3.4.1). Cells were transfected on day 1, then 18-20 hours later they were passaged to ~20-30% confluency via trypsinization using vigorous tituration to break up cell clusters. Plates were observed on day 3 and assessed for numbers of cells growing in clusters or single cells on the plates.

3.5 Mammalian Cell Culture Protein Expression Determinations

3.5.1 Preparation of Protein Extracts From Cultured Cells

Cultured cells were washed with room temperature PBS. Three hundred to 500 µL of lysis buffer was added to the washed cell layer for 100 mm plates; for general use the cells were lysed in 0.5% CHAPS in PBS although depending upon the downstream application the lysis buffer could vary (see sections 3.5.5 and 3.5.6). The cells were then incubated at 4°C for 15-30 minutes with gentle agitation. After incubation the cell layer was gently scraped using a cell lifter and then collected in a suitable tube. The cells were further subjected to fine needle

aspiration using a 25 gauge needle to disrupt the cell membranes. The tubes were centrifuged at 14,000 X g for 30 minutes to pellet debris and the supernatant was collected for subsequent applications.

3.5.2 Protein Quantification of Cellular Extracts

The supernatant from extract prepared as described in section 3.5.1 were subjected to the Bradford assay of protein detection using the BioRad protein assay system. Depending upon the concentration of protein in the initial lysate, an appropriate amount of lysate was added to the wells of a 96 well plate to obtain a final protein concentration between the ranges of 1 µg/µL and 8 µg/µL. One hundred µL of protein assay reagent was added to each well and gently mixed. The color change was read on a Molecular Devices 96 well microplate reader at 595 nm. The absorbance values were compared to a reference curve created from known BSA protein standards to obtain the final concentration of protein in the sample, and adjusted for the amount loaded into the wells.

3.5.3 SDS Polyacrylamide Gel Electrophoresis

Electrophoresis was performed using two BioRad systems, the Bio-Rad Mini Protean II Apparatus (Bio-Rad) for mini and the Bio-Rad Criterion system for midi format gels. The gels used for the mini system were handcast 4%/8% stacking/resolving SDS polyacrylamide gels (8% N,N'-methylene-bis-acrylamide (29:1), 390 mM Tris-HCl, pH 8.8, 0.1% SDS, 0.001% ammonium persulfate, and 0.0006% N,N,N',N'-tetramethylethanediamine, with a 4% stacking gel [4% N,N'-methylene-bis-acrylamide (29:1), 130 mM Tris-HCl, pH 6.8, 0.1% SDS, 0.001% ammonium persulfate, and 0.001% N,N,N',N'-tetramethylethanediamine]). The gels used for the midi system were precast 4-15% gradient gels from BioRad. PageRuler 10-170 kDa (Thermo) protein markers were used to determine the approximate molecular weight of detected proteins. The protein lysis samples were incubated at 95°C in SDS loading buffer (50 mM Tris, pH 6.8, 2% SDS, 10% glycerol, 0.02% bromophenol blue, and 1% β-mercaptoethanol) for 5 minutes prior to loading on the gel to denature samples. The protein samples were then loaded onto the gel and electrophoresed at 110 V (mini format) or 200 V

(midi format) until the dye front reached the bottom of the gel in 1X SDS running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS, pH 8.3).

3.5.4 Western Blotting

After electrophoresis, the SDS-PAGE gel was released from the running plates and incubated at RT in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol) for 10-20 minutes prior to transfer. The wet transfer method was used to transfer the proteins from the gel to the nitrocellulose membrane. The transfers were performed in a Bio-Rad Mini Trans-Blot transfer cell (mini format) and a Bio-Rad Criterion Blotter (midi format, plate electrodes) using transfer buffer at 70 V for 1 or 2 hours (midi and mini, respectively) or 25 V overnight at 4°C.

After wet transfer, the membrane was blocked in 5% milk in PBST (1X PBS, 0.05% Tween-20) for 1 hour at RT with gentle agitation. The membrane was subsequently probed with primary antibody diluted in 5% milk in PBST at RT for two hours with gentle agitation or overnight at 4°C (see table 6 for antibody dilutions used). After primary antibody the membrane was washed 3X with PBST for 10 minutes/wash. If required, a secondary HRP-conjugated antibody incubation was performed similar to the primary antibody followed by a wash step. The membrane was then developed using HyBlot® chemiluminescent film (Harvard Biosciences) and SuperSignal® West Pico Chemiluminescent Substrate (Thermo). For a full list of antibodies used and dilutions see Table 6.

Table 6: List of Antibodies Used

Each antibody was diluted as appropriate for its application. Catalog numbers and suppliers are also indicated. WB indicates Western Blot analysis; IF indicates immunofluorescence; Co-IP indicates Co-immunoprecipitation analysis.

Name	Dilution	Application	Supplier
Alexa Fluor® 405 Goat Anti-rabbit IgG (H+L) (A31556)	1:1000	IF	Invitrogen
Alexa Fluor® 405 Goat Anti-mouse IgG (H+L) (A31553)	1:1000	IF	Invitrogen
Alexa Fluor® 488 Goat Anti-Mouse IgG (H+L) *highly cross-adsorbed* (A11029)	1:1000-1:4000	IF	Invitrogen
Alexa Fluor® 594 Goat Anti-Mouse IgG (H+L) *highly cross-adsorbed* (A11032)	1:1000-1:4000	IF	Invitrogen
Alexa Fluor® 647 Goat Anti-Rabbit IgG (H+L) *highly cross-adsorbed* (A21245)	1:1000-1:4000	IF	Invitrogen
Mouse anti-SC35 (ab11826)	1:1000	IF	Abcam
Goat Anti-Rabbit IgG-HRP (sc-2004)	1:10,000	WB	Santa Cruz
Rabbit anti-C/EBPα (14aa) (sc-61)	1:10,000	WB	Santa Cruz
Rabbit anti-FLAG magnetic beads	10 uL/rxn	Co-IP	Sigma-Aldrich
Rabbit anti-HA HRP conjugated (A190-108P)	1:10,000	WB	Bethyl Labs
Rabbit anti-T7 HRP conjugated (A190-117P)	1:10,000	WB	Bethyl Labs

3.5.5 Co-Immunoprecipitation

Tissue culture cells were cultured in High glucose DMEM supplemented with 10% FBS in 100 mm dishes. Cells were transfected using 42 µg Polyethylenimine "MAX" MW 40,000 (Polysciences, Warrington, PA) in 100 mm culture dishes for HEK293T cells. The amount of DNA transfected was 7 µg total with 5 µg of BOST7-TCERG1 plasmid transfected along with 2 µg of FLAG-N1-C/EBPα or pTZ19R. The PEI, DNA and 308 µL of 150 mM NaCl were incubated at RT for 20 min and then added to each plate containing 10 mL complete media. Following incubation, the mixture was added to the cell culture well containing media and incubated at 37°C for 4-5 hours. Subsequently, the media was replaced with fresh complete media and cells were incubated for 24-48 hours depending upon the construct transfected. Cells were harvested and lysed at 4°C. The lysis/binding buffer used contained 25 mM HEPES pH 8.0, 0.5% Nonidet P-40, 200 mM KCl, 1 mM EDTA, 1 mM PMSF, and 1X cOmplete® protease inhibitor (Roche). Crude lysate was subjected to fine needle aspiration and then

incubated under gentle agitation for 1 hour, followed by centrifugation at 10,000 X g for 30 min. Input samples were collected at this time. The supernatant was then pre-cleared using 50 uL Sepharose beads (Sigma-Aldrich) for 1 hour under gentle agitation, then centrifuged for 20 min at 10,000 X g. Supernatant was incubated using 10 uL anti-FLAG magnetic beads (Sigma-Aldrich) for 2 hours at 4°C. Flow-through samples were collected and the bead fractions were washed 5X using lysis buffer. All fractions were incubated at 95°C in SDS-PAGE loading buffer (10% glycerol, 0.1% bromophenol blue, 2% SDS, 100 mM DTT, 50 mM Tris-Cl, pH 6.8) for 5 minutes. Samples were loaded onto a 4%/8% SDS-PAGE acrylamide gel and electrophoresed at 80 V until the dye front entered the resolving gel, then at 100V until the dye front reached bottom of the gel. Gel transfer was performed using a wet transfer method at 70 V for 110 mins onto nitrocellulose membrane (see section 3.5.4). The membrane was processed using standard western blotting technique (see section 3.5.4) using 1:10,000 rabbit anti-T7 HRP antibody (Bethyl) or 1:10,000 Rabbit anti-C/EBPα (14aa) and subsequently 1:10,000 Goat anti-Rabbit HRP in 5% milk in PBST.

3.5.6 Luciferase Assay

Reporter gene assays were performed in 12-well culture dishes with three biological replicates for each condition assayed. COS7 cells were transiently transfected using LipoD293 (see section 3.4.1.2) using 0.75 µg of DNA per well using pTZ19R as filler DNA to obtain equal amounts of DNA transfected per condition as per manufacturer's directions. The amounts of each individual plasmid transfected were as follows (per well of a 12-well plate):

.077 µg -68Fx4-luc
1.67 pg pGL4.71-CMV
0.047 µg C/EBPα expression plasmid*
0.227 µg BOST7-TCERG1 expression plasmid*
to 0.75 µg pTZ19R

*(exact plasmid varied from experiment to experiment, see figure legend of each experiment for details)

Forty-eight hours after transfection, the cells were washed once with PBS. Two hundred-fifty µL of 1X passive lysis buffer (Promega) was added to the cells. The plates were subsequently incubated at RT for 15 minutes on a rocking platform. Twenty µL of the resulting lysate was assayed using the Promega dual luciferase kit using 50 µL for each of the luciferase reagent and "Stop and Glo®" reagent to measure firefly and renila luciferase, respectively. The readings were performed on a Promega Gloxmax® luminometer.

3.6 Laser Scanning Confocal Microscopy

3.6.1 Transient Transfection and Microscope Slide Preparation

Tissue culture slides and immunostaining were prepared similar to the method presented in Moazed *et al.,* 2011. Cells were transfected using one of the methods presented in section 3.4.1. Cells used for microscope applications were grown on 22 mm coverslips in 35 mm culture dishes. Protein expression was allowed to proceed for 24-48 hours depending upon the expression plasmid transfected, at which time the coverslips containing the cells were harvested and fixed.

3.6.2 Preparation of Cells for Immunostaining

Cell slide fixing and immunostaining was performed according to the immunostaining protocol as described by Moazed *et al.,* (2011). The amount and type of antibody used for each experiment varied (See table 6 for dilution ranges). Cells were fixed by washing once with PBS and then incubated in 4% PFA in PBS for 30 minutes at RT. The cells were then washed 3X with PBS. At this point if the cells were not to be immunostained they were mounted to glass slides using Prolong® Diamond antifade reagent (Life Technologies). For cells which required immunostaining, the cells were further incubated at RT for 10 minutes in 0.15% Triton X-100 in PBS. The cells were then incubated in 3% BSA in PBS for 45 minutes at RT. After incubation, the BSA would be aspirated off and then 150 µL of primary antibody diluted in 3% BSA in PBS was added directly to each coverslip and incubated at RT for 1 hour. Following incubation, the coverslips were washed 3X using PBS, followed by adding 150 µL of secondary

antibody diluted in 3% BSA in PBS to each coverslip and incubated at RT for 45 minutes. Following incubation, the coverslips were washed 3X with PBS and then removed from the dishes and allowed to dry. Following drying, the cells were affixed to glass slides using Prolong® Diamond antifade reagent (Life Technologies). The slides were stored in the dark at RT and after 24 hours the slides were cleaned gently with 70% ethanol and stored at 4°C until required.

3.6.3 Image Acquisition and Manipulation

Imaging was performed using a Leica SP5 confocal microscope at the Western College of Veterinary medicine, or a Zeiss LSM 700 confocal microscope provided by Dr. Deborah Anderson and Dr. Erique Lukong in the cancer research cluster in the College of Medicine at the University of Saskatchewan. Images were obtained using Zen black edition (Zeiss) and further manipulated using FIJI image editing software (Fiji.sc).

4. Results

4.1 Role of the QA domain of TCERG1 in inhibiting C/EBPα-induced growth arrest

McFie *et al*., (2006) and Banman *et al.,* (2010) showed that TCERG1 inhibited C/EBPα-mediated growth arrest. While they were able to demonstrate that the inhibitory region was somewhere within amino acids 32-668 of TCERG1 the exact domain mediating this inihibition was unknown. As the foundation experiment for this thesis, our lab in collaboration with Nick Timchenko from Baylor University were able to demonstrate that deletion of the QA domain of TCERG1 disabled TCERG1 from repressing the G1/S growth arrest properties of C/EBPα. As shown in figure 6, it was demonstrated that as the QA domain was deleted, TCERG became unable to inhibit C/EBPα-mediated growth arrest. This lack of inhibition appeared when there were somewhere around 11 to 17 QA repeats remaining in TCERG1.

This experiment used a fluorescent-tracked growth arrest assay in which cells which were transfected with various expression plasmids and then the number of eGFP-expressing cells were evaluated. The pAdTrac plasmid used expresses both eGFP and a gene of interest, in this case C/EBPα, under the control of separate promoters to provide the researcher with an effective method to determine which cells were successfully transfected. The transfected, eGFP expressing cells were then evaluated whether they were in clusters of proliferating cells or as single, growth arrested cells. Using this data it is possible to determine if the cells were undergoing growth arrest or not.

For both HEK293T cells as well as COS7, control transfected cells were demonstrated to be ~90% proliferating. When C/EBPα was expressed in cells alone it mediated ~88% of the cells to go into growth arrest. When WT TCERG1 was expressed in cells along with C/EBPα it appeared that TCERG1 was able to effectively inhibit the growth arrest abilities of C/EBPα, such that ~90% of the cells were proliferating, similar to the control condition. By removing QA repeat residues from WT TCERG1 it was possible to obtain a point where TCERG1 was no longer able to inhibit the growth arrest potential of C/EBPα. Somewhere between 17 and 11 QA repeats TCERG1 lost its ability to inhibit the growth arrest potential of C/EBPα. Using QA_{17} TCERG1 the cells were able to effectively mediate C/EBPα growth arrest repression although QA_{11} TCERG1 and ΔQA TCERG1 lost this ability and the cells underwent growth

arrest at levels similar to those observed in the C/EBPα alone condition. The experiments were performed in both HEK293T cells as well as COS7 cells to ensure that the results presented could be replicated in different cell lines.

These data suggest that the QA domain of TCERG1 is involved in the interaction between TCERG1 and C/EBPα and formed the basis of the rest of the work undertaken in this thesis.

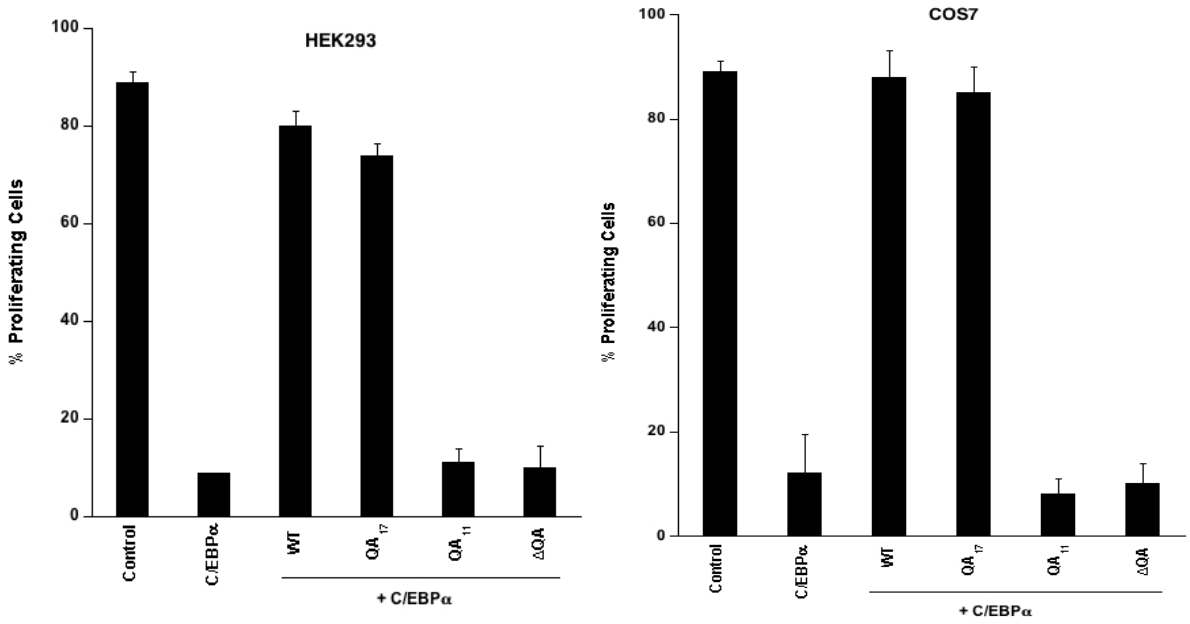

Figure 6 – The QA domain from TCERG1 affects the growth arrest potential of C/EBPα.

Growth arrest data in COS7 and HEK293T cells showing the percentage of cells still proliferating after transfection with TCERG1 QA mutants and C/EBPα. Cells were transiently transfected using pAdTrac alone for the control experiment or with pAdTrac-C/EBPα for the remaining experiments. BOST7-TCERG1 was transfected into the final four conditions using WT, QA_{17}, QA_{11} or ΔQA mutants respectively. Experiments were performed by Dr. Nick Timchenko from Baylor University in triplicate. Data is presented as the mean of the three experiments performed with standard error indicated.

4.2 The Length of the QA Domain in TCERG1 Influences the C/EBPα-Mediated Relocalization of TCERG1

Banman *et al.*, (2010) demonstrated that TCERG1 (both endogenous and over-expressed) underwent a relocalization from nuclear speckles to pericentromeric domains where C/EBPα normally resides when C/EBPα is expressed in the cell. They were further able to identify that this relocalization property resided within the first half of TCERG1 encompassing amino acids 32-668. We sought to further characterize this relocalization by identifying the domain inside TCERG1 which mediates this relocalization event. TCERG1 contains several important domains of interest in the amino terminus including a polyproline, STP, QA, KE, and 3 WW domains (see figure 3). Initially, work was focused upon the WW domains inside TCERG1 as potential mediators of this activity due to their being normally associated with protein interactions. By deleting the WW domains individually as well as in combination it was demonstrated that TCERG1 was unable to lose its C/EBPα-mediated relocalization (data not shown). Upon confirming that the WW domains were unable to mediate the relocalization of TCERG1, the QA domain was investigated further.

For this project, several versions of TCERG1 and C/EBPα were required. Where practical, experiments were undertaken to confirm the functionality of epitope tagged versions of TCERG1. Figure 7 highlights the major epitope tagged variants of TCERG1 used in this research. All of the fluorescent tagged versions of TCERG1 used in this thesis such as mKO2, mNeptune and mOrange2 all used amino terminus tags similar to the mCherry mutants shown, therefore for brevity are not included inside the figure.

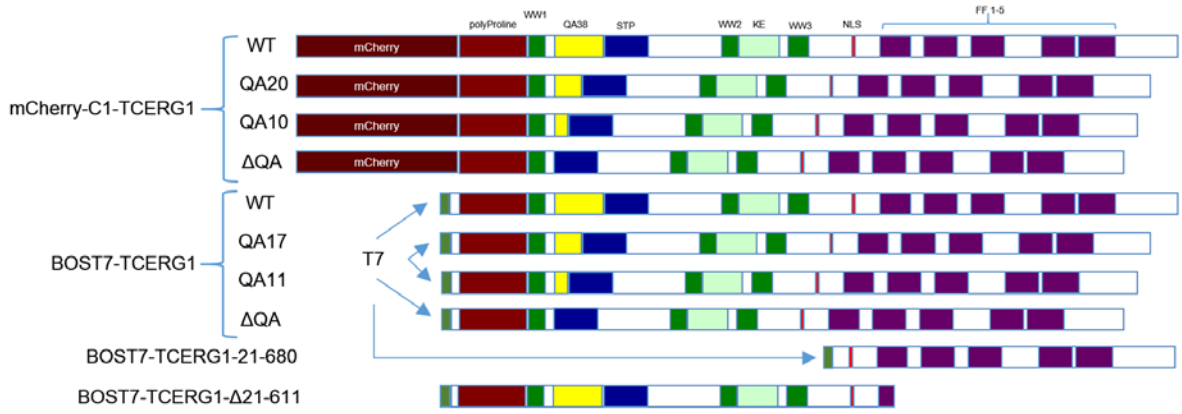

Figure 7 – TCERG1 deletion mutants.

Schematics of the major TCERG1 mutants used in this thesis. The domains shown in the image are pictured to scale with respect to each other. The specific domain information for TCERG1 can be found in figure 3. All fluorophore-C1 TCERG1 mutants are similar to mCherry-C1-TCERG1 with the specific fluorophore exchanged in each case. The orange box in the BOST7 mutants indicates the T7 tag.

The versions of C/EBPα used in this thesis are more varied in their composition than the TCERG1 mutants. It was discovered that depending upon the specific property of C/EBPα being tested, the epitope tag could potentially interfere with the activity of C/EBPα (this work, see section 4.2.5). For the most common used epitope tags for C/EBPα in this thesis, FLAG and eGFP, both an amino and carboxy tagged version of C/EBPα was generated as shown in figure 8. This allowed us to test the functionality of C/EBPα in the various assays to determine if an amino or carboxy terminal tag affected the activity of C/EBPα. The HA tagged version of C/EBPα was only created in an amino tagged version. In order to test untagged versions of C/EBPα, two methods were used to produce a C/EBPα expression plasmid which also expressed eGFP in order to track cells which were transfected. Firstly, an eGFP-P2A-C/EBPα construct was created which contained a self-cleaving P2A peptide between the eGFP and C/EBPα. This results in the cleavage of eGFP from C/EBPα that left only a single proline residue on the amino terminus of C/EBPα. While Western blotting indicated that cleavage was near 100% efficient (see figure 15) it was decided that using an internal ribosome entry site (IRES) would be a better option as there would be no chance to have uncleaved product affecting the results. To accomplish this the AcGFP-IRES-C/EBPα construct was created which contained a C/EBPα coding sequence with an IRES to separately translate an AcGFP fluorescent marker which is not fused to C/EBPα. AcGFP is a different version of GFP which

is derived from *Aequorea coerulescens* whereas eGFP (enhanced GFP) is derived from *Aequorea victoria*. Both GFP fluorophores are monomeric.

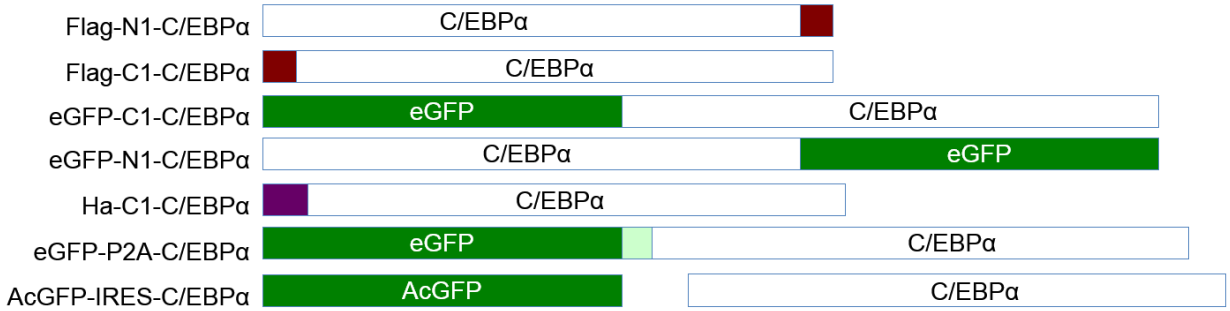

Figure 8 – C/EBPα mutant schematics.

Schematic representations of the major C/EBPα eptitope tagged versions used in this thesis. The FLAG tag is depicted in red, the HA tag in purple, the P2A peptide in light green and GFP in dark green. Domains are pictured to scale with each other. In the AcGFP-IRES-C/EBPα construct, the AcGFP and C/EBPα are translated as separate proteins, therefore are depicted separately.

4.2.1 The QA Domain is Required for the Relocalization of TCERG1 from Nuclear Speckles to Pericentromeric Regions

Since it had been demonstrated that the QA domain was affecting the ability of TCERG1 to inhibit C/EBPα-mediated growth arrest, the ability for the QA domain to mediate the relocalization of TCERG1 from nuclear speckles to pericentromeric regions was examined. Prior to performing subsequent experiments, an experiment was performed to determine if deleting the QA domain from TCERG1 affected its inherent nuclear speckle localization pattern. For each of the generated mutants, localization was compared to SC35 localization patterning that was determined by immunostaining (see figure 9 – panel "A"). SC35 was used in these experiments since it in not known to interact with either TCERG1 or C/EBPα and is commonly used as a nuclear speckle marker. For each of the four TCERG1 protein versions tested it was observed that TCERG1 co-localized to the nuclear speckle compartments. This suggested that deleting the QA domain from TCERG1 did not affect its ability to localize to or remain in the nuclear speckle compartment in the absence of C/EBPα expression. Although not examined further, it was noted the in the ΔQA TCERG1 images that the nuclear speckle compartments seemed to be enlarged.

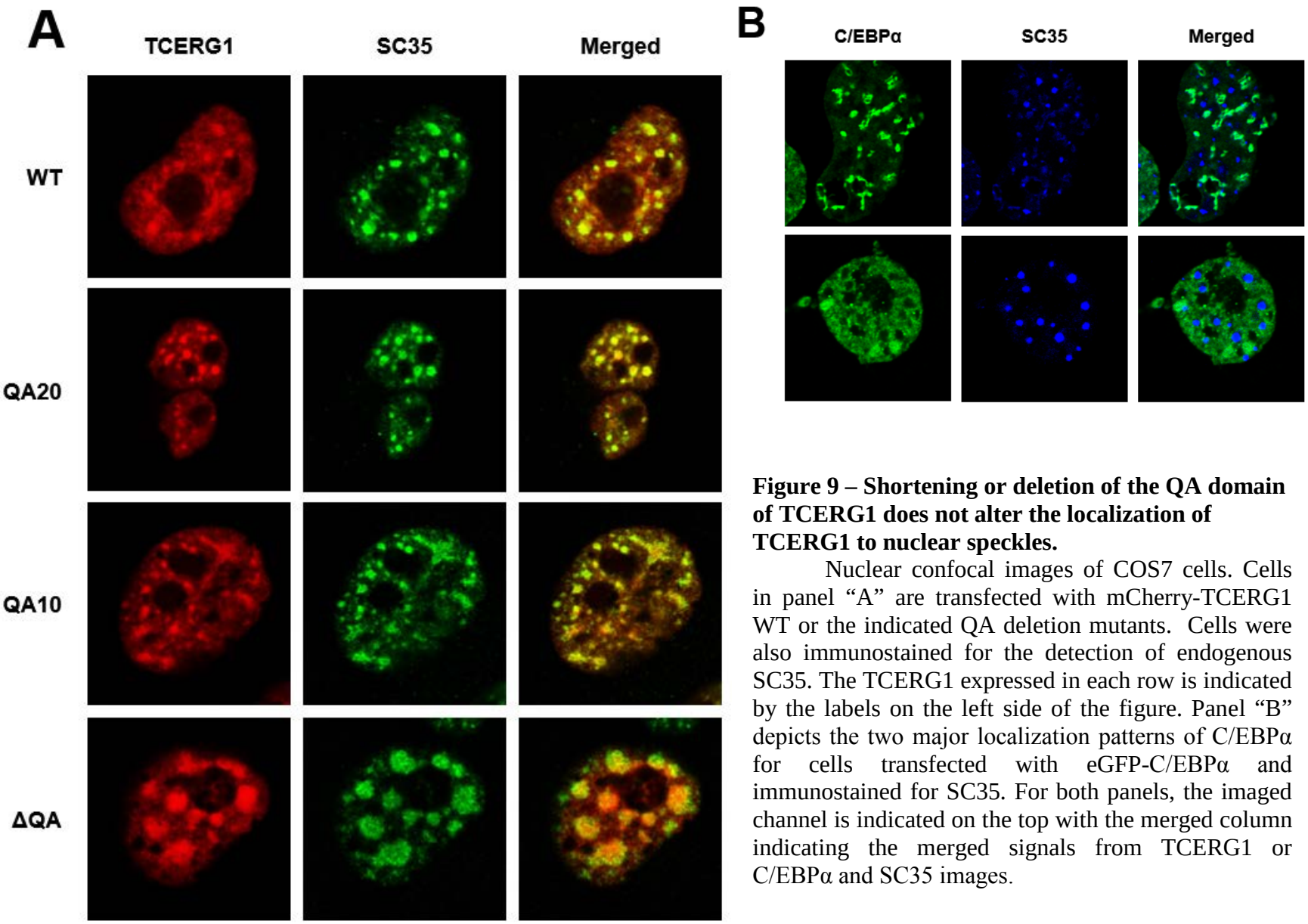

Figure 9 – Shortening or deletion of the QA domain of TCERG1 does not alter the localization of TCERG1 to nuclear speckles.

Nuclear confocal images of COS7 cells. Cells in panel "A" are transfected with mCherry-TCERG1 WT or the indicated QA deletion mutants. Cells were also immunostained for the detection of endogenous SC35. The TCERG1 expressed in each row is indicated by the labels on the left side of the figure. Panel "B" depicts the two major localization patterns of C/EBPα for cells transfected with eGFP-C/EBPα and immunostained for SC35. For both panels, the imaged channel is indicated on the top with the merged column indicating the merged signals from TCERG1 or C/EBPα and SC35 images.

Panel "B" of figure 9 depicts the two major observed patterns of C/EBPα expression. The top row of figure 9 – panel "B" is the "spongy" patterning of C/EBPα which was previously observed as the major patterning of C/EBPα in Banman *et. al.*, 2010, although as discovered in this thesis, this pattern becomes the minority of observed expression patterns using amino tagged eGFP-C1-C/EBPα or small epitope tagged versions (See figure 17). The majority of cells imaged demonstrated the fairly dispersed patterning present in the second row of figure 9 – panel "B". In both patterns, the signal from C/EBPα and SC35 did not overlap.

Next, C/EBPα was co-expressed to determine if it was able to mediate the relocalization of TCERG1 from nuclear speckles to pericentromeric regions. The TCERG1-C/EBPα merged column in figure 10 demonstrates that somewhere between 20 and 10 QA repeats, the relocalization of TCERG1 mediated by C/EBPα was lost. In the TCERG1-C/EBPα comparison column for the WT and QA_{20} TCERG1 conditions there is significant overlap between the signals for C/EBPα and TCERG1. For the last two rows of the TCERG1-C/EBPα merged column for the QA_{10} and ΔQA conditions there is a clear lack of overlap observed between the C/EBPα and TCERG1 signals.

Examination of the TCERG1-SC35 merged panels indicated that TCERG1 was unable to be relocalized away from the nuclear speckle compartments in the QA_{10} and ΔQA conditions. While the WT and QA_{20} TCERG1 conditions are similar to what has been demonstrated previously wherein TCERG1 moves away from nuclear speckles to the pericentromeric sites where C/EBPα resides (Banman *et al.*, 2010), the QA_{10} and ΔQA were unable to be relocalized by C/EBPα.

The merged C/EBPα-SC35 panels confirmed that the movement that occurred was by TCERG1 and not C/EBPα. In all conditions there is a clear separation between the C/EBPα signal and the SC35 signal with no overlap.

This experiment demonstrates that when the number of TCERG1 QA repeats decreased below 20 QA repeats that TCERG1 was unable to become relocalized from nuclear speckles to pericentromeric domains. Therefore, similar to the growth arrest assay in figure 6 there appears to be a threshold to the number of QA repeats – in this case, somewhere between 20 and 10 QA repeats – which are required to mediate the C/EBPα-mediated relocalization of TCERG1. This would suggest that in order for C/EBPα mediated relocalization of TCERG1 requires greater than 10 QA repeats with the actual number being somewhere between 11 and 20.

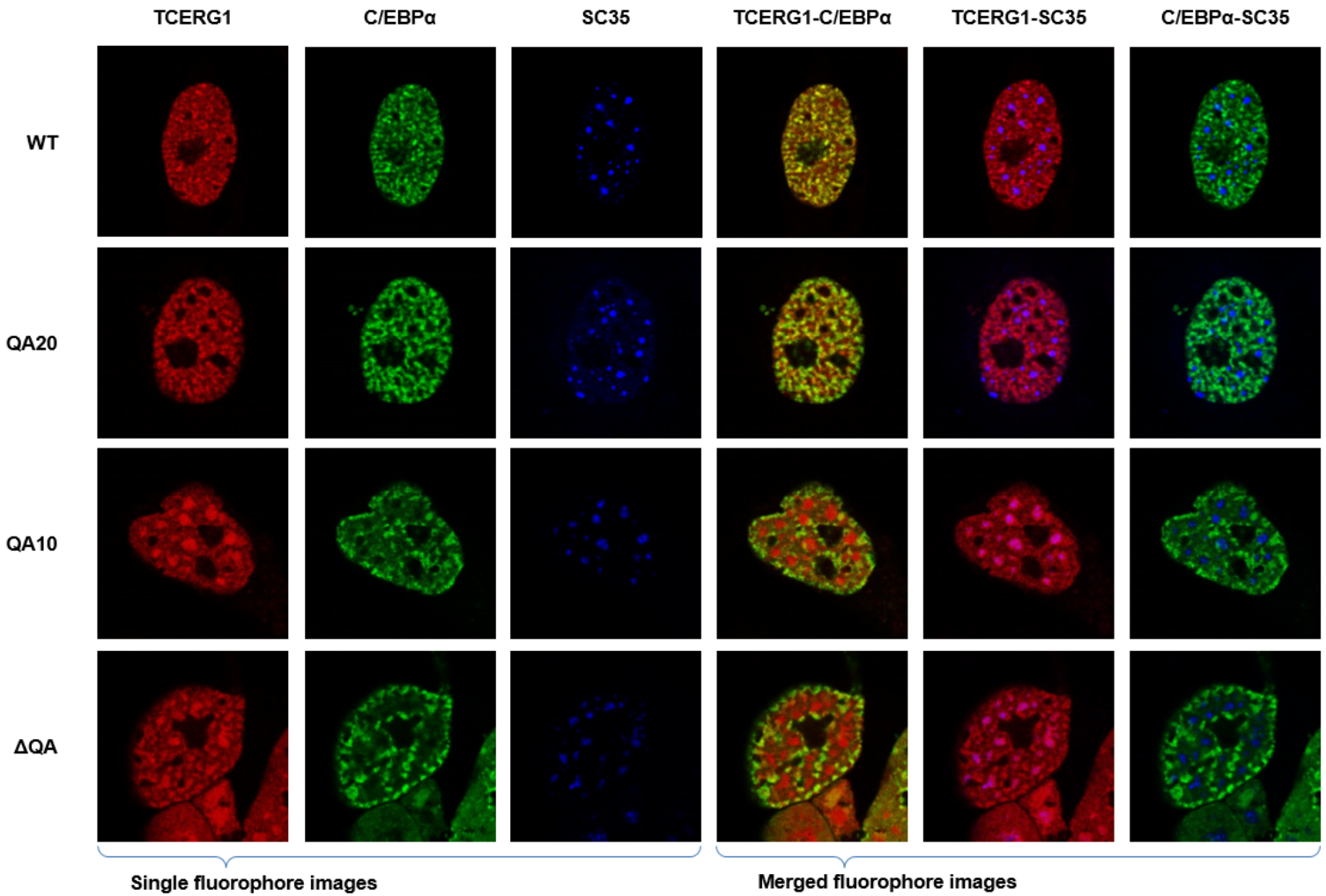

Figure 10 – Deletion of the TCERG1 QA domain abrogates the relocalization of TCERG1 from nuclear speckles to pericentromeric domains in the presence of C/EBPα.

Cells were transiently transfected with mCherry-TCERG1 WT or QA deletion mutants QA_{20}, QA_{10} and ΔQA along with eGFP-C1-C/EBPα. Cells were immunostained using SC35 to visualize the nuclear speckle compartment. The left side labels indicate the TCERG1 construct transfected and the top labels indicate which protein or pair of proteins are depicted in the image.

4.2.2 The Relocalization of TCERG1 by C/EBPα is Not Fluorophore Specific

mCherry was used initially for as a fluorescent epitope tag since it is one of the brightest red fluorophores and has a high extinction coefficient as well as minimal overlap with the excitation/emission spectra of eGFP. Initially, the mCherry fluorophore was sufficient for testing the interactions between TCERG1 and C/EBPα. It became apparent for the experiments planned, additional fluorophores would be required. Since we were using Alexa 405, eGFP and mCherry it only allowed for blue, green and red imaging. mCherry is too close to the far red spectrum to allow the use of far red fluorophores and too close to the orange spectrum to be able to use these as well.

Fluorophores used for *in vivo* imaging of cells usually have a bell-shaped distribution of excitation and emission ranges. In order to image several different proteins in the cell one needs to make use of different fluorescent marker variants which contain different wavelengths for their emission and excitation curves to minimize off-target excitation and emission. Since mCherry sits halfway between the far red and orange spectrums it didn't allow the proper imaging of either fluorophore. To overcome this, two new fusions of TCERG1 were created to allow imaging of four colors in the cell simultaneously. mNeptune2-TCERG1 and mKO2-TCERG1 which are far red and orange fluorophores, respectively were generated to replace the mCherry-TCERG1. This would allow imaging throughout the range of spectrum using Alexa 405 to image in the blue spectrum, eGFP for green, mKO2 for orange and mNeptune or Alexa 647 for far red emission. In order to test if there were any fluorophore specific problems, an experiment similar to figure 9 was performed to ensure that the expected localization patterns of TCERG1, C/EBPα and SC35 were observed; the data is presented in figure 11 for the mKO2 mutants.

As observed in the first two rows of figure 11, TCERG1 WT and ΔQA were both able to maintain nuclear speckle localization as was demonstrated by the TCERG1-SC35 merged column wherein the signals from both the TCERG1 and SC35 panels overlapped. This nuclear speckle localization supports the data presented in figure 10 and demonstrated that using different fluorophores attached to TCERG1 produced similar patterning. The eGFP signal was included in these conditions to demonstrate that there was no interaction between the eGFP fluorophore and either TCERG1 or SC35. This lack of interaction was demonstrated by a lack of patterning or speckling observed in the eGFP panels. Untagged eGFP did not have a nuclear

localization signal and was therefore found throughout the cell, whereas TCERG1 and SC35 were found exclusively in the nucleus. In the eGFP-SC35 and the eGFP-TCERG1 merged columns, the lack of overlap of signal was further observed since the signal of eGFP can be clearly observed outside the nucleus whereas both the SC35 and TCERG1 signals were confined to the nucleus.

Rows three and four of figure 11 demonstrate that, similar to the data presented in figure 10 WT TCERG1 was able to be relocalized by C/EBPα, whereas ΔQA TCERG1 was not. The eGFP-TCERG1 merged column, demonstrated that WT TCERG1 overlapped with C/EBPα whereas the ΔQA TCERG1 did not. As well, the TCERG1-SC35 merged column demonstrated that WT TCERG1 was moved from the nuclear speckle compartment whereas the ΔQA TCERG1 remains localized to nuclear speckles. Once again, to demonstrate that the movement was by TCERG1, the GFP-SC35 comparison column shows no overlap of signal in both conditions.

Taken together, these data suggest that there were no fluorophore specific problems observed when the fluorophore of TCERG1 was switched to allow for differential imaging. The mNeptune2-TCERG1 constructs were tested in a similar manner and the results obtained were consistent with the results presented for the mKO2 and mCherry mutants of TCERG1 (data partially presented in figure 14).

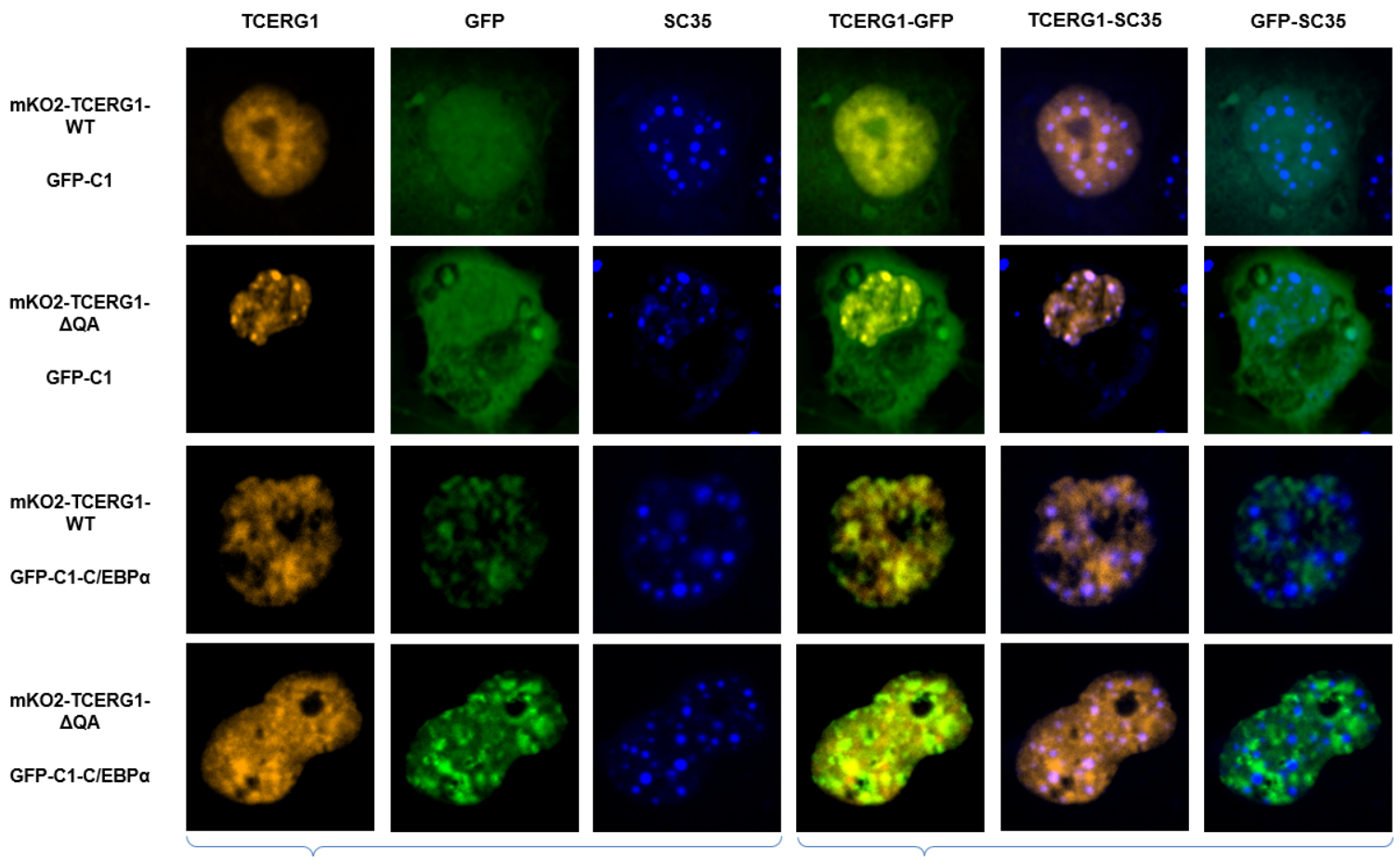

Figure 11 – The nuclear localization pattern of TCERG1 does not change dependent upon the fluorophore fused to TCERG1.

COS7 cells were transiently transfected with mKO2-TCERG1 WT or ΔQA along with eGFP-C1 +/- C/EBPα and immunostained for SC35. The left hand column describes the TCERG1 as well as the GFP construct transfected. The first three image columns are the isolated signal whereas the last 3 columns are merged signals from two of the first three panels to demonstrate interaction. Rows one and two demonstrate the localization of TCERG1 in the absence of C/EBPα. Rows three and four demonstrate the relocalization patterns of TCERG1 WT and ΔQA in the presence of C/EBPα expression.

4.2.3 ΔQA TCERG1 Acts as a Dominant Negative With Respect to Relocalization of TCERG1

The experiments described in this thesis involve the ectopic expression of proteins in the cell. This thesis makes extensive use of transiently transfected overexpressed proteins. When over-expressing proteins in a cell the issue that often arises is that the over-expressed proteins have different expression characteristics than their endogenous counterparts (Gibson *et al.*, 2013). To ensure that the problems of overexpression were minimized, an experiment was undertaken to further test if there were artifacts due to overexpression of the proteins of interest. An experiment was performed using all three overexpressed proteins mNeptune2-TCERG1-WT with mKO2-TCERG1-ΔQA with eGFP-C1 +/-C/EBPα.

The first row of figure 12 depicts the localization of overexpressed TCERG1 proteins in the absence of C/EBPα expression. The TCERG1 WT-ΔQA merged image depicts that both the WT and ΔQA version of TCERG1 co-localized. In the TCERG1-SC35 merged panels for both WT and ΔQA there is complete overlap between both signals in both images. This demonstrated that both WT and ΔQA TCERG1 are localized to nuclear speckles when co-expressed together.

The second row of figure 12 compares the expression patterns of overexpressed WT and ΔQA TCERG1 when C/EBPα was co-expressed. Previously it had been demonstrated in figure 10 that TCERG1 WT was able to undergo C/EBPα-mediated relocalization, whereas TCERG1 ΔQA was not. We wanted to determine what effect there would be if both proteins were expressed in the cell at the same time.

Surprisingly, it was discovered that when WT and ΔQA TCERG1 were expressed simultaneously that C/EBPα was unable to relocalize either protein. As depicted in the second row of the fluorophore comparisons of figure 12 in the WT-ΔQA TCERG1 merged image, both signals overlap each other. Additionally, comparing the TCERG1 WT or ΔQA-GFP images there is a clear separation of signal in both images, demonstrating that neither WT nor ΔQA TCERG1 was able to be relocalized. In fact, using the WT TCERG1 or ΔQA-SC35 images it is discovered that TCERG1 remains localized to nuclear speckles in both cases. Once again, the C/EBPα and SC35 signals did not overlap as demonstrated by the GFP-SC35 merged image.

These results suggest that ΔQA TCERG1 acts as a dominant negative protein which prevented relocalization of WT TCERG1. Over-expression of proteins can drive protein interactions which are not normally present in cells. To test this interaction an experiment

examining endogenously expressed TCERG1 along with transiently expressed ΔQA TCERG1with and without C/EBPα was performed. Moazed *et al.,* (2010) demonstrated that endogenous TCERG1 was able to be re-localized when C/EBPα is co-expressed. Since endogenous TCERG1 is expressed in the cell we could observe its movement along with ΔQA TCERG1 to determine if C/EBPα is able to relocalize either TCERG1. The experiment was performed using over-expressed ΔQA TCERG1 with eGFP-C1 +/-C/EBPα. Cells were then immunostained for endogenous TCERG1 as well as SC35 (figure 13). To ensure staining of only endogenous TCERG1 and not the over-expressed ΔQA TCERG1 we used an antibody which is specific to the first 20 amino acids of TCERG1, which are not present in the fluorescent tagged versions of TCERG1.

The first row of figure 13 demonstrates the interaction between endogenous TCERG1 and ΔQA TCERG1 without C/EBPα expression. In the endoTCERG1-TCERG1 ΔQA merged panel, it can be observed that there is complete overlap between the two signals, suggesting that they are co-localized in the nucleus. When observing both the endoTCERG1-GFP and TCERG1 ΔQA-GFP panels the TCERG1 signal can be observed completely localized to the nucleus whereas the eGFP signal is expressed homogenously throughout the cell. This demonstrates that there is no interaction between eGFP and either TCERG1 construct as there was no movement of either protein. Observing the panels for endoTCERG-SC35 as well as the TCERG1 ΔQA-SC35 it is possible to determine that both signals are located within the nuclear speckle compartment. This demonstrates that both of these proteins are localizing to nuclear speckles, as expected, in the absence of C/EBPα.

The second row of figure 13 demonstrates the localization of endogenous TCERG1, ΔQA TCERG1 and C/EBPα. Interestingly, it was observed similar to figure 12 that the endogenous wild type TCERG1 and ΔQA versions of TCERG1 completely overlapped. Although the staining of endogenous TCERG1 did not produce the nice punctate specks of fluorescence that the fluorophore-tagged TCERG1 showed, observing the endoTCERG1-TCERG1 ΔQA panel it can be demonstrated that the areas of higher signal from the endogenous TCERG1 corresponded to high expression areas of ΔQA TCERG1. This suggested that endogenous TCERG1 and ΔQA TCERG1 were co-localized. When the TCERG1-C/EBPα comparison panels were examined, it became easier to determine that the TCERG1 and C/EBPα signals did not overlap. In both the endoTCERG1-GFP and the

TCERG1 ΔQA-GFP merged panels a clear distinction between the two signals could be observed. When comparing the TCERG1 expression patterns with the immunostained SC35 it can be observed that TCERG1 remains inside nuclear speckles. While not as clear as the TCERG1 ΔQA-SC35 panel, the endoTCERG1-SC35 panel signals overlap significantly, suggesting that TCERG1 remains localized to nuclear speckles. Once again, the GFP-SC35 panel demonstrates no overlap of signal and therefore no interaction between C/EBPα and SC35.

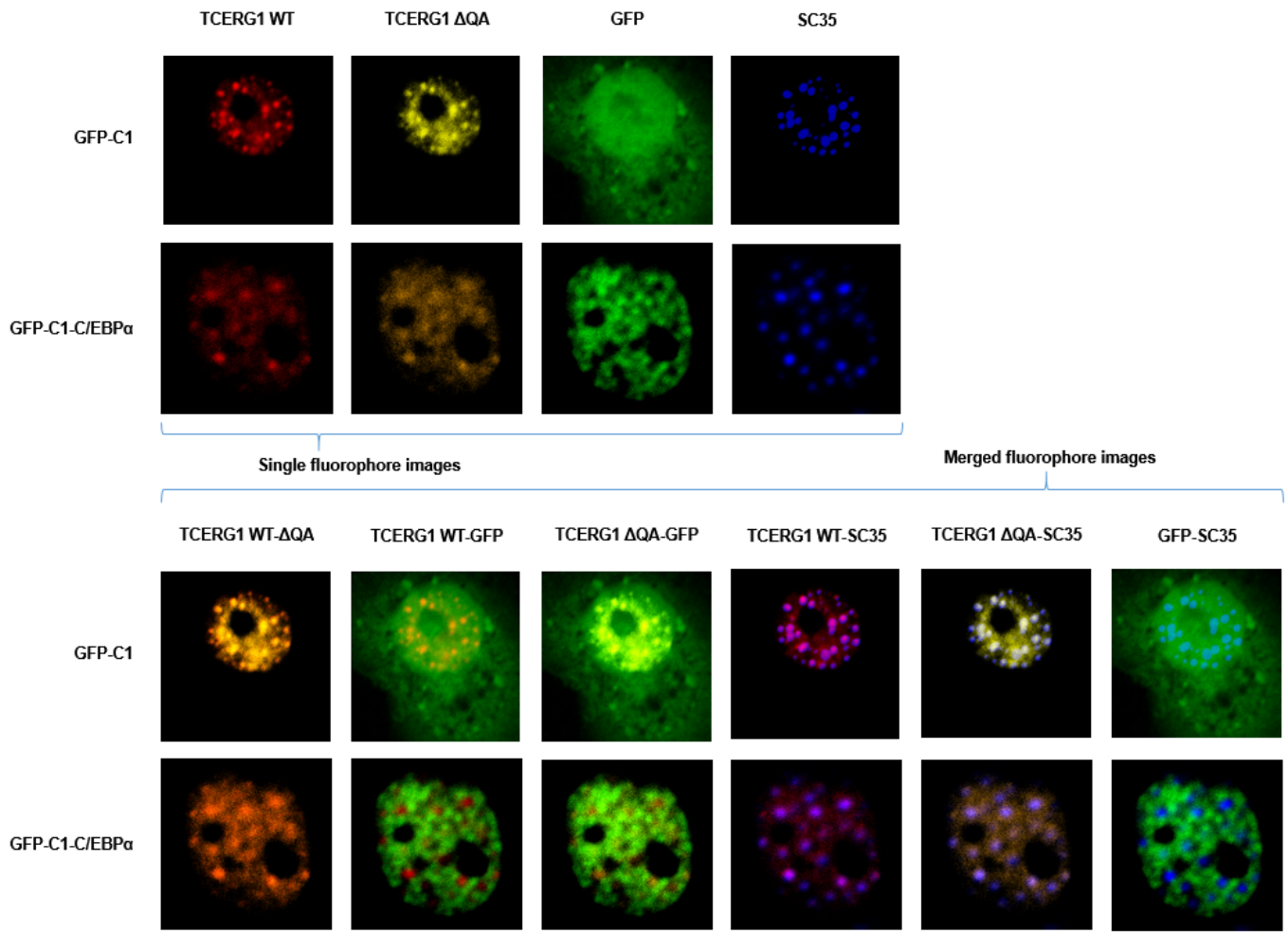

Figure 12 – TCERG1 ΔQA prevents TCERG1 WT from being relocalized by C/EBPα.

Confocal image of transiently expressed mNeptune2-TCERG1-WT and mOrange2-TCERG1-ΔQA along with eGFP-C1 +/-C/EBPα in COS7 cells. Cells were immunostained for endogenous SC35. The top four columns depict individual signals from the fluorophore as labelled while the bottom six columns contain merged signals from two of the top panels as depicted. In both top and bottom sections the first row contains eGFP-C1 as the green channel signal, labelled as "GFP" and the second row contains eGFP-C1-C/EBPα for the green channel labelled as "GFP".

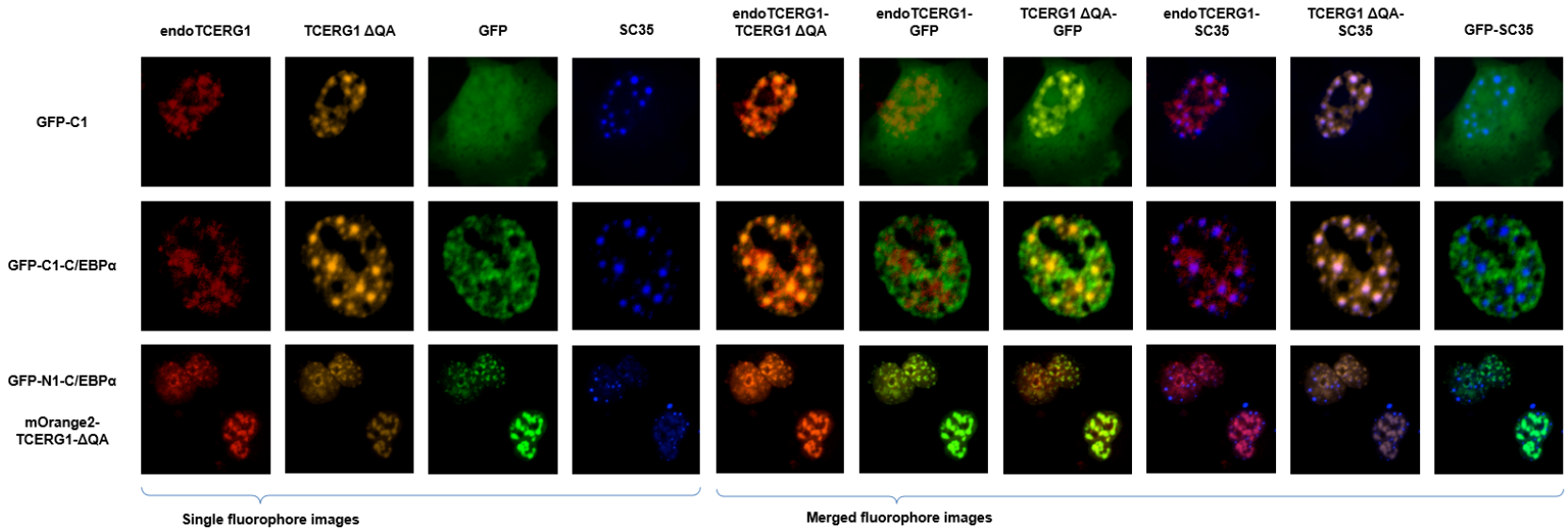

Figure 13 – Endogenous TCERG1 along with the ΔQA mutant of TCERG1 shows overlap of expression patterns.

Confocal image of immunostained endogenous TCERG1 along with mKO2-TCERG1-ΔQA and eGFP-C1 +/-C/EBPα in COS7 cells. The last row depicts cells expressing mOrange2-TCERG1 ΔQA instead of the mKO2 version along with eGFP-N1-C/EBPα to demonstrate the difference between amino and carboxy tagged versions of C/EBPα. The left side labels indicate the TCERG1 and eGFP constructs transfected. The first four top labels indicate which of the individual protein signals are being imaged and the last six top labels indicate which pair of the individual signals from the first four images are combined for comparison.

4.2.4 Differential Localization Patterns of C/EBPα and TCERG1 are Demonstrated Through Amino or Carboxy-End Epitope Tagging of C/EBPα

In performing the experiments shown in figure 13 it was noted that previous experiments in the lab had been performed using eGFP-N1-C/EBPα. This plasmid contained the eGFP fused to the carboxy terminus of C/EBPα (refer to figure 8). eGFP-C1-C/EBPα had been used in all prior experiments in this thesis which had the C/EBPα fused to the amino terminus of eGFP. To determine if the amino terminal fusion of C/EBPα was affecting the results, the experiment shown in the first two rows of figure 13 was repeated using eGFP-N1-C/EBPα. It was discovered that the ΔQA TCERG1 was able to become relocalized by eGFP-N1-C/EBPα, which is contrary to the results obtained with eGFP-C1-C/EBPα.

These results are depicted in the last row of figure 13. When the endoTCERG1-ΔQA images are merged the signals from both fluorophores were still overlapping, similar to the results presented in the second row of figure 10. The difference became apparent when examining the endoTCERG1-GFP and TCERG1 ΔQA-GFP panels. In both of these panels the TCERG1 and C/EBPα signals are overlapping each other, demonstrating relocalization. When the endoTCERG1-SC35, TCERG1 ΔQA-SC35 as well as GFP-SC35 panels were examined it is observed that both signals in each panel are separate from each other. This demonstrates movement out of nuclear speckle compartments to pericentromeric regions for both versions of TCERG1, and emphasized there was no interaction between C/EBPα and SC35. Taken together these results suggest that fluorescent tagging of C/EBPα on the amino or carboxy terminus can produce different results and these differential results will need to be examined in greater detail in later experiments.

To further investigate which relocalization pattern for ΔQA is correct, a comparison between the N1 and C1 tagged versions of eGFP-C/EBPα was undertaken (see figure 14). WT or ΔQA TCERG1 were co-expressed with amino or carboxy terminal eGFP fused C/EBPα and then the localization patterns were observed. As was suspected, it was discovered that relocalization of both the TCERG1 WT and ΔQA was induced by the N1 tagged versions of C/EBPα whereas only the TCERG1 WT was able to be relocalized by the eGFP-C1-C/EBPα version.

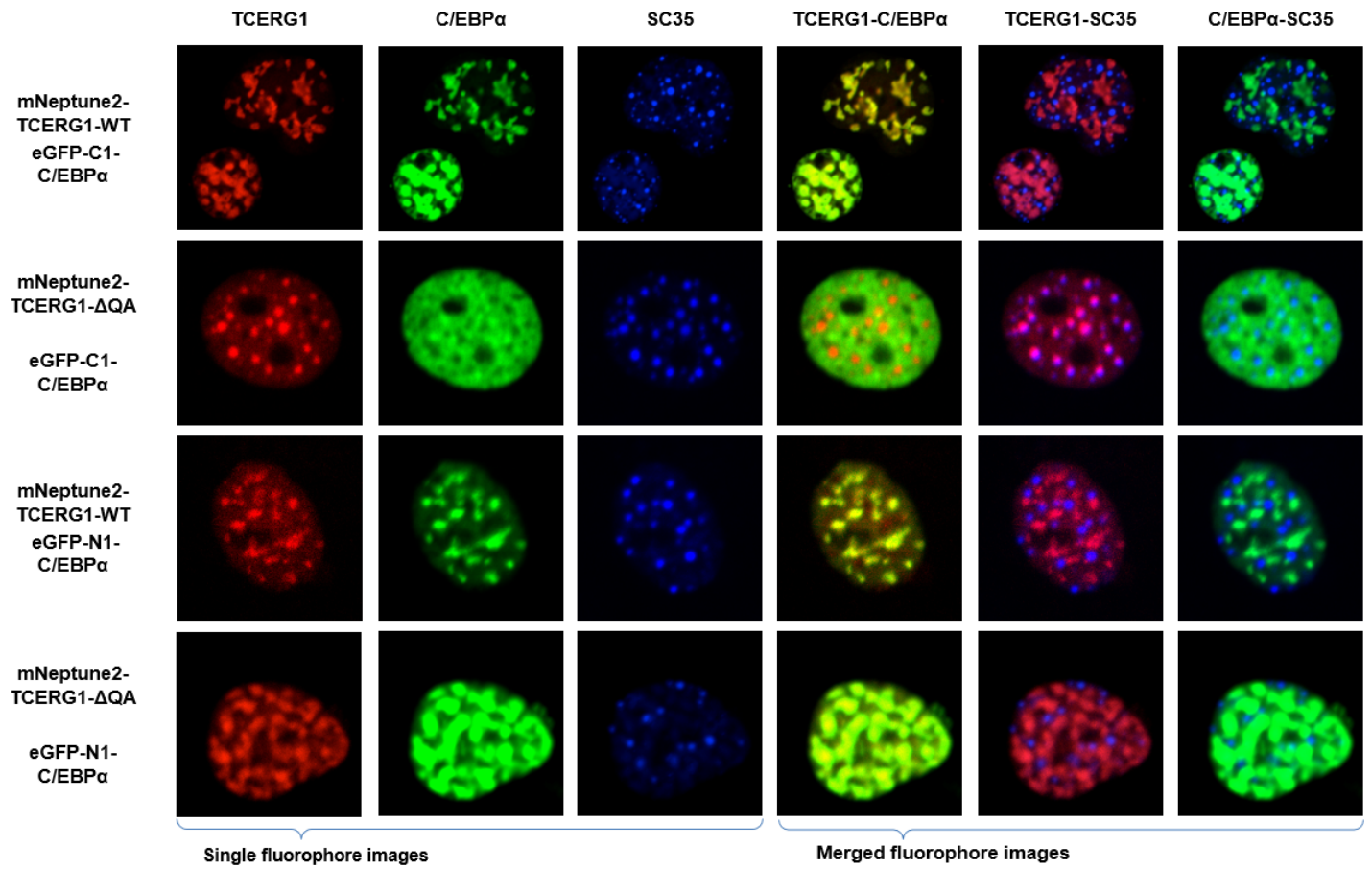

Figure 14 – Amino tagged eGFP-C/EBPα relocalizes TCERG1 in a QA dependent manner whereas carboxy tagged eGFP-C/EBPα does not.

Nuclear confocal comparison of amino and carboxy tagged eGFP-C/EBPα when compared to mNeptune2-TCERG1 WT or ΔQA in COS7 cells. eGFP-C1-C/EBPα contains eGFP fused to the amino end of C/EBPα whereas eGFP-N1-C/EBPα contains eGFP fused to the carboxy terminus of C/EBPα. The left side labels indicate the TCERG1 and C/EBPα constructs expressed and the top labels indicate which protein or pair of proteins are depicted in the image. The first three columns are single protein images whereas the last three depict pairs of merged signals from the initial three columns to facilitate signal comparison.

Comparing the TCERG1-C/EBPα column for all four conditions it was observed in all but the ΔQA TCERG1 and eGFP-C1-C/EBPα condition that the signals for TCERG1 and C/EBPα are overlapping. This suggests that C/EBPα is able to relocalize TCERG1 for both WT conditions as well as the ΔQA with eGFP-N1-C/EBPα but not the ΔQA with eGFP-C1-C/EBPα. The panels for TCERG1-SC35 comparison agree with these conclusions wherein the signals for SC35 and TCERG1 are separate in all conditions except ΔQA TCERG1 and eGFP-C1-C/EBPα. This demonstrates that C/EBPα is able to mediate the relocalization away from nuclear speckles in all conditions but the one. Once again, the signals for the C/EBPα-SC35 panels do not overlap, demonstrating no interaction between these two proteins.

Untagged versions of C/EBPα were next examined for their ability to induce relocalization and whether the relocalization was QA domain dependent. Two expression plasmids containing untagged C/EBPα were created for this experiment. The first plasmid contained AcGFP expression separate from C/EBPα via an internal ribosome entry site (IRES) such that transfection efficiency could be tracked using AcGFP to determine which cells had been successfully transfected. As well, the second expression plasmid for C/EBPα used a P2A self cleaving peptide between C/EBPα and eGFP such that upon cleavage only a single proline residue was left at the amino terminus of C/EBPα. Efficiency of transfection and cleavage of the P2A peptide was evaluated using Western blotting with no observable uncleaved product detected as shown in figure 15. The Western blot produced a blot containing a single band of expression close to the expected 37 kDa molecular weight for untagged C/EBPα.

COS7 cells were transfected with plasmids expressing either the IRES-C/EBPα or P2A-C/EBPα as well as either WT or ΔQA TCERG1. As demonstrated in figure 16, using cells which were expressing green fluorescence, both of the C/EBPα constructs demonstrated results similar to previous observations wherein wild type TCERG1 no longer localized to SC35-labelled nuclear speckles whereas the ΔQA versions were still found in nuclear speckle compartments.

By examining the WT TCERG1 expression pattern in both C/EBPα conditions in figure 16 a clear movement away from nuclear speckles can be observed. By comparing the signal patterns from TCERG1 and SC35 in the merged column it can be clearly seen that WT TCERG1 is no longer localized to SC35 containing nuclear speckles. Alternatively the ΔQA TCERG1 expression can be clearly seen to be inside the SC35 nuclear speckle signals in the

merged panels of figure 16. The green fluorescent signal from each cell was not included in the figure as it was used as a tracker of transfection and had no physiological relevance to the experiment otherwise.

Due to being unable to immunostain C/EBPα, the data is obtained indirectly through the movement of TCERG1 signal. By following the movement of TCERG1 away from nuclear speckles and by comparing and contrasting the behaviour of TCERG1 to the data presented in previous figures we can indirectly predict how TCERG1 is interacting with C/EBPα. The data presented in figure 16 indicates that the previous results obtained using eGFP-C1-C/EBPα are indeed the correct results whereas the eGFP-N1-C/EBPα results are artifactual by the inclusion of the eGFP fluorphore at the carboxy terminus of C/EBPα. Due to the inclusion of mCherry and GFP signals it was only possible to image a maximum of three fluorophores in this experiment, therefore it was impossible in this experiment to directly track the signal of C/EBPα.

To ensure that the green signal observed in the cells was indicative of sufficient levels of isolated C/EBPα expression, a Western blot was performed using each C/EBPα construct. The Western blot was used to ensure that the IRES site of AcGFP-IRES-C/EBPα was functioning correctly and expressing C/EBPα, as well as the P2A peptide of eGFP-P2A-C/EBPα was efficiently cleaving the eGFP from C/EBPα. As demonstrated in figure 15 both constructs were able to produce efficient C/EBPα expression levels at the expected molecular weights. This demonstrates that both of these constructs are expressing the correct untagged C/EBPα at desirable levels with highly efficient cleavage in eGFP-P2A-C/EBPα and non-fusion in AcGFP-IRES-C/EBPα. From these results we can conclude that the intensity of the green fluorescence observed in the cells imaged in figure 16 is correspondantly equivalent to expression of C/EBPα in those cells.

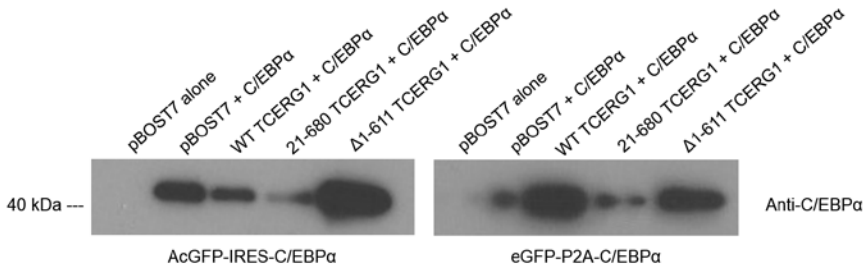

Figure 15 –AcGFP-IRES-C/EBPα and eGFP-P2A-C/EBPα produce a single band of the correct molecular weight.

Western blots probed with anti-C/EBPα to determine transfection expression levels of indicated C/EBPα expression plasmids. COS7 cells were transiently co-transfected with plasmids expressing either AcGFP-IRES-C/EBPα or eGFP-P2A-C/EBPα along with the indicated BOST7 plasmid and subsequently probed with anti-C/EBPα antibodies. No other bands were detected on the film therefore only a section of the film is included which contained the anti-C/EBPα signal. Western blotting performed by Kaitlyn Schick.

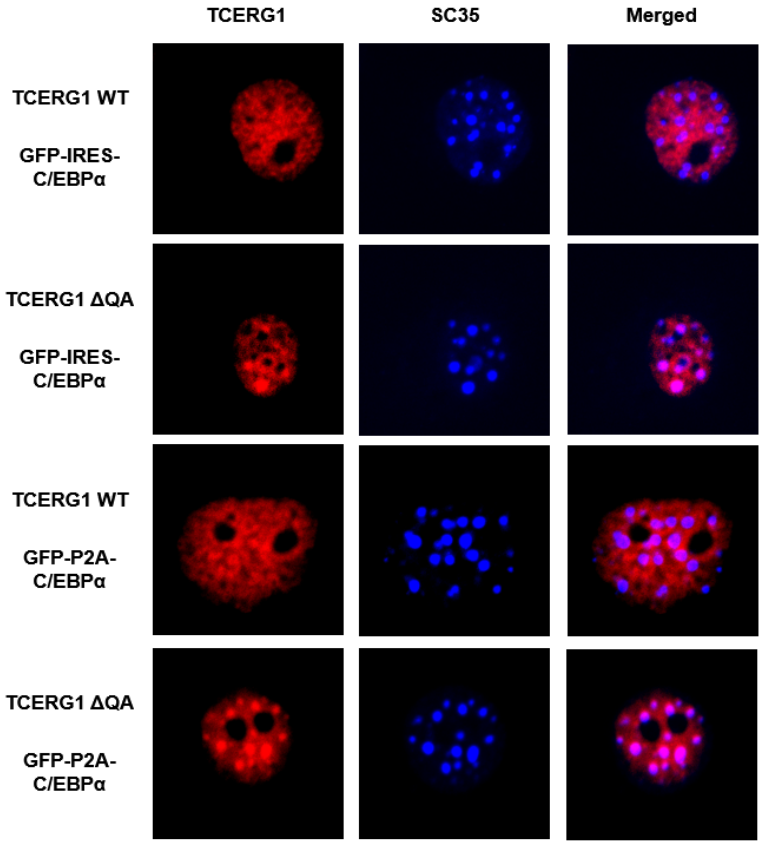

Figure 16 – Untagged C/EBPα displays similar TCERG1 relocalization characteristics to eGFP-C1-C/EBPα.

Confocal images of transiently transfected COS7 cells with plasmids expressing AcGFP-IRES-C/EBPα or eGFP-P2A-C/EBPα and either mCherry-TCERG1 WT or ΔQA. The fixed cells were subsequently immunostained for SC35. Green fluorescing cells were used for imaging. The left labels describe the transfected plasmids in each row and the top labels describe the fluorophore image or merged images presented in each column.

4.2.4.1 Differential Epitope Tagging of C/EBPα Does Not Change the Localization Pattern of C/EBPα Except for Large Carboxy-Terminal Tags

Given the finding that differential epitope tagging of C/EBPα could potentially produce different localization patterns of ΔQA TCERG1 it was decided to undertake a study to determine the localization patterns of C/EBPα present with each epitope tag present. As shown in figure 17 it was determined that the C/EBPα localization pattern present in most of the constructs were similar, with the exception of eGFP-N1-C/EBPα. In this construct, the localization of C/EBPα deviated from the mostly diffuse patterning with punctate specks of fluorescence seen with other C/EBPα versions to a more "spongy" patterning of fluorescence. In this patterning the fluorescence forms oblong, irregularly shaped spots of highly bright fluorescence with minimal fluorescence elsewhere in the cell. It was previously noted in Banman *et al.,* (2010) that the spongy patterning seen in eGFP-N1-C/EBPα was the majority of the C/EBPα patterning found in the cell with a minority of cells being in the forms found for the remaining C/EBPα versions shown in figure 17. In the remaining C/EBPα versions, including the untagged C/EBPα, IRES-C/EBPα, Ha-C/EBPα and FLAG-C/EBPα, the C/EBPα expression patterns were seen as somewhat diffuse green fluorescence with varying sizes of specks of higher intensity fluorescence (see figure 8 for diagram of mutants). These results confirmed the C/EBPα patterning results for eGFP-N1-C/EBPα described by Banman but uncovered different results for the rest of the constructs tested. In the other constructs the "spongy" patterning could be found in the slides but they were a minority only occurring in about 5-10% of the cells. The rest of the 90-95% of cells were found with diffuse nuclear fluorescence with varying amounts of round, punctate specks of increased fluorescence. These results further support our conclusions that the patterning seen in eGFP-C1-C/EBPα reflects that of untagged versions of C/EBPα.

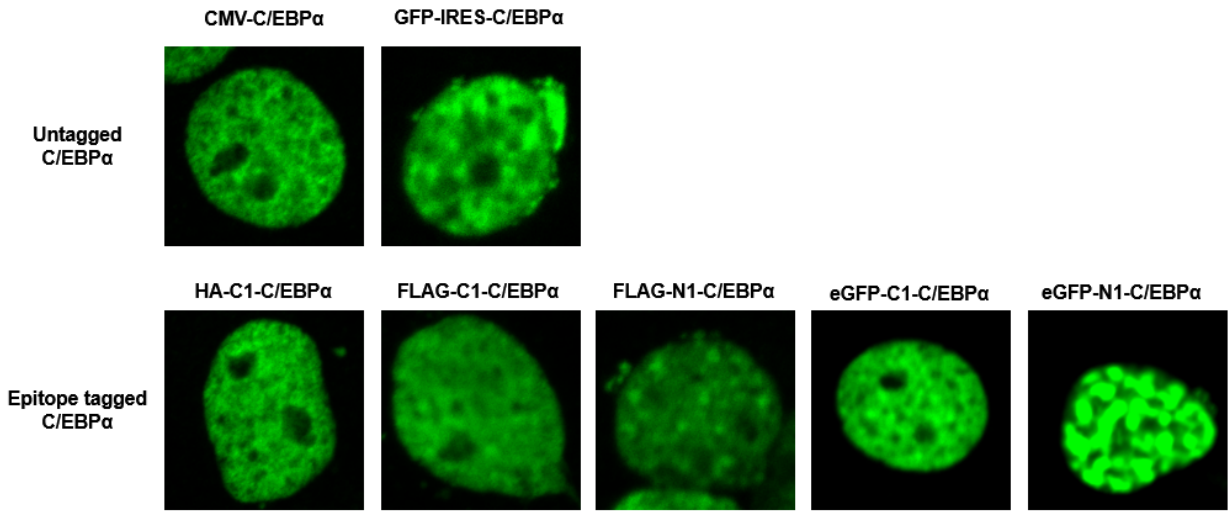

Figure 17 – Differential epitope tagging of C/EBPα produces patterns of C/EBPα expression in COS7 cells except for fluorescent markers fused to the carboxy terminus.

Confocal images of COS7 cells of various epitope tagged versions of C/EBPα. The untagged and epitope tagged C/EBPα constructs were immunostained using anti-C/EBPα antibodies. COS7 cells were transiently transfected with each version of C/EBPα in COS7 cells. Images for CMV-C/EBPα, HA-C1-C/EBPα, FLAG-C1-C/EBPα, and FLAG-N1-C/EBPα provided by Kaitlyn Schick.

4.3 The QA Domain of TCERG1 is Important For Some of The Interactions Between TCERG1 and C/EBPα

McFie *et al.,* (2006) obtained data from a yeast two-hybrid assay which suggested that TCERG1 and C/EBPα were able to physically interact with each other via the amino terminal half of TCERG1 which contains the QA domain. Thus far in this thesis it has been demonstrated that the QA domain of TCERG1 is involved in TCERG1-mediated inhibition of cellular growth arrest by C/EBPα as well as the relocalization of TCERG1 by C/EBPα. While these experiments are important to establish a role for the QA domain in the interaction between C/EBPα and TCERG1, these interactions thus far have only been studied at a cellular level. To obtain a clearer understanding of what is happening at a protein level it was decided to use co-immunoprecipitation to assess the role of the QA domain in the physical interaction between TCERG1 and C/EBPα.

4.3.1 The QA Domain of TCERG1 is Involved in the Physical Interaction Between TCERG1 and C/EBPα

To determine whether the QA repeat domain is involved in the physical interaction between TCERG1 and C/EBPα, a co-immunoprecipitation (Co-IP) was performed. Using FLAG-N1-C/EBPα to pull down proteins using anti-FLAG magnetic beads it was demonstrated that TCERG1 could be pulled down in a QA dependent manner.

Figure 18 contains a Co-IP blot in which it was demonstrated that as the number of QA repeats were reduced from 38 (WT) down to zero (ΔQA) that the corresponding amount of TCERG1 which was able to bind to the immobilized C/EBPα decreased. The WT TCERG1 as well as QA_{17} pulled down almost equal amounts of TCERG1 as indicated by the anti-T7 probing of the Co-IP condition of figure 18. There was a significant decrease in the amount of signal for QA_{11} TCERG1 and almost no signal detected for ΔQA. The amount of C/EBPα detected in the anti-C/EBPα blot remained similar across the four test conditions. The TCERG1 protein expression was strong across all conditions in the anti-T7 input blot and the C/EBPα expression levels were consistent and strong across the input blots of anti-C/EBPα.

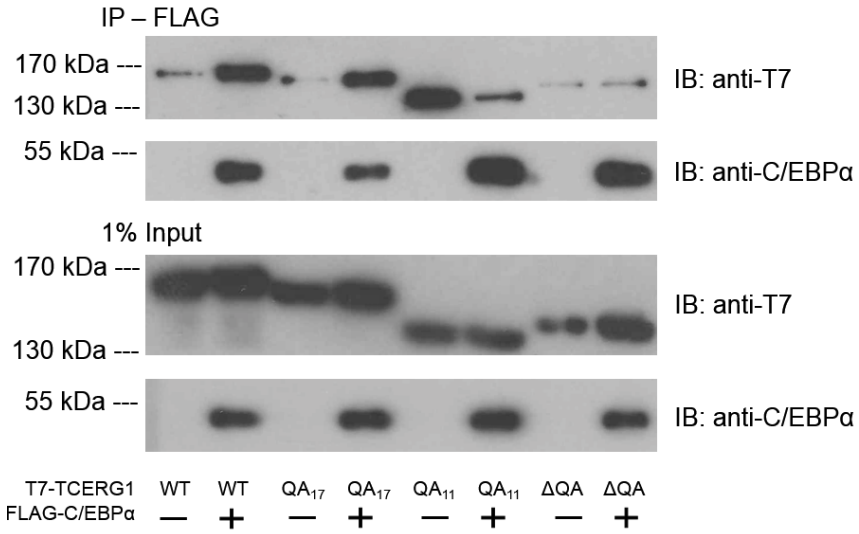

Figure 18 – The QA domain of TCERG1 is important for the physical interaction between TCERG1 and C/EBPα.

Co-immunoprecipitation pulldown assay against anti-FLAG. HEK293T were transiently transfected using BOST7-TCERG1-WT, BOST7-TCERG1-QA_{20}, BOST7-TCERG1-QA_{10} and BOST7-TCERG1-ΔQA using FLAG-N1-CEBPa as the pull-down interaction partner. Each of the input and Co-IP blots were probed using anti-T7 antibody to visualize the BOST7 tagged mutants as well as anti-C/EBPα antibody to visualize the FLAG-N1-C/EBPα.

There are several issues with the Co-IP we were unable to remedy. First, the blot shows an expression band for the QA_{11} TCERG1 condition without C/EBPα expression. When there was no FLAG-N1-C/EBPα expression, only low, background binding should bind to the anti-FLAG beads. Therefore, the protein band detected in the QA_{11} TCERG1 -C/EBPα condition should not be present. This Co-IP experiment was attempted repeatedly without success of getting a blot which showed low binding in all of the negative control conditions and clear levels of expression in all the test conditions. Through the various blotting attempts it was demonstrated several times for each condition that QA-specific pulldown was repeatedly demonstrated with minimal binding in the conditions without FLAG-N1-C/EBPα expression, just not for all eight conditions present in the experiment on the same blot. The second issue we found with these blots is that the pulldown of TCERG1 only equated to approximately 0.5% of the total lysate, which is low. This can be rationalized by the fact that both TCERG1 and C/EBPα are transcription factors and as such are interacting with many partners inside the cellular milleu, therefore would only be able to bind a small amount of each transcription factor present. Moreover, the interaction could be very weak to transient. Lastly, although not a

problem with this experiment, it was noted that the molecular weight of the ΔQA TCERG1 is inconsistent with the expected molecular weight. Thus far it is still unknown why the ΔQA mutant migrates at a higher molecular weight than the QA_{11}.

4.3.2 The Isolated QA Domain is Unable to Mediate Relocalization

Since TCERG1 and C/EBPα were demonstrated to interact in a QA-dependent manner, it was next examined whether the QA domain was sufficient to mediate relocalization. To accomplish this it was decided to fuse the QA38 domain from TCERG1 to a fluorescent reporter to be able to track its movements in the cell (refer to figure 19). It was also attempted to fuse the QA domain to SC35 but the fusion caused problems with the nuclear localization patterns of the SC35 fusion protein and therefore was not used (data not shown).

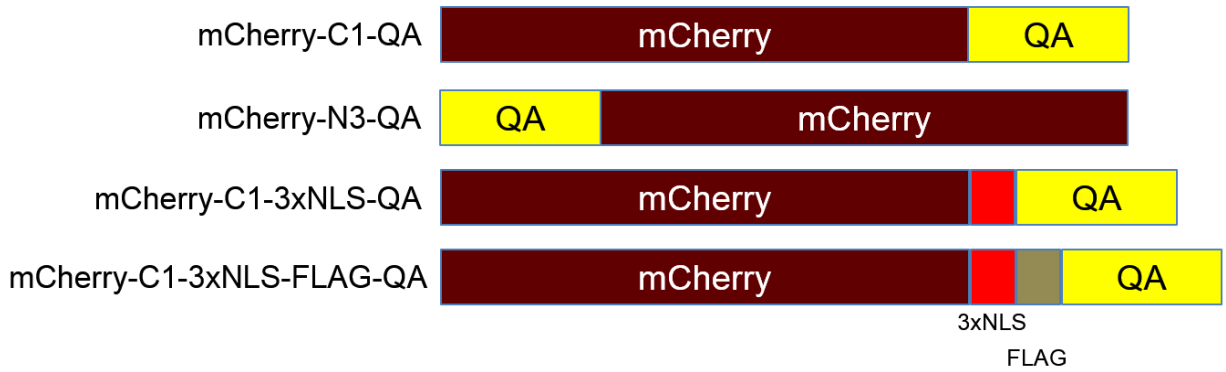

Figure 19 – Schematic of the mCherry-QA fusion constructs used.
Scale representations of the QA domain fusion proteins engineered. The QA38 domain is shown in yellow whereas the 3xSV40-NLS is depicted in red, the FLAG tag is depicted in the last construct in grey. The 3xSV40-NLS consisted of 3 synthetic tandem repeats of the SV40 nuclear localization signal (NLS).

Figure 20 details the results of an experiment which examined the localization patterns of the mCherry-C1-QA fusion. The top row of figure 20 contains images of mCherry-QA when expressed with eGFP to test localization patterns of mCherry-QA without C/EBPα expression. The mCherry-QA expression was not confined to the nucleus, nor was it localized to nuclear speckles in this condition as shown in the QA-SC35 comparison image. The eGFP expression were detected throughout the cell without any visible aggregation near the mCherry-QA, suggesting no interaction with the mCherry-QA.

The bottom two rows of figure 20 portray the test conditions for this experiment. In the second row mCherry-QA was expressed with eGFP-C1-C/EBPα whereas the third row show results using eGFP-N1-C/EBPα. The mCherry-QA was unable to co-localize with C/EBPα in any of the conditions as illustrated by the QA-eGFP comparison panels for these three conditions. Whereas the C/EBPα signal was localized inside the nucleus, the QA expression was mostly cytoplasmic with larger amounts of expression being present next to the nuclear membrane. Of the QA that is localized to the nucleus there appears to be no overlap of signal between C/EBPα and QA, therefore no observable interaction. The merged QA-SC35 panels demonstrate that SC35 expression continues to be nuclear localized in these conditions, with some of the nuclear localized QA being near the SC35 speckles but with little overlap of the two proteins.

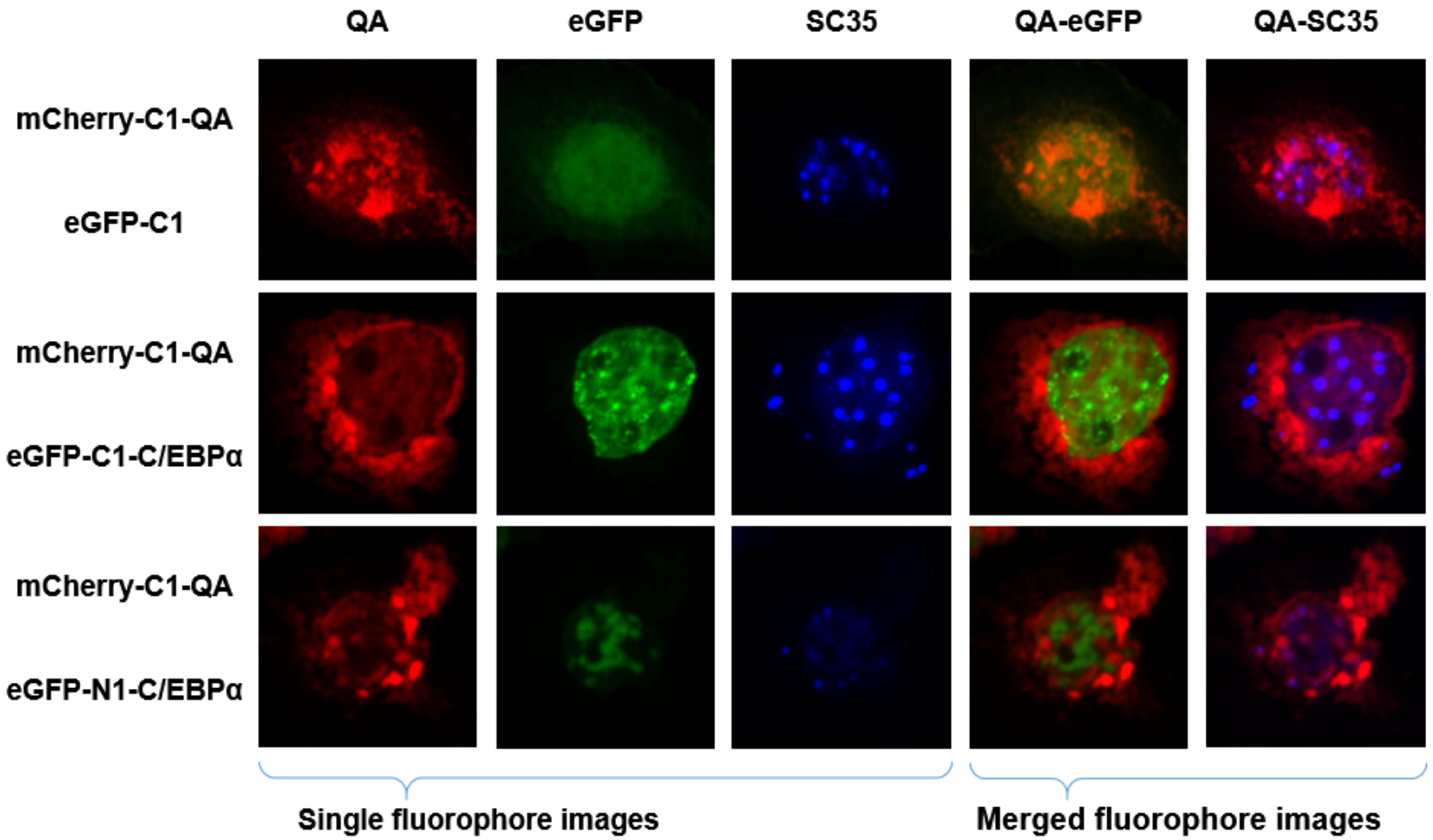

Figure 20 – The QA domain of TCERG1 when fused to the carboxy terminus of mCherry does not act as an NLS nor does it co-localize with C/EBPα.

Confocal localizations of transiently transfected mCherry-C1-QA along with eGFP-C1 +/- C/EBPα or eGFP-N1-C/EBPα in COS7 cells. The left labels describe the transfected plasmids in each row and the top labels describe the fluorophore image or images presented in each column.

Due to the differential localization patterns observed for amino or carboxy terminal tagging of C/EBPα, it was decided to also test the QA fusion to the amino terminus of the mCherry fluorophore. Using an mCherry-N3-QA fusion which fuses the QA domain to the amino terminus of mCherry, the localization patterns of QA were examined in the presence and absence of C/EBPα expression. The first row of figure 21 demonstrates the pattern of QA expression in the absence of C/EBPα expression. Interestingly, in this experiment there was much less distributed fluorescence and more punctate specks of QA expression than shown in figure 20. Similar to the results presented in figure 20, the QA expression was not confined to the nucleus as demonstrated in the QA-eGFP comparison panel nor was it localized to nuclear speckles as demonstrated by the QA-SC35 panel. Once again, the expression of the eGFP was spread evenly across the cell, suggesting no interaction with the QA in this condition. When co-expressed with N1 or C1 eGFP-C/EBPα there was very little change from the control condition. Examining the QA-eGFP expression panels for the last two rows of figure 21 it was observed that the QA localization did not overlap with the localization of C/EBPα, demonstrating that the QA did not interact with C/EBPα, confirming what was seen in figure 20. As well, the QA-SC35 panels indicated that C/EBPα did not overlap with SC35, demonstrating similar to the control condition that the QA is not localized to nuclear speckles.

Taken together, figures 21 and 22 suggest that the isolated QA domain is unable to act as a nuclear localization signal nor is it able to interact with C/EBPα, suggesting that another domain inside TCERG1 may be helping the QA to mediate the interaction with C/EBPα.

It was hypothesized that the lack of interaction between the QA domain and C/EBPα may be due to the fact that the mCherry-QA construct was not exclusively localized to the nucleus. To test this hypothesis a synthetic 3xSV40-NLS signal was fused between the mCherry epitope and the QA domain (see figure 22). The Cherry-3xNLS-QA was then tested in an experiment similar to figures 21 and 22 to determine if directing the QA fusion to the nucleus was able to mediate an interaction with C/EBPα.

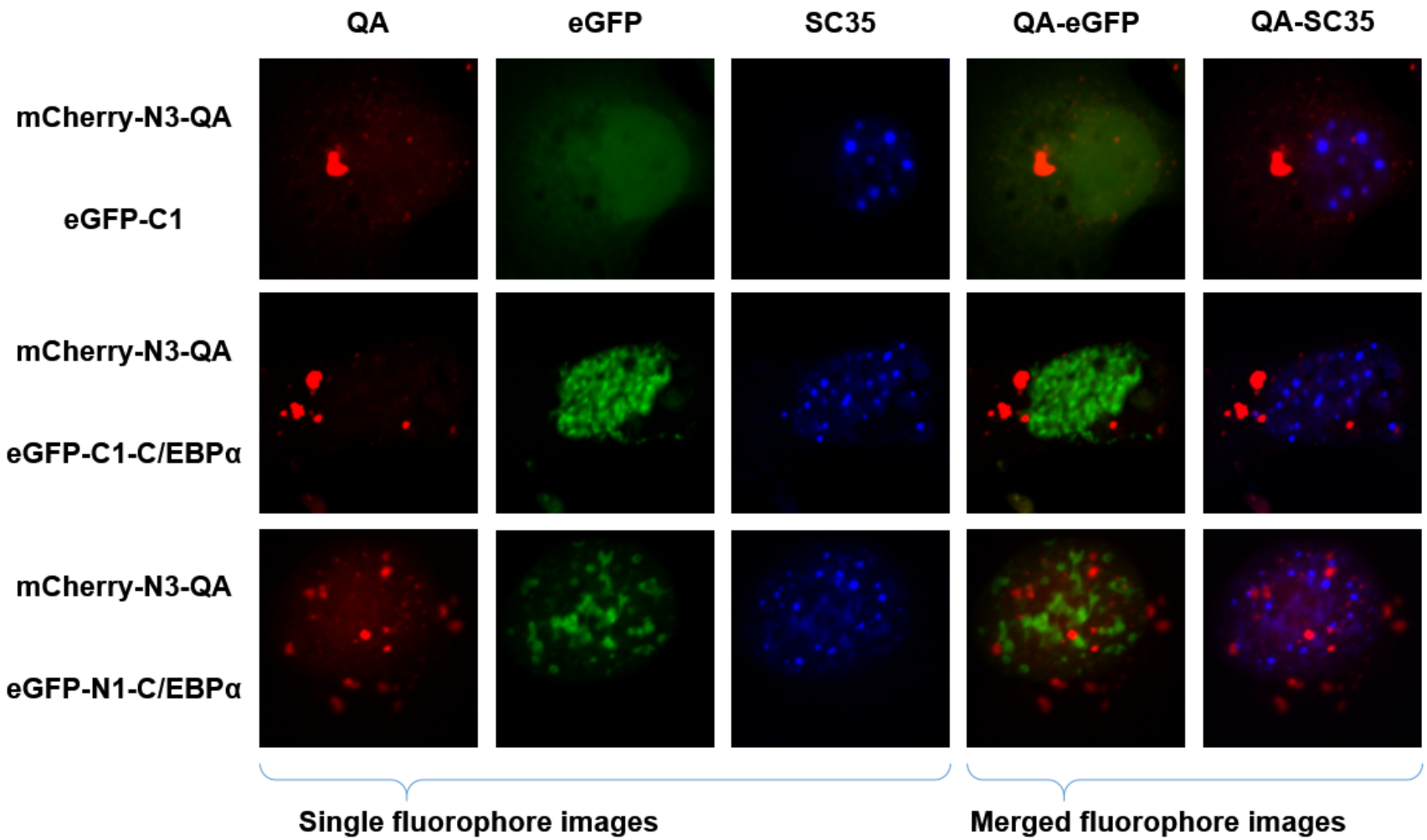

Figure 21 – Fusing the QA domain to the amino terminus of mCherry does not change the localization of the isolated QA domain.

Confocal localizations of mCherry-N3-QA in the presence or absence of eGFP-C1-C/EBPα or eGFP-N1-C/EBPα in COS7 cells. The left labels describe the transfected plasmids in each row and the top labels describe the fluorophore image or images presented in each column.

The first row of figure 22 demonstrates the control condition without C/EBPα expression. By examining the signals present in the QA-eGFP merged panel it can be seen that the QA signal was now localized to the nucleus. Furthermore, the eGFP signal is evenly dispersed across the cell, once again demonstrating no interaction with the QA. The QA-SC35 panel for the control condition demonstrated that the QA did not localize to nuclear speckles.

The second and third rows for figure 22 show experiments where eGFP-C/EBPα was expressed with the QA to determine if amino or carboxy-fused C/EBPα were able to interact with the nuclear localized QA. When examining the merged images for QA-eGFP it can be seen in both C/EBPα conditions that no overlap of signal was present. This demonstrates that neither amino or carboxyl terminal tagging of C/EBPα affected its ability to interact with the mCherry-QA, furthermore, suggesting once again that the isolated QA domain and C/EBPα do not interact. It is also noted in these conditions that the pattern of fluorescence of the mCherry-QA switched to a more punctate, mostly round configuration as compared to what was previously observed in figures 20 and 21. Similar to the test condition, in the QA-SC35 comparison images there was no overlap of signal between the QA and SC35, indicating that the isolated QA domain does not localize to nuclear speckles.

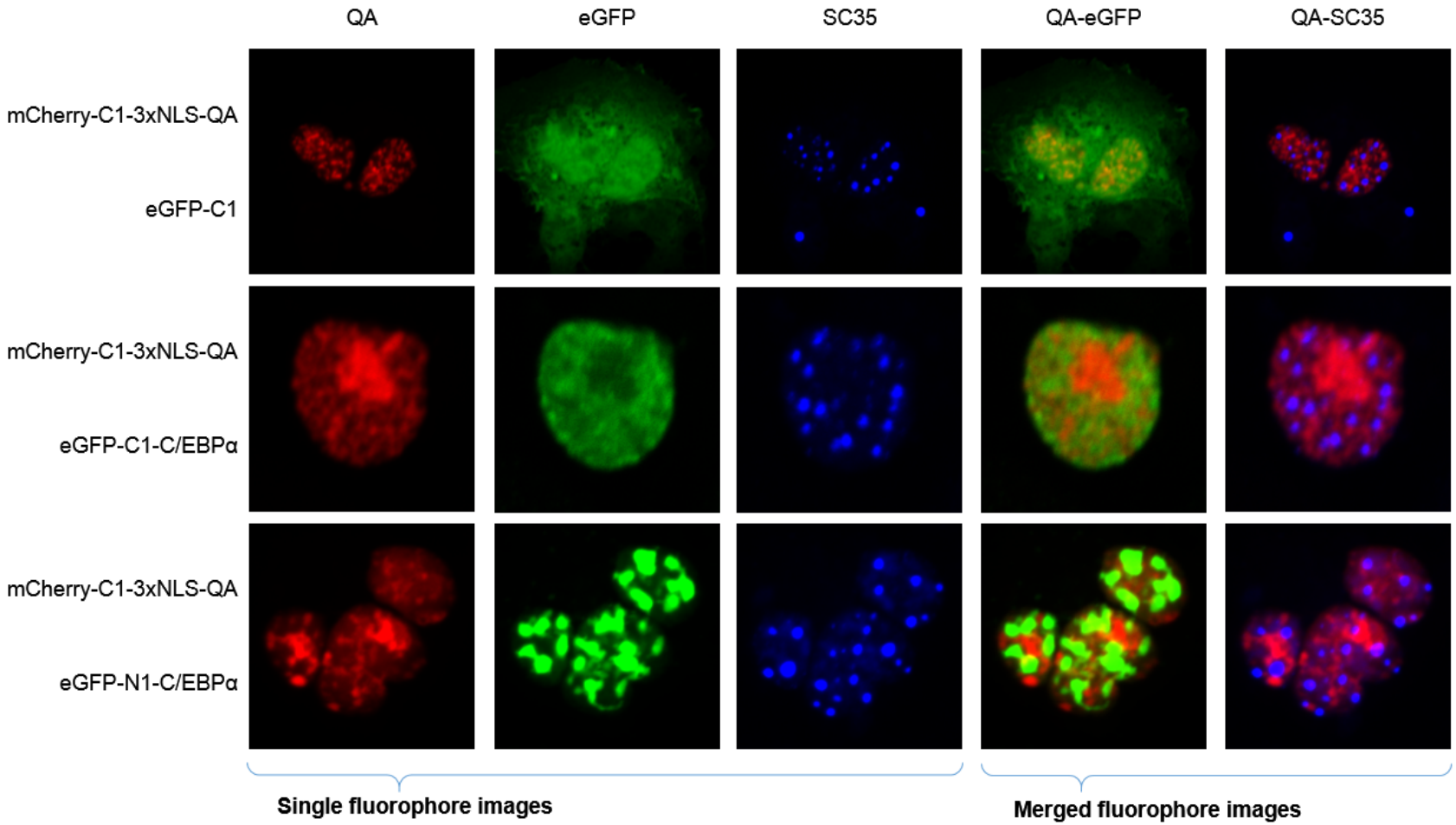

Figure 22 – Even when localized to the nucleus, the QA domain alone does not co-localize with C/EBPα.

Confocal localizations of transiently transfected mCherry-C1-3xNLS-QA along with eGFP-C1, eGFP-C1-C/EBPα or eGFP-N1-C/EBPα in COS7 cells. The left labels describe the transfected plasmids in each row and the top labels describe the fluorophore image or images presented in each column.

4.4 The QA Domain Does Not Play a Role in the Transactivation Inhibition of C/EBPα by TCERG1

C/EBPα is a well-known activator of transcription. TCERG1 has previously been demonstrated to be an inhibitor of C/EBPα-mediated transactivation. We strove to determine whether the QA domain of TCERG1 was involved in this inhibition.

Moazed *et al.,* (2010) suggested that the domain which mediates the transactivation inhibition of C/EBPα was located within the amino half of TCERG1, which is where the QA domain is located. Since the QA domain had been demonstrated in this thesis to be involved in C/EBPα growth arrest inhibition (see section 4.1), the relocalization of TCERG1 (see section 4.2.1) and the physical interaction between TCERG1 and C/EBPα (see section 4.3.1), this seemed like a logical domain to examine. The same firefly reporter gene as described in Banman *et al.,* (2010) and Moazed *et al.,* (2011), the -68Fx4-luc, which consists of a highly C/EBPα responsive promoter fused to a luciferase reporter gene was used in these experiments. A CMV promoter driven renilla reporter gene was co-transfected and a dual luciferase kit from Promega was then used to track transfection efficiency in all conditions. The results presented in figures 24 and 25 were all normalized using the renilla reporter gene to account for transfection efficiency differences. Each condition tested was assayed without C/EBPα to ensure the activation of the reporter gene was specific. As shown in figure 23, without C/EBPα expression, the -68Fx4-luc reporter gene showed only background level, low activity. When C/EBPα was co-expressed, an approximately 37 fold induction of luciferase activity from the same condition without C/EBPα was observed. When TCERG1 was co-expressed, it was demonstrated that C/EBPα-mediated activation of the reporter gene was inhibited approximately 60% from the control condition or a level of about 13 fold activation of the reporter gene above baseline. Sequential deletions of the QA domain from TCERG1 had no observable impact on the ability of TCERG1 to repress C/EBPα-mediated transactivation. These data suggest that the QA domain does not play a role in the transcription inhibition potential of TCERG1 on C/EBPα.

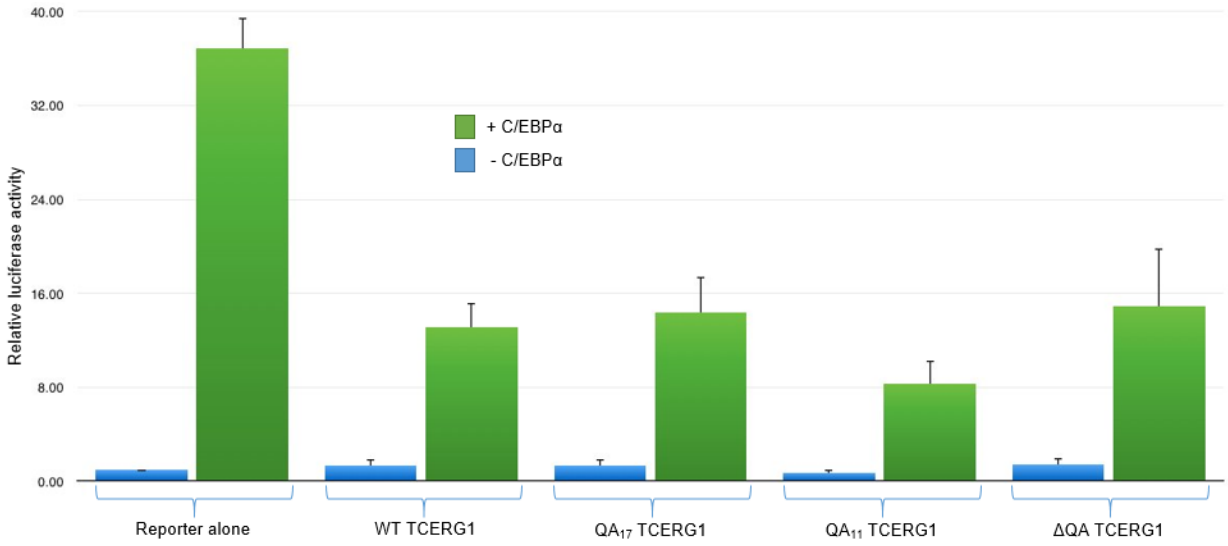

Figure 23 – The QA domain of TCERG1 is not required for the transactivation inhibition of C/EBPα by TCERG1.

Dual luciferase assay using the -68Fx4-luc reporter gene along with pGL4.71-CMV expressing firefly and renilla luciferase respectively in transiently transfected COS7 cells. Both reporter genes were transfected into each condition tested in the experiment. Empty pBOST7 plasmid was transfected with and without FLAG-N1-C/EBPα to determine baseline activity of -68Fx4-luc for the "reporter alone" condition. BOST7-TCERG1-WT, BOST7-TCERG1-QA_{20}, BOST7-TCERG1-QA_{10} and BOST7-TCERG1-ΔQA were transiently transfected along with or without FLAG-N1-C/EBPα in the last four conditions. Results were normalized to "1" being "reporter alone" without C/EBPα. Data is presented as the mean of three experiments with standard error indicated.

4.4.1 The C/EBPα Epitope Tag Does Not Affect the Ability of TCERG1 to Inhibit C/EBPα Transactivation Potential.

Due to the findings that the epitope tag and location on C/EBPα can affect some of the interaction properties of C/EBPα and TCERG1 (see section 4.2.4), an experiment was undertaken to determine if the results presented in figure 23 were influenced by the FLAG-N1-C/EBPα used in those experiments. In this experiment a variety of epitope tags were fused to the amino and carboxy terminus of C/EBPα (see figure 8) and then were assayed using the Promega dual luciferase kit using -68Fx4-luc, pGL4.71-CMV and using WT TCERG1, 21-680 TCERG1, or Δ1-611 TCERG1 as the TCERG1 test conditions (Panel B, Figure 24).

In order to initially determine activation levels of the various C/EBPα plasmids expressed the data was normalized with an RLU of "1" being the background expression levels of the reporter gene plasmids. As shown in Panel A of figure 24, each of the C/EBPα fusion

protein expression plasmids were examined for their ability to induce expression of the -68Fx4-luc reporter gene. Each of the C/EBPα fusion proteins were able to activate the reporter gene with fold activation levels from 34 to 113 from baseline levels. The FLAG epitope tag displayed consistently strong activation in both N1 and C1 versions with 64 and 62 times activation. The eGFP fusion plasmids produced higher activation levels with the C1 fusion activating 82 times above baseline levels and the N1 fusion producing the strongest signal of those tested at 113 fold activation above baseline values. The AcGFP-IRES-C/EBPα produced a very strong activation of expression with 88 fold above baseline values. Alternatively, the eGFP-P2A-C/EBPα activated with the second lowest values among those tested with 47 fold above baseline values. Lastly, the HA fusion produced the lowest activation among all the tested conditions with 34 fold activation above baseline values. Overall, each of the C/EBPα expression plasmids tested were able to effectively activate expression of the -68Fx4-luc reporter gene above baseline levels.

The range of activation values observed could be due to efficiency of transfection or levels of expression for each plasmid expressed. Therefore, the data would be better viewed by normalizing the activation for each C/EBPα fusion plasmid with "1" being the C/EBPα activated reporter gene prior to the addition of TCERG1. Since TCERG1 has been classified as an inhibitor of C/EBPα-mediated transactivation this allows the uninhibited activation levels of C/EBPα for each condition to be considered 100% and therefore comparisons can be made from this point as to the subsequent activation levels of the reporter gene by C/EBPα from the uninhibited condition. These results are presented in the "B" panel of figure 24. As in figure 23, there was no significant activation of the -68Fx4-luc promoter without co-expression of C/EBPα.

Regardless of the C/EBPα expressed, when WT TCERG1 was co-expressed with C/EBPα it was demonstrated that there was an inhibition of C/EBPα-mediated transactivation of the -68Fx4-luc reporter gene. FLAG-C1-C/EBPα, eGFP-C1-C/EBPα, GFP-IRES-C/EBPα and eGFP-P2A-C/EBPα all demonstrated 45%-to 31% levels of activation of the reporter gene with respect to the "reporter +C/EBPα" condition whereas FLAG-N1-C/EBPα, eGFP-N1-C/EBPα and HA-C1-C/EBPα were 70%, 69% and 78%, respectively. These results suggest that amino tagged versions or untagged versions of C/EBPα are able to be more effectively inhibited by TCERG1 than carboxy tagged versions with the sole exception of HA-C1-C/EBPα.

Since deleting the QA domain had no effect on the ability for TCERG1 to repress the transactivational activation ability of C/EBPα it was decided to use deletion mutants of TCERG1 similar to what Moazed *et al.*, (2011) used. In these experiments an amino terminal fragment of TCERG1 was used in the form of 21-680 TCERG1 and a carboxy terminal fragment of TCERG1 was used in the form of Δ1-611 TCERG1 in which the first 611 amino acids was deleted. These deletion mutants were chosen since the NLS of TCERG1 is located close to the middle of the protein (refer to figure 3) thus both mutants contain the TCERG1 NLS.

As shown in panel B of figure 24, it was discovered that the 21-680 TCERG1 mutant was unable to inhibit the transactivation activation potential of C/EBPα. Regardless of the C/EBPα expressed, with the exception of eGFP-C1-C/EBPα, all conditions containing 21-680 TCERG1 reduced the levels of C/EBPα-mediated transcriptional activation levels to close to or greater than the luciferase activity of the reporter +C/EBPα condition. Although the eGFP-C1-C/EBPα condition only had 62% of the activation compared to the reporter +C/EBPα condition it still displayed double the activation of the WT TCERG1 condition. All conditions presented at least approximately 2X the activation of the reporter gene when compared to the activation of the WT TCERG1 conditions. This demonstrates that there was minimal inhibition of C/EBPα-mediated transactivation compared to the WT TCERG1.

The carboxy terminal mutant of TCERG, the Δ1-611 TCERG1, demonstrated C/EBPα-mediated activation levels which across most conditions were fairly similar to those observed with the WT TCERG1 conditions. There were two conditions which did not follow these trends in the cases of eGFP-C1-C/EBPα and eGFP-P2A-C/EBPα. The activation levels of eGFP-C1-C/EBPα were similar in both the 21-680 TCERG1 and Δ1-611 TCERG1 conditions. The eGFP fluorophore is a fairly large tag and may be interfering with proper functioning of C/EBPα in this condition. eGFP-P2A-C/EBPα demonstrated a complete lack of inhibition by TCERG1 with levels equivalent to those seen with the reporter alone with C/EBPα when expressed with TCERG1-Δ1-611. Although this condition had higher levels of activation compared to TCERG1 WT the luciferase levels are still lower than TCERG1 21-680, which indicating that some inhibition of transactivation still occurred.

Taken together, these results demonstrate that although there are small variations in transactivation potential inhibition of C/EBPα by TCERG1 caused by differential epitope

tagging, when using small epitope tags the overall trends are similar. These results also suggest that the domain which is mediating the inhibition of transactivation of C/EBPα by TCERG1 is somewhere in the carboxy terminus half of TCERG1, potentially the FF domains which are known protein interaction domains. These results were unexpected since the previous work by Moazed *et al.,* (2010) had suggested that the domain in TCERG1 mediating C/EBPα transactivation inhibition was located within the amino terminus of TCERG1.

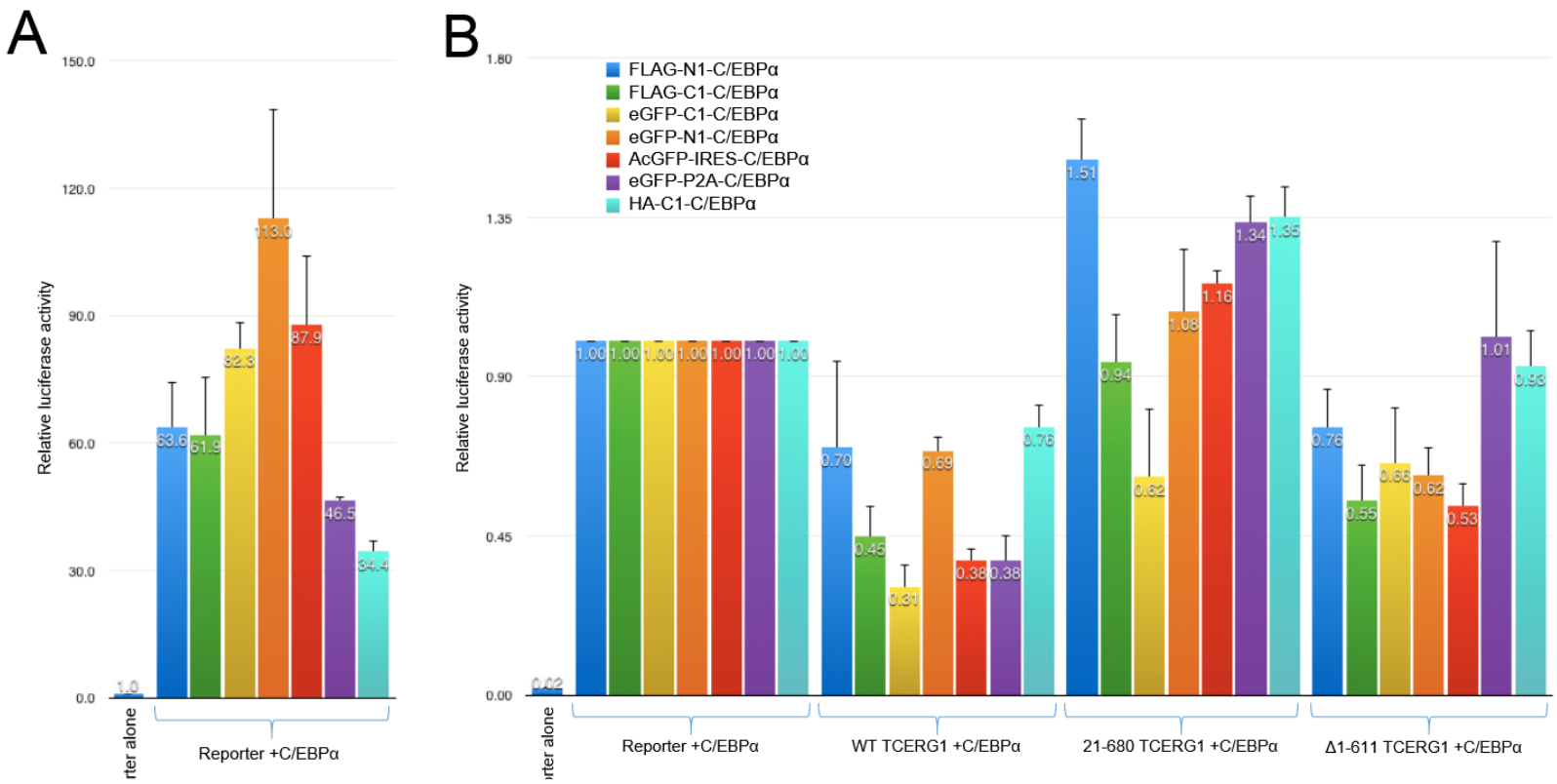

Figure 24 – Differential epitope tagging of TCERG1 does not affect the transactivation inhibition of C/EBPα using TCERG1 and amino and carboxy terminal deletion mutants.

Dual luciferase assay using the -68Fx4-luc reporter gene co-expressed with pGL4.71-CMV expressing firefly and renilla luciferase respectively in transiently transfected COS7 cells. Empty pBOST7 plasmid was co-transfected with and without the addition of the epitope tagged versions of C/EBPα under the reporter alone and reporter +C/EBPα conditions. Plasmids expressing WT TCERG1, 21-680 TCERG1, or Δ1-611 TCERG1 were transiently co-transfected along with the differential tagged C/EBPα in the last 3 conditions of panel "B", respectively. In orders to observe the results in regards to the activation potential of the -68Fx4-luc promoter in regards to the C/EBPα used in each condition, in panel "A" results are normalized to "1" being the reporter alone condition. In panel "B" the results are normalized to "1" for the pBOST7 with C/EBPα for each condition. Data is presented as the mean of three experiments with standard error indicated.

5. Discussion

5.1 The Growth Arrest and Transactivation Inhibition of C/EBPα by TCERG1 are Mediated by Different Domains in TCERG1

The inhibition of C/EBPα-mediated growth arrest by TCERG1 had originally been characterized to be in a region encompassing amino acids 32-668 of TCERG1 (Banman *et al.*, 2010). As well, the C/EBPα-mediated relocalization domain in TCERG1 had been localized to this same region. In this thesis, the results presented in section 4.1 demonstrated that deletion of the QA domain in TCERG1 rendered it unable to reverse C/EBPα-mediated growth arrest. Furthermore, the data presented in section 4.2.1 demonstrated that the relocalization of TCERG1 to pericentromeric regions did not occur when the QA domain was deleted. Both of these findings suggest a role of the QA domain being important for these activities. Interestingly, it was discovered that while the relocalization of TCERG1 was QA-dependent using an amino fused eGFP-C1-C/EBPα, the relocalization was not QA specific in the case of a carboxy terminal fused eGFP-N1-C/EBPα wherein WT TCERG1 as well as ΔQA TCERG1 were able to become relocalized (see section 4.2.4). The implications of this are discussed further in section 5.4.

The domain of TCERG1 which inhibited the transactivation ability of C/EBPα had originally been characterized to the same 32-668 amino acids of TCERG1 (Moazed *et al.*, 2011). In this thesis it was found that the QA domain was not the domain which mediated the transactivation inhibition of C/EBPα (see section 4.4). Furthermore, figure 24 demonstrated that the C/EBPα inhibitory activities appeared to be located in the carboxy terminal half of TCERG1 rather than the amino terminal half as previously reported. This is further discussed in section 5.3.

Taken together, these data would suggest that the growth arrest inhibitory and transactivation inhibitory activities of TCERG1 on C/EBPα reside in different domains of TCERG1. These two activities are mediated by different domains in C/EBPα as well. The growth arrest domain of C/EBPα is located around the S193 residue whereas the transactivation domains are located on the amino half of C/EBPα. Since both TCERG1 and C/EBPα contain a

separate domain to control each interaction it could be suggested that each activity could be controlled and regulated separately. The complexities of domain structure of C/EBPα and TCERG1 on the function of each protein is further discussed in section 5.5.

5.2 The TCERG1 QA Repeat Domain is Important for Inhibitory Activity Toward C/EBPα

5.2.1 The Relocalization of TCERG1 by C/EBPα is QA Length Dependent

The relocalization of TCERG1 induced by C/EBPα was initially characterized by Banman *et al.*, (2010) when they noticed that TCERG1 became relocated from the nuclear speckle compartments to the pericentromeric domains where C/EBPα resides. Furthermore, they narrowed down the domain responsible for relocalization to being somewhere in the amino terminus of TCERG1. In this thesis we were able to further refine this site of interaction to being dependent upon the QA domain of TCERG1. Furthermore, it was discovered that the relocalization event required between 11 and 20 QA repeats for the relocalization event to take place. Since the relocalization and growth arrest properties appear to be dependent upon the QA domain it could be hypothesized that they are inter-related. By moving into the pericentromeric domains with C/EBPα, this would allow TCERG1 to inhibit the growth arrest properties of C/EBPα as discussed in section 5.5. Since they are both QA-dependent, by integrating the findings of the growth arrest and relocalization experiments it can be suggested that they are accomplishing the same goal of inhibiting C/EBPα-mediated growth arrest. The difference between these experiments would be that the growth arrest experiment (section 4.1) is the physiological consequence and the relocalization experiment (section 4.2.1) is the physical manifestation of this interaction. The results from Moazed *et al.*, (2011) agree with this hypothesis as they were able to demonstrate that a V296A mutant of C/EBPα, which is dispersed throughout the nucleus was still able to be inhibited by TCERG1. As well, they demonstrated that TCERG1 adopts a dispersed pattern of expression when co-expressed with V296A C/EBPα. Therefore, this suggests that the relocalization and growth arrest inhibitory properties are related.

The data described in section 4.1 indicated that the C/EBPα-mediated growth arrest inhibitory activities of TCERG1 are QA length dependent. It was discovered that somewhere between 11 and 20 QA repeats that TCERG1 were required to inhibit the growth arrest functions of C/EBPα. The results from the relocalization experiments reported in section 4.2 agree with these results. As well, the Co-IP experiment described in section 4.3.1 demonstrated that as the number of QA repeats decrease, so does the interaction between TCERG1 and C/EBPα.

All three of these experiments suggest that not only is the QA domain important for the functions performed but there is also a dependence upon the length of the QA domain. There is only one other known QA containing protein in humans, zinc nuclear factor 384 (ZNF384). This protein contains a smaller QA domain compared to TCERG1 at 14 QA repeats in humans. Interestingly, 14 repeats is within the 11 to 20 QA repeats identified previously for proper functioning of the QA domain in the relocalization and growth arrest inhibition of C/EBPα. While there are no links currently between ZNF384 and C/EBPα or TCERG1, as our knowledge about the functions of the QA domain increases it could be of importance to keep ZNF384 in mind. Since the QA domain has been implicated in this thesis as being required for protein relocalization as well as protein inhibition, the QA domain of ZNF384 may be mediating similar activities for other proteins in the cell.

The evolutionary aspects of the QA domain as presented in section 2.4.4.1 are also interesting to note in terms of the information gathered in this thesis. It was noticed in figure 5 that there was a sharp increase of QA repeats in the evolution of the TCERG1 QA domain. It was noted that with the exception of chickens, the evolution of complex multicellular eukaryotes with complex splicing reactions that the QA domain expanded rapidly from one or two copies to 20 copies. This rapid expansion to above the 11-20 QA repeat threshold identified in the previously mentioned experiments suggests that the expanded QA domain conferred an evolutionary advantage to the higher organisms. The chicken, without an expanded QA repeat section may provide an interesting contrast to determine further function of the QA domain.

5.2.2 The Isolated QA Domain is Unable to Target Proteins to the Nucleus, Nuclear Speckles or Mediate Relocalization by C/EBPα

Arango *et al.,* (2006) suggested that the QA domain of TCERG1 was able to localize TCERG1 to the nucleus. The results presented in this thesis disagree with these findings. If the QA domain acted as an NLS, the ΔQA TCERG1 signal in figures 10 and 11 would likely not be localized to the nucleus. Additionally, the experiments detailed in section 4.3.2 using the isolated QA domain would have not required the use of a synthetic NLS to be added to the protein to obtain nuclear localization. Through the use of both deletion of the QA domain as well as expression of the isolated QA domain it has been demonstrated that the QA domain does not act as an NLS.

Further, the data presented in section 4.3.2 indicated that the mCherry-QA fusion did not appear to be sufficient to interact with C/EBPα or nuclear speckles inside the cells. Neither the N1 nor C1 fusions of eGFP-C/EBPα were also able to mediate co-localization. As well as not localizing with C/EBPα it was further observed that the QA fusions were also unable to efficiently localize to nuclear speckles as immunostained using the nuclear speckle marker, SC35. This result is not surprising since it has been suggested that the FF4 and FF5 domains of TCERG1 control nuclear speckle localization inside the nucleus (Sanchez-Hernandez *et al.,* 2012). Interestingly, the QA domain fusion does appear to mediate some form of protein aggregation or speckling as the fluorescence appears to form areas of higher fluorescence inside the cell at several different points.

As discussed in section 4.3.2, the isolated QA domain was unable to relocalize to C/EBPα when in isolation. This suggests that there is another domain or multiple domains inside TCERG1 which is helping the QA mediate it's interactions with C/EBPα since there are several experiments performed in this thesis which suggest that C/EBPα interact through the QA domain.

Insights into this can be obtained from results presented by Banman *et al.,* (2010). In this paper they demonstrated that an amino terminal fragment of TCERG1 which contained the QA domain, amino acids 32-293, was unable to mediate the relocalization of TCERG1 when expressed with C/EBPα. The fragment of 32-293 contained the polyproline region, WW1, the QA domain and approximately half of the STP domain (see figure 3). Alternatively, they also showed that a fragment of TCERG1 containing amino acids 281-1098 also did not relocalize

correctly. This fragment did not contain the QA domain and contained a slightly larger fragment of the STP domain (see figure 3). The results of this thesis can provide insights into both of these observations. In the first fragment, the 32-293, TCERG1 was unable to mediate relocalization, even though the QA domain was present, similar to the results provided by the experiments performed in section 4.3.2. Both the paper and these results suggest that the QA domain requires another domain inside TCERG1 to function properly. The results suggest that the STP domain could be functioning in this capacity. In the 32-293 fragment, the STP domain was cut in half and therefore may not have been large enough to facilitate the QA domain functionality. The experiments described in sections 4.3.2 in this thesis which use the isolated QA domain were unable to be relocalized by C/EBPα, further suggesting that there was another domain required to help the QA domain. The second piece of evidence which supports the hypothesis that it is the STP domain which is helping the QA domain function correctly comes from the 281-1098 mutant presented in Banman *et al.*, (2010). In this mutant they were unable to obtain proper relocalization of TCERG1. Although it appears there was a slight amount of relocalization of TCERG1, the majority of the signal remained inside the nuclear speckles. The QA domain was deleted in this fragment but the STP domain was mostly intact, from this it could be suggested that the STP domain provides some relocalization properties along with the QA domain but both may be required for the complete relocalization of TCERG1.

The initial two-hybrid clone pulled out by Mcfie *et al.*, (2006) contained amino acids 89-490 of TCERG1. Since this clone was able to successfully interact with C/EBPα this provides further support that there is another domain in close proximity to the QA domain which provides help for it to function. This clone was missing most of the polyproline domain but contained intact WW1, QA, STP, and WW2 domains (see figure 3). Based upon these finding as well as the previous discussed properties of the Banman clones the STP domain appears to be a prime candidate to be providing the QA domain with additional properties to mediate its activities.

5.3 The Carboxy Terminus of TCERG1 Mediates the Transactivation Inhibition of C/EBPα

Moazed *et al.*, (2011) suggested that the C/EBPα-mediated transcriptional activation was inhibited by TCERG1 through a domain in the amino terminus of TCERG1. Since the QA domain of TCERG1 had been demonstrated to play a role in the C/EBPα-mediated relocalization of TCERG1 and growth arrest inhibition of C/EBPα by TCERG1 it was initially tested to determine if there was any role of the QA domain involved in the transcriptional inhibition of C/EBPα. The experiments performed in figure 23 suggests that the QA domain is not involved in this activity of TCERG1. Across all conditions it was demonstrated that there was no impact of deleting the QA domain upon the transcriptional inhibition of C/EBPα.

Furthermore, it was discovered that not only was there no requirement for the QA domain of TCERG1 for this inhibition, that indeed the functional interaction site between TCERG1 and C/EBPα for transcriptional inhibition appeared to be the carboxy terminus of TCERG1. As demonstrated in panel B of figure 24, across all the conditions it was demonstrated that the amino terminal fragment of TCERG1 was unable to inhibit C/EBPα transactivation, whereas the carboxy terminal fragment was able to inhibit. These data suggest that the TCERG1 inhibitory domain is most likely located in the carboxy terminus of TCERG1. The carboxy terminus of TCERG1 contains five FF domains which have been implicated in protein interactions (Carty *et al.*, 2000; Sanchez-Hernandez *et al.*, 2012; Smith *et al.*, 2004). Therefore, it can be hypothesized that this activity of C/EBPα is mediated by one or more of the TCERG1 FF domains.

Interestingly, even though the carboxy terminal half of TCERG1 appears to be able to affect the transactivation ability of C/EBPα, according to the Co-IP results presented in figure 18 the physical interaction between TCERG1 and C/EBPα requires the QA domain. Furthermore, unpublished Co-IP results demonstrated that there was no interaction between TCERG1 and C/EBPα when using a carboxy terminal fragment of TCERG1 (Sheng-Pin Hsiao, unpublished observations). These results suggest that the transactivation inhibition of C/EBPα occurs through another mechanism in which the proteins do not interact directly. The mechanism of how this occurs is currently unknown, although since FF domains are known protein interaction domains it is hypothesized that TCERG1 is able to recruit a protein which is then able to inhibit the transactivation ability of C/EBPα.

These results disagree with the results presented by Moazed *et al.*, (2011) in which they claimed the domain responsible for the inhibition of C/EBPα-mediated transactivation lay in the amino half of TCERG1. Through the use of the dual luciferase kit and Western blotting of the protein extracts we are confident in the results presented here as any variations in the expression levels of each product in the reaction should have been compensated by the expression of the renilla reporter gene and therefore were taken into account in the results. We did not use the exact same TCERG1 plasmid constructs as in the prior publication so there could have been variation introduced as to how the plasmids were created or the proteins expressed. As well, our amino terminus mutant of TCERG1 contained the endogenous NLS of TCERG1, whereas in the Moazed *et* al. paper the TCERG1 protein contained a synthetic NLS which was not in the original protein, which could have affected the results. As with all deletions of proteins we also have to consider that any deletion, even if it is very small, may have disturbed proper folding of the protein and therefore its interaction properties. As well, although we took great care to replicate the prior results, there always is the possibility of cell line specific changes due to passage number or handling which produced the prior results so even though we may have been unable to reproduce them, the previous results may provide insights later.

5.4 The Epitope Tagging of C/EBPα Has Implications For its Function

Whereas untagged versions of the proteins of interest will continue to be the gold standard when performing experiments, sometimes there is a need to use epitope tags to perform specific experiments. As discovered in sections 4.2.4 and 4.4.1, the position of the epitope tag on C/EBPα appears to interfere with different aspects of the function of C/EBPα. The structure of C/EBPα, as shown in figure 1, is such that there are binding domains on each terminus of the protein. The amino terminus of C/EBPα acts as a transactivation domain, binding other proteins to effect the regulation of genes. The carboxy terminus of C/EBPα contains a bZIP, DNA binding domain.

It has been hypothesized that the QA-mediated, C/EBPα-mediated growth arrest inhibition and the TCERG1 relocalization properties are two aspects of the same process as discussed in section 5.2.1. Furthermore, it was discussed in section 4.2.4 that the ability of

C/EBPα to relocalize TCERG1 in a QA-dependent manner could be modified by fusing the eGFP epitope tag to either the carboxy or amino terminus of C/EBPα. By fusing the eGFP to the amino terminus of C/EBPα it was demonstrated that TCERG1 was relocalized in a QA dependent manner. Furthermore, by fusing the eGFP to the carboxy terminus of C/EBPα, the relocalization of TCERG1 occurred in QA independent manner. Further examination using untagged C/EBPα concluded that the QA dependent relocalization of TCERG1 appears to be the biologically relevant activity of C/EBPα. This conclusion would suggest that the eGFP when fused to the carboxy terminus of C/EBPα is somehow blocking the activity which mediates the QA specific relocalization of TCERG1. Furthermore, in section 4.3.1 it was described that the QA domain appears to mediate the physical interaction between C/EBPα and TCERG1, suggesting that the relocalization may actually be mediated by a protein other than C/EBPα. If the QA domain is the sole mediating factor for the relocalization, when the physical interaction is disrupted by deleting the QA domain, the relocalization should no longer take place. Since this is not that case it suggests that there is more to this interaction which has not been discovered. Determinations of the differences between the protein complexes which form when WT TCERG1 or ΔQA TCERG1 may help to determine the other proteins playing a role in this interaction.

The other major activity which was tested to determine epitope tagging issues was the transactivational activity of C/EBPα. The transactivation activities of C/EBPα are located on the amino terminal half of C/EBPα. The experiments performed in section 4.4.1 demonstrated that epitope tagging can have implications on the transactivational ability of C/EBPα. While the rest of the epitope tags tested had minimal effects on C/EBPα, one of the eGFP-C/EBPα fusions caused the C/EBPα to interact with TCERG1 differently. When eGFP was fused to the carboxy terminal of C/EBPα (eGFP-N1-C/EBPα), this fusion protein was able to activate the reporter gene similarly to the other constructs tested; however, the amino terminal fusion did not. While WT TCERG effectively inhibited the eGFP-C1-C/EBPα, there was no difference between the inhibition patterns for the 21-680 TCERG1 as well as the Δ1-611 TCERG1. In all of the other conditions tested, the WT TCERG1 and the Δ1-611 TCERG1 inhibited the C/EBPα-mediated transactivation while the 21-680 TCERG1 was unable to repress the C/EBPα-mediated transactivation. This suggests that the large eGFP epitope is interfering with C/EBPα and is causing this fusion protein to behave different from what would be expected.

Taken together, this suggests that care must be taken when using epitope tags with C/EBPα as in this thesis it was demonstrated that fusing a fluorescent tag to the amino terminus of C/EBPα disrupted the transactivation properties of C/EBPα with respect to TCERG1 (see section 4.4.1) and when fused to the carboxy terminus, the QA-dependent relocalization properties of TCERG1 was affected (see section 4.2.2). This is especially true when using larger epitope tags such as eGFP since it appears that smaller epitope tags such as FLAG or HA have a lesser effect on the optimal function of the protein. Although the smaller epitope tags have a lesser effect upon the function of C/EBPα, care must still be taken that the epitope tags being used are not adversely affecting protein function.

5.5 Complex Domain Structure Provides Opportunities to Coordinate Different Nuclear Events

Both TCERG1 and C/EBPα contain many different domains. Of the two proteins, TCERG1 is far less characterized than C/EBPα and as such there are many unknowns in terms of what functions these domains mediate.

Even though C/EBPα has been the best characterized of the two proteins, there are still many unknowns as to how precisely the domains of C/EBPα function (see section 2.3.1). It appears that in C/EBPα there are three separate regions which mediate different processes, each containing one or more domains. The amino terminus is involved in protein-protein interactions, the central part of C/EBPα is involved in growth arrest and the carboxy terminus is involved in sequence-specific DNA binding.

The TCERG1 domain structure, alternatively, remains relatively unexplored. While there has been investigation on the general mechanisms of transcription and splicing that TCERG1 mediates, the majority of the protein remains completely unknown in terms of function. Thus far, only some of the WW and FF domains have been explored and characterized (see section 2.4). No functions have been described for the polyproline, STP and KE domains (see figure 3).

The QA domain researched in this thesis had only been briefly mentioned previously in the literature as a potential NLS and more importantly as a potential determinant of age of onset of Huntington's disease (see section 5.5). The domains surrounding the QA domain, the poly proline and the STP domain remain unexplored in terms of function.

This thesis makes use of deletion mutants of the QA domain of TCERG1, and it is recognized that deletion or shortening of the TCERG1 could induce conformational changes in other parts of the protein that may have contributed to the observations made. These conformational changes have the potential to disrupt native protein-protein interactions by changing the properties from the native protein. Crystal structures of wild type TCERG and the ΔQA mutant would need to be obtained in order to fully address this possibility.

Although there haven't been any structures solved for a polyQA domain there has been some research undertaken on polyQ and polyA domains. While there is still no consensus as to the conformation adopted by these domains it is generally accepted that they most likely form β-sheets when a certain length has been exceeded and they are exhibiting amyloid plaque characteristics (Bauer *et al.*, 2011; Pelassa *et al.*, 2014). Alternatively, it has been suggested that these domains form coiled-coil configurations while in normal conformation although, once again, due to issues of solubility these domains have been problematic in producing a crystal structure (Pelassa *et al.*, 2014). Due to similarities of both polyQ and polyA domains, it can be suggested that the conformation of the polyQA domain could therefore be similar to its isolated parts.

Although difficult to research, the complex domain structure and interaction patterns of TCERG1 and C/EBPα allow us to bridge seemingly unrelated mechanisms in the cell and piece together a larger picture of the regulation and interrelation of signaling pathways.

To begin to piece together the interactions between C/EBPα and TCERG1, the reasons for their interactions need to be considered. In considering these reasons, several proposals about the interactions between these two proteins can be made. The first proposal is that TCERG1 is a native regulator of the growth arrest potential of C/EBPα.

C/EBPα is only expressed in cells at very specific time points in the growth cycle of organisms or cells (McKnight *et al.*, 1989). When the levels of C/EBPα become high enough in the cell, the cell undergoes G1/S growth arrest or differentiation, depending upon cellular conditions. Through the recruiting of TCERG1 to C/EBPα, the cell could potentially overcome the growth arrest placed upon it by C/EBPα. By depleting the available pool of TCERG1 this could in turn drive the alternative splicing of transcripts in the cell until the C/EBPα could be effectively degraded or acted upon to regulate another reaction.

The second proposal is that C/EBPα, through its growth arrest properties allows the buildup of transcription factors so the cell can be better prepared to enter the next cell cycle. This arrest would then allow the cell to continue to build up the levels of transcription and splicing factors, one of them being TCERG1. When the cell builds up enough TCERG1 to overcome the growth arrest of C/EBPα, the cell has built up enough transcription/splicing factors and is ready to go into the next part of the cell cycle. Potentially, the depletion of TCERG1 from nuclear speckles may trigger the cell to produce more TCERG1. This buildup of TCERG1 could play an important role when C/EBPα is either released from the pericentromeric regions with TCERG1 bound to it or is degraded. Further testing will need to be performed to determine if either of these hypotheses have merit.

Based upon the observations presented in this thesis along with previous observations, the physical interactions between TCERG1 and C/EBPα can be theorized. The first theorized set of interactions between TCERG1 and C/EBPα could occur between the QA domain of TCERG1 and a currently unknown domain within C/EBPα. Based upon the findings of this thesis we have discovered that deletion of the QA domain is able to abrogate the C/EBPα-mediated relocalization of TCERG1 as well as the C/EBPα-mediated G1/S growth arrest. Assuming that the QA domain is only able to bind to one domain in C/EBPα the most likely candidate for this binding would be the S193 domain in C/EBPα which has previously been demonstrated to mediate the growth arrest properties of C/EBPα (see section 2.3.1). Since the major functional interaction demonstrated between TCERG1 and C/EBPα is the inhibition of C/EBPα-mediated growth arrest it can be hypothesized that the S193 growth arrest domain of C/EBPα is the site of interaction for C/EBPα with the QA domain of TCERG1.

Most likely there is a second site of interaction between C/EBPα and TCERG1 since the epitope tagging of C/EBPα was able to produce differential relocalization properties based upon amino or carboxy tagging of C/EBPα using a large fluorescent epitope (see section 4.2.5). Since the carboxy terminal tagging of C/EBPα produced no QA-specific relocalization it could be assumed that the large epitope tag on the carboxy terminus of C/EBPα was able to block the recognition site which allowed the relocalization to take place in a QA dependent manner. As well, the transactivation inhibition of C/EBPα by TCERG1 appeared to be mediated by a domain in the carboxy terminal half of TCERG1 rather than the QA domain. This suggests that there are two separate domains in TCERG1 that interact with C/EBPα.

By this logic it would suggest that TCERG1 and C/EBPα could be interacting in one of two ways:

1) TCERG1 and C/EBPα interact directly or indirectly through a mediator protein via an unknown domain in the carboxy terminus of TCERG1 with an unknown domain in C/EBPα along with the direct interaction of the TCERG1 QA domain along with likely the S193 region of C/EBPα.

2) TCERG1 and C/EBPα interact similar to 1) but the QA domain interaction is mediated by an external protein which modulates post-translational modifications on one or both of the proteins. The relocalization of TCERG1 and the transcriptional inhibition of C/EBPα are by-products of the actions of this protein or proteins.

Without further testing, a model of the interaction of these two proteins is difficult to propose without making too many assumptions. Furthermore, the work presented in this thesis focused solely upon the domains in TCERG1 that are involved in the interaction. In order to obtain a clearer picture, the domains inside C/EBPα that mediate these interactions will have to be probed as well.

5.6 Summary of Findings Presented in This Thesis

While normally found in nuclear speckle compartments, TCERG1 has been demonstrated in this thesis and prior work by Banman *et al*., 2010 and Moazed *et al*., 2011 that when co-expressed in a cell with C/EBPα that TCERG1 is able to move from the nuclear speckles to where C/EBPα resides; figure 25 illustrates this interaction. Furthermore, this relocalization activity has been demonstrated in this thesis to be dependent upon the QA domain in TCERG1 with between 11 and 20 QA repeats being required for relocalization. When expressed in a cell, C/EBPα has previously been demonstrated to mediate cellular growth arrest. When TCERG1 is co-expressed in the cell it has been demonstrated to inhibit the growth arrest activity of C/EBPα. This thesis was able to demonstrate that the inhibition of C/EBPα-mediated growth arrest activity by TCERG1 was QA dependent. As well, as for the relocalization activity it was demonstrated that between 11 and 20 QA repeats were required for TCERG1 to meditate the inhibition. Through co-immunoprecipitation it was demonstrated in this thesis that the physical interaction between TCERG1 and C/EBPα was also QA dependent

with the interaction becoming weaker as the amount of QA repeats in TCERG1 decreased. Since the relocalization, growth arrest activity, as well as the physical interaction, between TCERG1 and C/EBPα were QA dependent it is hypothesized that all three of these interactions were extensions of each other being the physical, physiological and molecular interactions, respectively. Although only tested for relocalization of TCERG1, there also appears to be a dominant negative affect of the ΔQA. When co-expressed in a cell with ΔQA TCERG1, WT TCERG1 is unable to become relocalized from nuclear speckles to pericentromeric regions by C/EBPα.

Alternatively, the transactivational activities of C/EBPα were demonstrated to be inhibited by the carboxy terminal half of TCERG1, somewhere within amino acids 612-1098. When C/EBPα is expressed in a cell it is able to mediate the transactivation of various genes, but when TCERG1 is co-expressed it inhibits the transactivational activities of C/EBPα. When a carboxy terminal mutant of TCERG1 was co-expressed with C/EBPα, it was observed that there was no inhibition of C/EBPα-mediated transactivation. While this work was unable to precisely determine the domain in TCERG1 that mediates this repression, this activity was mapped to the amino terminal half of TCERG1 which contains 5 FF domains.

This work on the QA domain of TCERG1 is the first time the QA domain of TCERG1 has been assigned a functional role inside the cell. TCERG1 and its function is mostly unknown as there has been little research performed on it. Previously the QA domain had only been implicated in the age of onset of Huntington's disease, but no functional role was assigned in that research. Assigning a functional role to the QA domain opens up the potential to discover the further functionality of the QA domain as it relates to other proteins in the cell. By potentially interacting with and changing the characteristics of proteins in the cell the QA domain could be demonstrated to be an important modulator of function of its interaction partners.

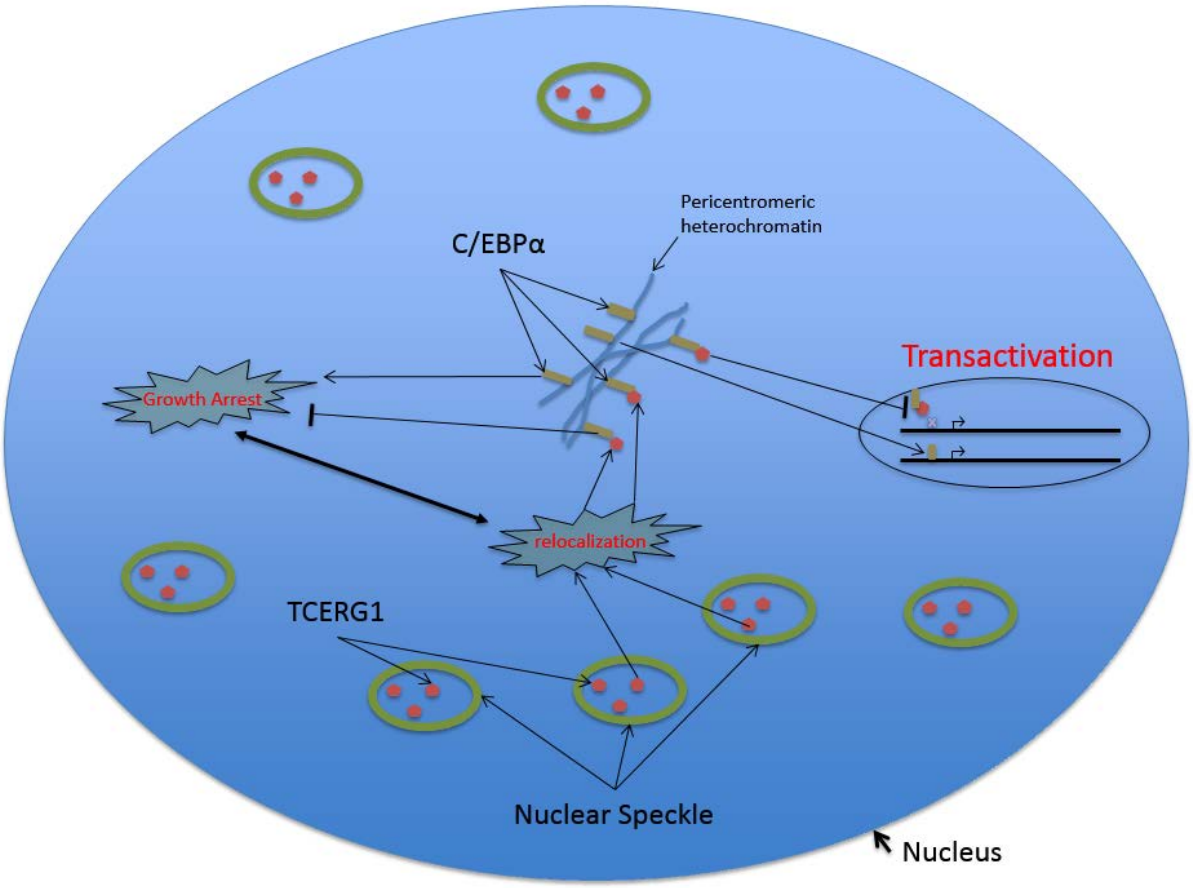

Figure 25 – Representation of the Interactions Between TCERG1 and C/EBPα Explored in This Thesis

6. Future Directions

6.1 Further QA Domain Exploration

There are several unanswered questions arising from the work performed in this thesis. These questions mostly arise from when TCERG1 or C/EBPα suddenly change their interaction properties when an epitope tag is placed on the amino or carboxy terminus of C/EBPα. Initially the domains on either end of C/EBPα would have to be examined to determine if they participate in the interaction between C/EBPα and TCERG1. The first example for this is when the fluorescent epitope tag was placed on the carboxy terminus of C/EBPα, it was observed that the relocalization of TCERG1 was no longer QA-dependent. Secondly, it was observed that when eGFP was fused to the amino terminus of C/EBPα that there was no difference in the transactivation potential from the amino and carboxy terminal halves of TCERG1, whereas all the other tested C/EBPα proteins demonstrated inhibition for the carboxy terminal half of TCERG1.

Following this, it would be suggested that there are other proteins involved in the various interactions between TCERG1 and C/EBPα, whether it be the gain or loss of these proteins following the removal of the QA domain in TCERG1. It would be beneficial to undertake a characteristic study of the proteins involved in the interaction complex between C/EBPα and TCERG1 in the presence and absence of the QA domain. One way to explore this question would be to identify proteins that interact with TCERG1 but not with the ΔQA mutant (or vice versa), using a combination of immunoprecipitation and mass spectrometry methodologies.

Determining whether proteins are gain or lost in the TCERG1-C/EBPα complex with and without the QA domain present in TCERG1 could help identify proteins which are involved in the interactions such as relocalization. The dominant negative qualities of the TCERG1 ΔQA protein could be of interest as well. Although this thesis never addressed the further implications of this dominance it could be of interest to determine what happens in growth arrest assays or reporter gene assays if both the WT and ΔQA versions of TCERG1 are co-expressed.

It was observed when performing the Co-IP experiments that the ΔQA TCERG1 migrated at a different size than expected. When comparing the sizes of the QA_{11} and ΔQA in figure 18 it is observed that the ΔQA migrates at a higher molecular weight than the QA11. Presumably this would mean that there is some post-translational modifications present in the ΔQA that may or may not be present in WT TCERG1. Determining what is causing ΔQA TCERG1 to migrate at a different than expected molecular weight may provide insight into the further function of the QA domain.

Although not addressed in this thesis it was noted that the nuclear speckle compartments appeared to increase in size when ΔQA TCERG1 was expressed (see figure 10). Since an increase in nuclear speckle size is usually correlated with an inhibition of transcription or splicing it may be worthwhile to investigate this phenomenon. By using targets of TCERG1 and performing microarray, RNAseq or similar analysis it should be possible to determine if the ΔQA mutant is able to affect the transcriptional or splicing products inside the cell. By performing this analysis it could give important insights into the further control of transcription and splicing by not only TCERG1 but the QA domain as well.

It has been suggested in this thesis that there could be another domain helping the QA domain achieve full functionality. Further experimentation will have to be performed to determine if the STP domain or other domains inside TCERG1 are helping to stabilize the interactions that the QA domain are catalyzing. These experiments could be performed similar to the isolated QA domain experiments.

6.2 C/EBPα Interaction Site

Whereas this research provides information on the site of interaction between C/EBPα and TCERG1 via TCERG1 there is very little known about the interaction domains involved in terms of C/EBPα. It was hypothesized in this research that TCERG1 and C/EBPα interact via the S193 domain as well as another unknown domain. Further research could be conducted to further this knowledge and specifically which sites are the sites of interaction.

6.3 Whole Animal Testing

Since little is known about the QA domain in general it is of unknown importance. As discussed in section 2.4.4.1 the QA domain appears to be a recently evolved motif. As such, it may be of importance to determine the significance of the reason this domain has arisen and the evolutionary advantage conferred. This study would ideally delete specifically the QA domain from TCERG1 to determine the physiological effects of such a deletion. Furthermore, as demonstrated in section 4.2.3 the QA domain may function in a dominant negative as therefore a heterozygous deletion of the QA domain could provide some important data if the dominant negative properties are once again observed in a whole animal model.

7. References

Andresen, J.M., Gayan, J., Cherny, S.S., Brocklebank, D., Alkorta-Aranburu, G., Addis, E.A., Group, U.S.-V.C.R., Cardon, L.R., Housman, D.E., and Wexler, N.S. (2007). Replication of twelve association studies for Huntington's disease residual age of onset in large Venezuelan kindreds. J Med Genet *44*, 44-50.

Banman, S.L., Mcfie, P.J., Wilson, H.L., and Roesler, W.J. (2010). Nuclear redistribution of TCERG1 is required for its ability to inhibit the transcriptional and anti-proliferative activities of C/EBPalpha. J Cell Biochem *109*, 140-151.

Bauer, M.T., Gilmore, K.A., and Petty, S.A. (2011). Formation of beta-sheets in glutamine and alanine tripeptides. Biochem Biophys Res Commun *406*, 348-352.

Bedford, M.T., and Leder, P. (1999). The FF domain: a novel motif that often accompanies WW domains. Trends Biochem Sci *24*, 264-265.

Bohne, J., Cole, S.E., Sune, C., Lindman, B.R., Ko, V.D., Vogt, T.F., and Garcia-Blanco, M.A. (2000). Expression analysis and mapping of the mouse and human transcriptional regulator CA150. Mamm Genome *11*, 930-933.

Carrero, G., Hendzel, M.J., and De Vries, G. (2006). Modelling the compartmentalization of splicing factors. J Theor Biol *239*, 298-312.

Carty, S.M., Goldstrohm, A.C., Sune, C., Garcia-Blanco, M.A., and Greenleaf, A.L. (2000). Protein-interaction modules that organize nuclear function: FF domains of CA150 bind the phosphoCTD of RNA polymerase II. Proc Natl Acad Sci U S A *97*, 9015-9020.

Chattopadhyay, B., Ghosh, S., Gangopadhyay, P.K., Das, S.K., Roy, T., Sinha, K.K., Jha, D.K., Mukherjee, S.C., Chakraborty, A., Singhal, B.S., *et al.* (2003). Modulation of age at onset in Huntington's disease and spinocerebellar ataxia type 2 patients originated from eastern India. Neurosci Lett *345*, 93-96.

Coiras, M., Montes, M., Montanuy, I., Lopez-Huertas, M.R., Mateos, E., Le Sommer, C., Garcia-Blanco, M.A., Hernandez-Munain, C., Alcami, J., and Sune, C. (2013). Transcription elongation regulator 1 (TCERG1) regulates competent RNA polymerase II-mediated elongation of HIV-1 transcription and facilitates efficient viral replication. Retrovirology *10*, 124.

Dundr, M., and Misteli, T. (2010). Biogenesis of nuclear bodies. Cold Spring Harb Perspect Biol *2*, a000711.

Ghazi, A. (2013). Transcriptional networks that mediate signals from reproductive tissues to influence lifespan. Genesis *51*, 1-15.

Ghazi, A., Henis-Korenblit, S., and Kenyon, C. (2009). A transcription elongation factor that links signals from the reproductive system to lifespan extension in Caenorhabditis elegans. PLoS Genet *5*, e1000639.

Gibson, D.G., Young, L., Chuang, R.Y., Venter, J.C., Hutchison, C.A., 3rd, and Smith, H.O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. Nat Methods *6*, 343-345.

Gibson, T.J., Seiler, M., and Veitia, R.A. (2013). The transience of transient overexpression. Nat Methods *10*, 715-721.

Goldstrohm, A.C., Albrecht, T.R., Sune, C., Bedford, M.T., and Garcia-Blanco, M.A. (2001). The transcription elongation factor CA150 interacts with RNA polymerase II and the pre-mRNA splicing factor SF1. Mol Cell Biol *21*, 7617-7628.

Harbour, J.W., and Dean, D.C. (2000). The Rb/E2F pathway: expanding roles and emerging paradigms. Genes Dev *14*, 2393-2409.

Harris, T.E., Albrecht, J.H., Nakanishi, M., and Darlington, G.J. (2001). CCAAT/enhancer-binding protein-alpha cooperates with p21 to inhibit cyclin-dependent kinase-2 activity and induces growth arrest independent of DNA binding. J Biol Chem *276*, 29200-29209.

Horvath, J.E., Viggiano, L., Loftus, B.J., Adams, M.D., Archidiacono, N., Rocchi, M., and Eichler, E.E. (2000). Molecular structure and evolution of an alpha satellite/non-alpha satellite junction at 16p11. Hum Mol Genet *9*, 113-123.

Hsin, J.P., and Manley, J.L. (2012). The RNA polymerase II CTD coordinates transcription and RNA processing. Genes Dev *26*, 2119-2137.

Huang, S., and Spector, D.L. (1996). Intron-dependent recruitment of pre-mRNA splicing factors to sites of transcription. J Cell Biol *133*, 719-732.

Johnson, P.F. (2005). Molecular stop signs: regulation of cell-cycle arrest by C/EBP transcription factors. J Cell Sci *118*, 2545-2555.

Keith, S.A., and Ghazi, A. (2014). Recent Discoveries in the Reproductive Control of Aging. Curr Gen Med Rep *3*, 26-34.

Khanna-Gupta, A. (2008). Sumoylation and the function of CCAAT enhancer binding protein alpha (C/EBP alpha). Blood Cells Mol Dis *41*, 77-81.

Kirstetter, P., Schuster, M.B., Bereshchenko, O., Moore, S., Dvinge, H., Kurz, E., Theilgaard-Monch, K., Mansson, R., Pedersen, T.A., Pabst, T., *et al.* (2008). Modeling of C/EBPalpha mutant acute myeloid leukemia reveals a common expression signature of committed myeloid leukemia-initiating cells. Cancer Cell *13*, 299-310.

Kurogi, Y., Matsuo, Y., Mihara, Y., Yagi, H., Shigaki-Miyamoto, K., Toyota, S., Azuma, Y., Igarashi, M., and Tani, T. (2014). Identification of a chemical inhibitor for nuclear speckle formation: implications for the function of nuclear speckles in regulation of alternative pre-mRNA splicing. Biochem Biophys Res Commun *446*, 119-124.

Lamond, A.I., and Spector, D.L. (2003). Nuclear speckles: a model for nuclear organelles. Nat Rev Mol Cell Biol *4*, 605-612.

Lejeune, E., Bayne, E.H., and Allshire, R.C. (2010). On the connection between RNAi and heterochromatin at centromeres. Cold Spring Harb Symp Quant Biol *75*, 275-283.

Liu, J., Fan, S., Lee, C.J., Greenleaf, A.L., and Zhou, P. (2013). Specific interaction of the transcription elongation regulator TCERG1 with RNA polymerase II requires simultaneous phosphorylation at Ser2, Ser5, and Ser7 within the carboxyl-terminal domain repeat. J Biol Chem *288*, 10890-10901.

Liu, W., Enwright, J.F., 3rd, Hyun, W., Day, R.N., and Schaufele, F. (2002). CCAAT/enhancer binding protein alpha uses distinct domains to prolong pituitary cells in the growth 1 and DNA synthesis phases of the cell cycle. BMC Cell Biol *3*, 6.

Liu, X., Wu, B., Szary, J., Kofoed, E.M., and Schaufele, F. (2007). Functional sequestration of transcription factor activity by repetitive DNA. J Biol Chem *282*, 20868-20876.

Mao, Y.S., Zhang, B., and Spector, D.L. (2011). Biogenesis and function of nuclear bodies. Trends Genet *27*, 295-306.

Mcfie, P.J., Wang, G.L., Timchenko, N.A., Wilson, H.L., Hu, X., and Roesler, W.J. (2006). Identification of a co-repressor that inhibits the transcriptional and growth-arrest activities of CCAAT/enhancer-binding protein alpha. J Biol Chem *281*, 18069-18080.

Mcknight, S.L., Lane, M.D., and Gluecksohn-Waelsch, S. (1989). Is CCAAT/enhancer-binding protein a central regulator of energy metabolism? Genes Dev *3*, 2021-2024.

Milo, R. (2013). What is the total number of protein molecules per cell volume? A call to rethink some published values. Bioessays *35*, 1050-1055.

Misteli, T. (2005). Concepts in nuclear architecture. Bioessays *27*, 477-487.

Misteli, T. (2007). Beyond the sequence: cellular organization of genome function. Cell *128*, 787-800.

Misteli, T., and Spector, D.L. (1997). Protein phosphorylation and the nuclear organization of pre-mRNA splicing. Trends Cell Biol *7*, 135-138.

Misteli, T., and Spector, D.L. (1998). The cellular organization of gene expression. Curr Opin Cell Biol *10*, 323-331.

Moazed, B., Banman, S.L., Wilkinson, G.A., and Roesler, W.J. (2011). TCERG1 inhibits C/EBPalpha through a mechanism that does not involve sequestration of C/EBPalpha at pericentromeric heterochromatin. J Cell Biochem *112*, 2317-2326.

Montes, M., Cloutier, A., Sanchez-Hernandez, N., Michelle, L., Lemieux, B., Blanchette, M., Hernandez-Munain, C., Chabot, B., and Sune, C. (2012). TCERG1 regulates alternative splicing of the Bcl-x gene by modulating the rate of RNA polymerase II transcription. Mol Cell Biol *32*, 751-762.

Morimoto, M., and Boerkoel, C.F. (2013). The role of nuclear bodies in gene expression and disease. Biology (Basel) *2*, 976-1033.

Muller, C., Calkhoven, C.F., Sha, X., and Leutz, A. (2004). The CCAAT enhancer-binding protein alpha (C/EBPalpha) requires a SWI/SNF complex for proliferation arrest. J Biol Chem *279*, 7353-7358.

Nerlov, C. (2004). C/EBPalpha mutations in acute myeloid leukaemias. Nat Rev Cancer *4*, 394-400.

Nerlov, C. (2008). C/EBPs: recipients of extracellular signals through proteome modulation. Curr Opin Cell Biol *20*, 180-185.

Pearson, J.L., Robinson, T.J., Munoz, M.J., Kornblihtt, A.R., and Garcia-Blanco, M.A. (2008). Identification of the cellular targets of the transcription factor TCERG1 reveals a prevalent role in mRNA processing. J Biol Chem *283*, 7949-7961.

Pelassa, I., Cora, D., Cesano, F., Monje, F.J., Montarolo, P.G., and Fiumara, F. (2014). Association of polyalanine and polyglutamine coiled coils mediates expansion disease-related protein aggregation and dysfunction. Hum Mol Genet *23*, 3402-3420.

Portugal, J., and Waring, M.J. (1988). Assignment of DNA binding sites for 4',6-diamidine-2-phenylindole and bisbenzimide (Hoechst 33258). A comparative footprinting study. Biochim Biophys Acta *949*, 158-168.

Prasanth, K.V., Sacco-Bubulya, P.A., Prasanth, S.G., and Spector, D.L. (2003). Sequential entry of components of the gene expression machinery into daughter nuclei. Mol Biol Cell *14*, 1043-1057.

Ramirez-Alvarado, M., Merkel, J.S., and Regan, L. (2000). A systematic exploration of the influence of the protein stability on amyloid fibril formation in vitro. Proc Natl Acad Sci U S A *97*, 8979-8984.

Ramji, D.P., and Foka, P. (2002). CCAAT/enhancer-binding proteins: structure, function and regulation. Biochem J *365*, 561-575.

Rieder, D., Ploner, C., Krogsdam, A.M., Stocker, G., Fischer, M., Scheideler, M., Dani, C., Amri, E.Z., Muller, W.G., Mcnally, J.G., *et al.* (2014). Co-expressed genes prepositioned in

spatial neighborhoods stochastically associate with SC35 speckles and RNA polymerase II factories. Cell Mol Life Sci *71*, 1741-1759.

Roe, J.S., and Vakoc, C.R. (2014). C/EBPalpha: critical at the origin of leukemic transformation. J Exp Med *211*, 1-4.

Roesler, W.J. (2001). The role of C/EBP in nutrient and hormonal regulation of gene expression. Annu Rev Nutr *21*, 141-165.

Sacco-Bubulya, P., and Spector, D.L. (2002). Disassembly of interchromatin granule clusters alters the coordination of transcription and pre-mRNA splicing. J Cell Biol *156*, 425-436.

Sanchez-Alvarez, M., Goldstrohm, A.C., Garcia-Blanco, M.A., and Sune, C. (2006). Human transcription elongation factor CA150 localizes to splicing factor-rich nuclear speckles and assembles transcription and splicing components into complexes through its amino and carboxyl regions. Mol Cell Biol *26*, 4998-5014.

Sanchez-Alvarez, M., Montes, M., Sanchez-Hernandez, N., Hernandez-Munain, C., and Sune, C. (2010). Differential effects of sumoylation on transcription and alternative splicing by transcription elongation regulator 1 (TCERG1). J Biol Chem *285*, 15220-15233.

Sanchez-Hernandez, N., Ruiz, L., Sanchez-Alvarez, M., Montes, M., Macias, M.J., Hernandez-Munain, C., and Sune, C. (2012). The FF4 and FF5 domains of transcription elongation regulator 1 (TCERG1) target proteins to the periphery of speckles. J Biol Chem *287*, 17789-17800.

Schaufele, F., Enwright, J.F., 3rd, Wang, X., Teoh, C., Srihari, R., Erickson, R., Macdougald, O.A., and Day, R.N. (2001). CCAAT/enhancer binding protein alpha assembles essential cooperating factors in common subnuclear domains. Mol Endocrinol *15*, 1665-1676.

Schor, I.E., Lleres, D., Risso, G.J., Pawellek, A., Ule, J., Lamond, A.I., and Kornblihtt, A.R. (2012). Perturbation of chromatin structure globally affects localization and recruitment of splicing factors. PLoS One *7*, e48084.

Sleeman, J.E., and Lamond, A.I. (1999). Nuclear organization of pre-mRNA splicing factors. Curr Opin Cell Biol *11*, 372-377.

Sleeman, J.E., and Trinkle-Mulcahy, L. (2014). Nuclear bodies: new insights into assembly/dynamics and disease relevance. Curr Opin Cell Biol *28*, 76-83.

Slomiany, B.A., D'arigo, K.L., Kelly, M.M., and Kurtz, D.T. (2000). C/EBPalpha inhibits cell growth via direct repression of E2F-DP-mediated transcription. Mol Cell Biol *20*, 5986-5997.

Smith, M.J., Kulkarni, S., and Pawson, T. (2004). FF domains of CA150 bind transcription and splicing factors through multiple weak interactions. Mol Cell Biol *24*, 9274-9285.

Spector, D.L., and Lamond, A.I. (2011). Nuclear speckles. Cold Spring Harb Perspect Biol *3*.

Sune, C., and Garcia-Blanco, M.A. (1999). Transcriptional cofactor CA150 regulates RNA polymerase II elongation in a TATA-box-dependent manner. Mol Cell Biol *19*, 4719-4728.

Sune, C., Hayashi, T., Liu, Y., Lane, W.S., Young, R.A., and Garcia-Blanco, M.A. (1997). CA150, a nuclear protein associated with the RNA polymerase II holoenzyme, is involved in Tat-activated human immunodeficiency virus type 1 transcription. Mol Cell Biol *17*, 6029-6039.

Wang, G.L., and Timchenko, N.A. (2005). Dephosphorylated C/EBPalpha accelerates cell proliferation through sequestering retinoblastoma protein. Mol Cell Biol *25*, 1325-1338.

Wang, H., Iakova, P., Wilde, M., Welm, A., Goode, T., Roesler, W.J., and Timchenko, N.A. (2001). C/EBPalpha arrests cell proliferation through direct inhibition of Cdk2 and Cdk4. Mol Cell *8*, 817-828.

Wang, K.C., Cheng, A.L., Chuang, S.E., Hsu, H.C., and Su, I.J. (2000). Retinoic acid-induced apoptotic pathway in T-cell lymphoma: Identification of four groups of genes with differential biological functions. Exp Hematol *28*, 1441-1450.

Wang, N.D., Finegold, M.J., Bradley, A., Ou, C.N., Abdelsayed, S.V., Wilde, M.D., Taylor, L.R., Wilson, D.R., and Darlington, G.J. (1995). Impaired energy homeostasis in C/EBP alpha knockout mice. Science *269*, 1108-1112.

Yogesha, S.D., Mayfield, J.E., and Zhang, Y. (2014). Cross-talk of phosphorylation and prolyl isomerization of the C-terminal domain of RNA Polymerase II. Molecules *19*, 1481-1511.

Zaragoza, K., Begay, V., Schuetz, A., Heinemann, U., and Leutz, A. (2010). Repression of transcriptional activity of C/EBPalpha by E2F-dimerization partner complexes. Mol Cell Biol *30*, 2293-2304.

Zhang, W.H., Srihari, R., Day, R.N., and Schaufele, F. (2001). CCAAT/enhancer-binding protein alpha alters histone H3 acetylation at large subnuclear domains. J Biol Chem *276*, 40373-40376.

Zhou, J., Liu, Y., Zhang, W., Popov, V.M., Wang, M., Pattabiraman, N., Sune, C., Cvekl, A., Wu, K., Jiang, J., *et al.* (2010). Transcription elongation regulator 1 is a co-integrator of the cell fate determination factor Dachshund homolog 1. J Biol Chem *285*, 40342-40350.