

Development of a system to probe the differential production of eIF4A from
the paralogous genes *TIF1* and *TIF2* in *Saccharomyces cerevisiae*

Darryl Jones

A thesis

Submitted to the Victoria University of Wellington
in fulfilment of the requirements for the degree of
Master of Biomedical Science

Victoria University of Wellington

2012

Abstract

The aim of this Master's thesis was to evaluate the hypothesis that the *Saccharomyces cerevisiae* homologs for Eukaryotic Initiation Factor 4A, *TIF1* and *TIF2*, can be individually regulated. This may be an expectation if the retention of the *TIF1* and *TIF2* duplicates arises from the requirement to respond to a wide variety of cellular needs in *S. cerevisiae*. These paralogs show an almost identical sequence in their coding region (barring six synonymous changes in the gene sequence) but are substantially different in their 5' and 3' untranslated regions that are the probable sites of regulatory functions. To identify differences in use of the *TIF1* and *TIF2* paralogs, a dual fluorescent reporter strain expressing plasmid borne *TIF1*-RFP, incorporating the endogenous *TIF1* 3' and 5' untranslated regions, and chromosomally integrated *TIF2*-GFP was created in order to probe for any differential regulation between *TIF1* and *TIF2*. To create the fluorescent reporters it was necessary to learn and execute sophisticated molecular biology and molecular genetics which are described in this thesis. The generated fluorescence reporter strains were shown to be stable over multiple generations and subjected to high throughput and high content automated confocal microscopy. The commercially available LOPAC¹²⁸⁰ Library of Pharmacologically-Active Compounds was used to probe for differential regulation where a "hit" was defined as a significant change in the expression of at least one of *TIF1*-RFP or *TIF2*-GFP four hours after application. For *TIF2*-GFP, 2 compounds out of the 1280 library showed evidence of regulation under stringent thresholding criteria. For *TIF1*-RFP, 43 compounds were identified as regulators. There was no overlap of compounds. This screen provides *prima-facie* evidence that the *TIF1* and *TIF2* are differentially regulated, as assessed by the criteria of the experimental system described in this thesis.

Acknowledgements

I would like to express my gratitude to my supervisor, Dr. Paul Teesdale-Spittle for all the imparted knowledge, unwavering support and calming words. Thank you for the much needed guidance throughout this journey not to mention, your potentially limitless patience levels without which I may never have submitted this thesis. I would also like to extend my gratitude to Prof. Paul Atkinson and Dr. David Maass without whom the lab would not function. You have both been inspirational in my development as a scientist.

I would like to extend my gratitude to Liz Richardson and the entirety of Te Ropu Awhina whanau. Your faith in me as well as your continued support throughout the years has made studying away from home, even though it's less than an island away, a lot easier.

To Dr. James Matthews, Dr. Veronica Venturi and all those who throughout my time in the Chemical Genetics lab have worked under Paul's supervision, thank you for being great teachers, collaborators and all around quality individuals. Thank you for being exceptional sounding boards for this research especially when the progress came grinding to a halt and the sounding board evolved into the complaining board.

To Peter and Bede, thanks for sharing your knowledge and quality lab skills, without your help I probably wouldn't have been able to turn the microscope on let alone the rest of what is contained in this thesis. To Katie, thanks for being good at art as well as remembering that homeless people need food too. In particular you three have made my time in the lab significantly more enjoyable if not potentially less efficient. Much appreciated.

To the wider Chemical Genetics family, as well as the late night workers, thanks for sharing this journey with me. There are many of you I wish I could name individually as you have been major contributors to the success of this research.

I would also like to acknowledge all those past and present who I have flatted with during my studies. You have made the transition between work and play seamless allowing me to study a lot longer than I probably should have.

Finally I would like to acknowledge the support of my Mum, Dad and the rest of my ever increasing family. Without your support I definitely would not be where I am today and I

am eternally grateful. Hopefully this “giant doorstep”, that you will each receive a copy of as a Christmas present, reminds you of the things you have given up for my successes.

List of Figures

Figure 1.1: The families of SF1 and SF2.	2
Figure 1.2: The scanning model of cap dependent translation initiation.	4
Figure 1.3: The process leading to long-term duplicate survival.	10
Figure 2.1: Illustration of LOPAC screen 384-well plate plan.....	28
Figure 3.1: Primer design employed when creating TIF1-RFP via PCR mediated module construction.....	32
Figure 3.2: PCR mediated component module construction.	34
Figure 3.3: Sites of sequence homology between modules and plasmid.....	35
Figure 3.4: Plasmid borne <i>TIF1</i> -RFP.	36
Figure 3.5: Confocal microscopy images of the 4 strains utilised in the HTS of the LOPAC library.....	38
Figure 3.6: Overview of TIF1p-RFP TIF2p-GFP strain generation.	41
Figure 3.7: Whole cell fluorescence intensities of TIF1p-GFP treated with the LOPAC library.....	48
Figure 3.8: Whole cell fluorescence intensities of TIF2p-GFP treated with the LOPAC library	51
Figure 3.9: Comparison of the GFP components from the single TIF1p-GFP screen compared with that from the TIF2p-GFP screen.....	54
Figure 3.10: Comparison of the GFP components from the single TIF1p-GFP screen compared with that from the TIF2p-GFP screen on a restricted range..	55
Figure 3.11: Whole cell fluorescence intensities of the GFP component of TIF1p-GFP TIF1p-RFP treated with the LOPAC library.....	59
Figure 3.12: Whole cell fluorescence intensities of the RFP component of TIF1p-GFP TIF1p-RFP treated with the LOPAC library.....	62
Figure 3.13: Comparison of the GFP and RFP components from the TIF1p-GFP TIF1p-RFP LOPAC library screen.....	65
Figure 3.14: Comparison of the GFP and RFP components from the TIF1p-GFP TIF1p-RFP LOPAC library screen on a restricted range.....	66
Figure 3.15: Whole cell fluorescence intensities of the GFP component of TIF2p-GFP TIF1p-RFP treated with the LOPAC library.....	69

Figure 3.16: Whole cell fluorescence intensities of the RFP component of TIF2p-GFP TIF1p-RFP treated with the LOPAC library.....	72
Figure 3.17: Comparison of the GFP and RFP components from the TIF2p-GFP TIF1p-RFP LOPAC library screen.....	75
Figure 3.18: Comparison of the GFP and RFP components from the TIF2p-GFP TIF1p-RFP LOPAC library screen on a restricted range	76
Figure 3.19 : Comparison of the RFP components from the TIF1p-GFP TIF1p-RFP screen compared with that from the TIF2p-GFP TIF1p-RFP screen.	79
Figure 3.20: Strategy for GFP library construction.	83
Figure 5.1: Whole cell fluorescence intensities of TIF1p-GFP treated with the LOPAC library at t = 0.....	96
Figure 5.2: Whole cell fluorescence intensities of TIF2p-GFP treated with the LOPAC library at t = 0.....	98
Figure 5.3: Whole cell fluorescence intensities of the GFP component of TIF1p-GFP TIF1p-RFP treated with the LOPAC library t = 0	100
Figure 5.4: Whole cell fluorescence intensities of the GFP component of TIF1p-GFP TIF1p-RFP treated with the LOPAC library at t = 0 on a restricted range	101
Figure 5.5: Whole cell fluorescence intensities of the RFP component of TIF1p-GFP TIF1p-RFP treated with the LOPAC library at t = 0	103
Figure 5.6: Whole cell fluorescence intensities of the GFP component of TIF2p-GFP TIF1p-RFP treated with the LOPAC library at t = 0	105
Figure 5.7: Whole cell fluorescence intensities of the RFP component of TIF2p-GFP TIF1p-RFP treated with the LOPAC library at t = 0	107

List of Tables

Table 2.1: Reaction setup for the component module construction PCR.....	25
Table 2.2: Primers used in the TIF1-RFP component module construction PCR	25
Table 3.1: The construct design of the TIF1p-RFP modules as assembled through PCR mediated module construction..	33
Table 3.2: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF1p-GFP screen	49
Table 3.3: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF2p-GFP screen	52
Table 3.4: Compounds that elicited a decrease in the observed ratio of compound-treated GFP compared to control GFP in both the single <i>TIF1</i> and <i>TIF2</i> GFP screens.....	53
Table 3.5: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF1p-GFP TIF1p-RFP screen.	60
Table 3.6: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated RFP compared to control RFP from the TIF1p-GFP TIF1p-RFP screen.....	63
Table 3.7: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF2p-GFP TIF1p-RFP screen.....	70
Table 3.8: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated RFP compared to control RFP from the TIF2p-GFP TIF1p-RFP screen.....	73
Table 3.9: The primer set used in the construction of TIF1-RFP colour coded to highlight the origin of each component	80
Table 5.1: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF1p-GFP screen at $t = 0$	97
Table 5.2: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF2p-GFP screen at $t = 0$	99

Table 5.3: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF1p-GFP TIF1p-RFP screen at $t = 0$	102
Table 5.4: The top compounds that elicited an increase or a decrease in the observed ratio of compound-treated RFP compared to control RFP from the TIF1p-GFP TIF1p-RFP screen at $t = 0$	104
Table 5.5: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF2p-GFP TIF1p-RFP screen at $t = 0$	106
Table 5.6: The top compounds that elicited an increase or a decrease in the observed ratio of compound-treated RFP compared to control RFP from the TIF2p-GFP TIF1p-RFP screen at $t = 0$	108

List of Abbreviations

ATP	Adenosine triphosphate
BLAST	Basic Local Alignment Search Tool
BSA	Bovine serum albumin
CFP	Cerulean fluorescent protein
ddH ₂ O	Double distilled water
DEAD box	D-E-A-D (Asp-Glu-Ala-Asp) motif
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleotide triphosphates
DO	Drop-out
EDTA	Ethylenediaminetetraacetic acid
eIF	Eukaryotic initiation factor
Fwd	Forward
G418	Geneticin
GFP	Green fluorescent protein
IRES	Internal Ribosome Entry Site
Kb	Kilobase
KO	Knockout
LB	Luria-Bertani
LOPAC	Library of Pharmacologically-Active Compounds
m ⁷ G	7-methylguanosine

mRNA	Messenger RNA
MSG	Monosodium glutamate
NLS	Nuclear localisation signal
NMD	Non-sense mediated decay
OD	Optical density
ORF	Open reading frame
PABP	Poly (A) binding protein
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PIC	43S pre-initiation complex
Rev	Reverse
RFP	Red fluorescent protein
RNA	Ribonucleic acid
Rpm	Revolutions per minute
SC	Synthetic complete
SD	Synthetic dropout
SF	Superfamilies
SGD	<i>Saccharomyces</i> genome database
SSD	Small scale duplication
TBE	Tris/Borate/Ethylenediaminetetraacetic acid
TE	Tris-Ethylenediamine Tetraacetic Acid

Tris	Tris(hydroxymethyl)aminomethane
UTR	Untranslated region
UV	Ultraviolet
v/v	Volume/volume
w/v	Weight/volume
WGD	Whole genome duplication
YEAstract	Yeast Search for Transcriptional Regulators And Consensus Tracking
YGDS	Yeast Genome Deletion Set
YKO	Yeast knock out
YPD	Yeast peptone dextrose
X _g	Times gravity

Table of Contents

Abstract.....	ii
Acknowledgements	iii
List of Figures	v
List of Tables	vii
List of Abbreviations	ix
1 Introduction	1
1.1 RNA helicases and the DEAD box proteins.	1
1.1.1 RNA helicases.....	1
1.1.2 DEAD box protein family of helicases.....	2
1.2 Translation initiation and eIF4A.....	3
1.2.1 Translation Initiation.....	3
1.2.2 Cap dependent initiation	4
1.2.3 Cap independent Initiation	6
1.2.4 Translation initiation and eIF4A in <i>Saccharomyces cerevisiae</i>	6
1.3 Duplicate genes.....	9
1.3.1 The need for duplication	9
1.3.2 Duplication methods	9
1.3.3 Gene functional bias	10
1.3.4 Surviving duplication – Gene fixation.....	12
1.3.5 Genetic redundancy.....	12
1.3.6 Duplication in <i>S. cerevisiae</i>	14
1.4 Regulation of protein synthesis	14
1.5 Aims and Objectives.....	16
1.5.1 Aims	16
1.5.2 Experimental Objectives.....	17

2	Methods	18
2.1	Materials and Equipment	18
2.1.1	Reagents	18
2.1.2	Chemical libraries and individual chemicals.....	18
2.1.3	Yeast strains.....	18
2.1.4	Growth media	19
2.1.5	Bacterial growth media	20
2.1.6	Plasmids used in this study.....	20
2.2	Methodology.....	21
2.2.1	DNA preparation and manipulation	21
2.2.2	Dual fluorescent strain generation.....	26
2.2.3	LOPAC library HTS of dual fluorescent probe	27
3	Results and Discussion.....	31
3.1	Creation of a <i>TIF1p</i> -RFP fusion protein.....	31
3.1.1	Primer design and module creation	31
3.1.2	Introduction of modules into a functional host	35
3.2	Construction of a dual fluorescent reporter strain.....	36
3.2.1	Application of dual fluorescent reporter strain.....	42
3.3	Screening of the LOPAC ¹²⁸⁰ library.....	44
3.3.1	Rational of statistical analysis.....	45
3.3.2	Single GFP screening – Reading at 4 Hours	46
3.3.3	Dual Fluorescent GFP/RFP screening – Reading at 4 Hours	57
3.4	Advantages and limitations of this screen.....	80
3.4.1	Advantages	80
3.4.2	Limitations	85
3.5	Conclusions	87

3.6	Future directions	89
4	References	91
5	Appendix A – LOPAC Library HTS time zero readings	95
5.1	Single GFP screening	96
5.1.1	TIF1p-GFP screen	96
5.1.2	TIF2p-GFP screen	98
5.2	Dual Fluorescent GFP/RFP screening	100
5.2.1	TIF1p-GFP TIF1p-RFP screen	100
5.2.2	TIF2p-GFP TIF1p-RFP screen	105

1 Introduction

The study described in this thesis aims to understand the evolutionary benefit to the retention of two paralogous genes which produce an identical protein, eIF4A, in *Saccharomyces cerevisiae*. It is proposed that this benefit is associated with the ability to independently regulate the abundance of eIF4A in response to cellular needs through the paralogous genes. A fluorescence-based reporter strain has been constructed to test this hypothesis, which has been evaluated through treatment of the reporter strain with small molecule perturbants.

The role of eIF4A is presented below, specifically its position within the superfamily of RNA helicases and the initiation of translation. This study sits within a wider programme of evaluation of the role of the duplicate genes for eIF4A in yeast, and so the principles of gene duplication are introduced and the implications relating to eIF4A and translation are discussed.

1.1 RNA helicases and the DEAD box proteins.

1.1.1 RNA helicases

Enzymes that use ATP to bind or remodel RNA and RNA-protein complexes are termed RNA helicases, and are involved in almost all aspects of RNA metabolism, from transcription and translation to mRNA decay (Linder and Jankowsky 2011). A classification system established by Gorbalenya and Koonin, sorts both RNA and DNA helicases into superfamilies based on the occurrence and characteristics of conserved motifs in their primary sequence (Gorbalenya and Koonin 1993; Cordin, Banroques et al. 2006). Although the structural conservation within each superfamily is high, two distinct types of helicases exist which divide the superfamilies; those that form multimeric ring-like assemblies as seen in superfamilies 3-6, and those which do not form rings, superfamilies 1 and 2 (SF1 and SF2) (Fairman-Williams, Guenther et al. 2010). The families which make up SF1 and SF2 are illustrated in Figure 1.1.

SF1 and SF2 helicases are characterised by the presence of at least 12 consensus sequences motifs situated around a highly conserved core region of two virtually identical domains resembling the bacterial recombination protein recombinase A (RecA) (Linder

and Jankowsky 2011). Sequence conservation across both superfamilies is at its highest in the residues between the two RecA like domains which coordinate ATP binding and hydrolysis (motifs I, II and VI) (Rocak and Linder 2004); (Jankowsky, Guenther et al. 2011).

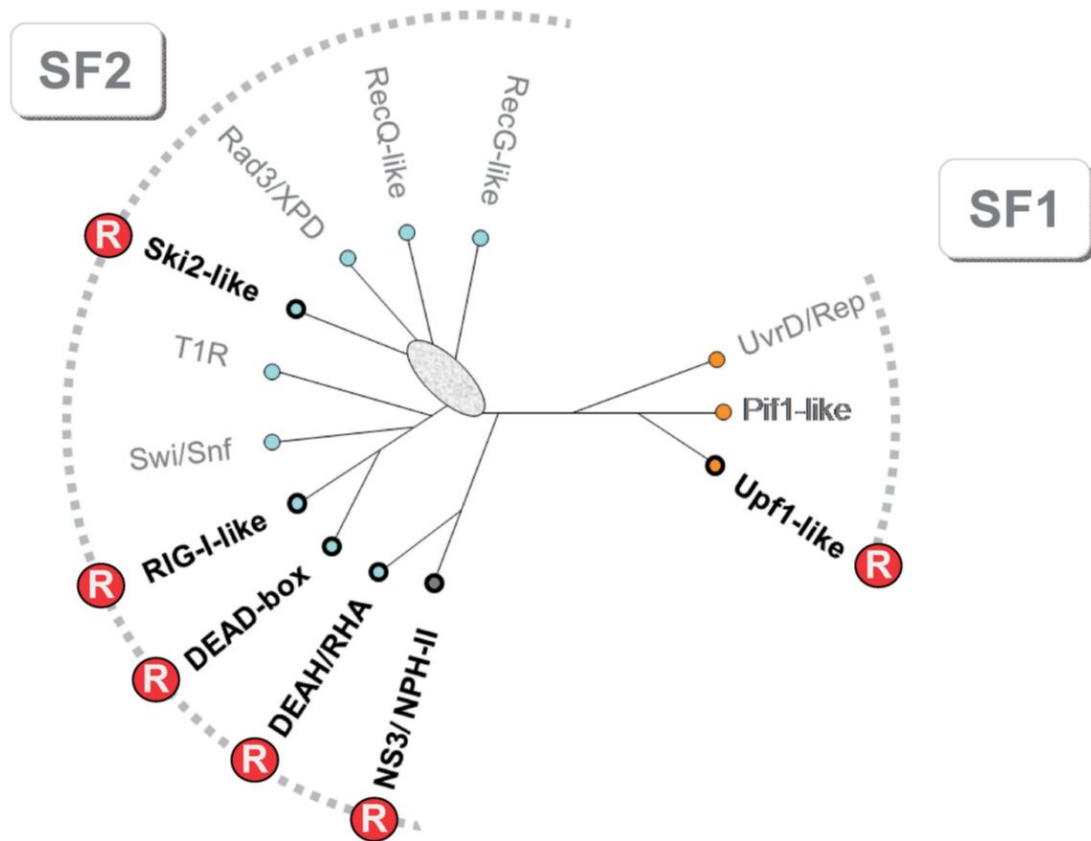


Figure 1.1: The families of SF1 and SF2. Families containing only RNA helicases are marked with a circled R, all other families include DNA helicases as well as RNA helicases. The oval represents uncertainty in the topology. Figure from Jankowsky 2011 (Jankowsky, Guenther et al. 2011).

1.1.2 DEAD box protein family of helicases

First described by Linder and Slonimski, the DEAD box protein family of RNA helicases are the largest sub class of Superfamily 2 (SF2) (Linder and Slonimski 1988; Linder and Slonimski 1989) and, although known to bind DNA, are the only family to contain exclusively RNA helicases. Although DEAD box proteins have been found in bacteria and archaea they are most prevalent in eukaryotes (Rocak and Linder 2004) and are directly

involved in nuclear transcription, pre mRNA splicing, and ribosome biogenesis (de la Cruz, Kressler et al. 1999). Dead Box proteins are characterised by the presence of at least 12 consensus sequences /conserved motifs in the core region and notably it is the presence of the amino acid sequence D-E-A-D in Motif II (or Walker B motif) from which their name is derived (Linder 2006; Linder and Jankowsky 2011). This core region is flanked by N and C terminal extensions with varying length and composition (Cordin, Banroques et al. 2006). Moreover, the core region maintains ATP dependent RNA helicase activities, whilst the flanking regions provide additional interactions with substrates and cofactors leading to the substrate specificity and wide functional variation of the DEAD box family (Benz, Trachsel et al. 1999; de la Cruz, Kressler et al. 1999; Rocak and Linder 2004).

Eukaryotic Initiation Factor 4A (eIF4A), one of the most abundant proteins in many cell types (Linder and Jankowsky 2011), is the archetypical DEAD box protein representing the minimum core helicase elements common to all DEAD box proteins (Benz, Trachsel et al. 1999; Rocak and Linder 2004). It is generally accepted that eIF4A has a major function in the translation initiation process.

1.2 Translation initiation and eIF4A

Translation, the conversion of genomic information from mRNA into an amino acid sequence, matured and folded into protein, can be sub divided into three major sections; initiation, elongation, and termination (Sonnenberg and Hinnebusch 2009).

1.2.1 Translation Initiation

The initiation phase, the rate limiting step of translation (Lackner and Bähler 2008), involves recruitment of the 40S ribosomal subunit at the initiator AUG codon and the subsequent assembly of a translational /elongation competent 80S ribosome with the union of the 40S and 60S ribosomal subunits (Lackner and Bähler 2008). The initiation phase, for most eukaryotic mRNAs, is a highly conserved process catalysed by at least 11 eukaryotic initiation factors (eIFs), many of which form multi-protein complexes with up to a minimum of 25 proteins (Berthelot, Muldoon et al. 2004). The process is currently best described by the scanning model proposed by Kozak and Shatkin (Kozak and Shatkin 1978; Kozak 1999).

1.2.2 Cap dependent initiation

The scanning model (illustrated in Figure 1.2) is a cap dependent process that outlines initiation from the interactions of the 40S ribosomal subunit with mRNA through to generation of an 80S ribosome competent for polypeptide chain elongation.

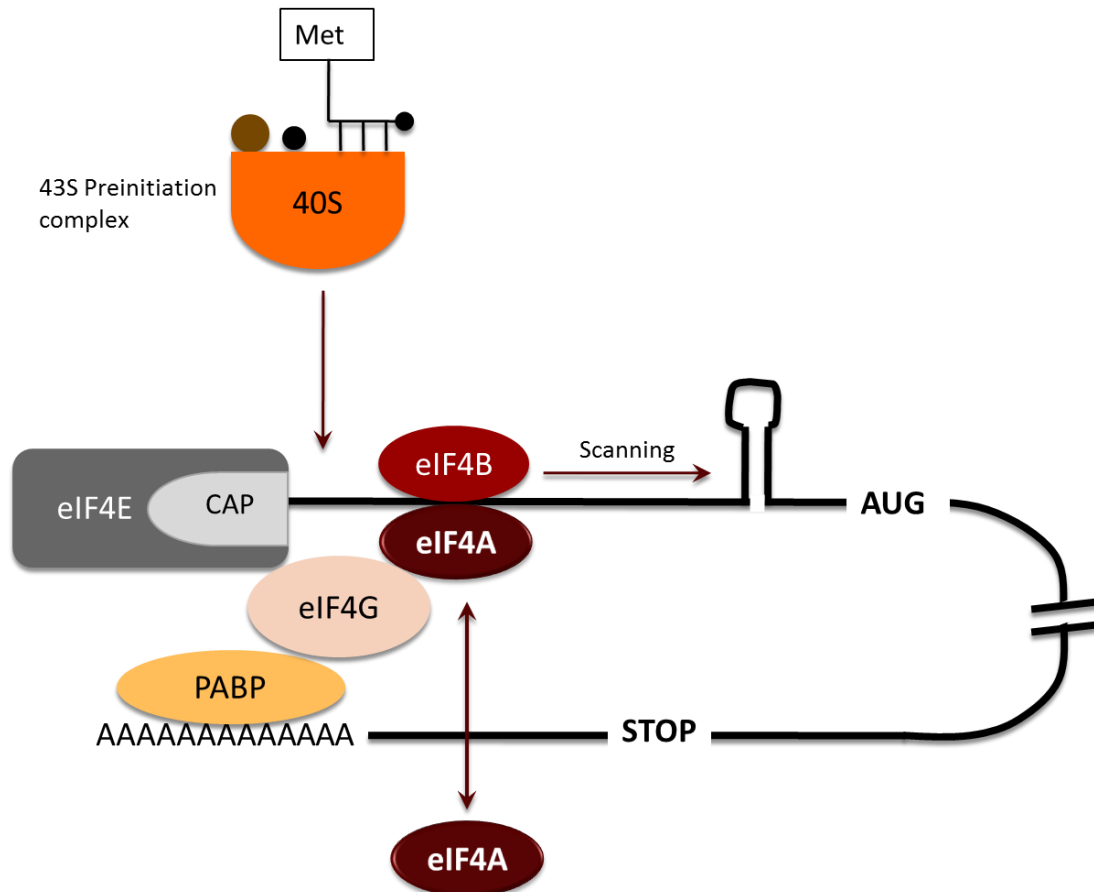


Figure 1.2: The scanning model of cap dependent translation initiation. Figure adapted from de la Cruz *et al* (1999) and Altman *et al* (2010).

The process of scanning begins before any ribosomal input with recognition of a 5' terminal modification, the 7-methylguanosine (m⁷GpppN or m⁷G) cap structure, on the mRNA by the eIF4F complex which consists of three components eIF4E, eIF4G and eIF4A. The cap binding protein eIF4E is responsible for the direct interaction with the aforementioned cap structure and is held in position near the mRNA through the duration of the scanning for the AUG codon by the scaffolding protein eIF4G.

For successful scanning, the initiation procedure requires an unfolded section of mRNA which is often not the lowest energy conformation. Subsequently any and all secondary

structures present in the leader mRNA must be removed. Unwinding and/or rearrangement of the duplex mRNA in the 5' UTR (5' untranslated region) are mediated by eIF4A's ability to "melt" the secondary structure with its ATP dependent RNA helicase activity. However, eIF4A has low affinity for RNA (*in vitro*) hence, the RNA stimulated ATPase and unwinding activities shown by eIF4A have been described as low (Linder and Jankowsky 2011). Furthermore, Rogers, Richter *et al* (2001) illustrated that the RNA helicase activity observed during translation initiation can be stimulated by co factors eIF4B, eIF4H, or as a constituent of the eIF4F complex.

The scaffolding protein eIF4G has been proposed to promote translation through mRNA circularisation. The interactions of eIF4G with both the cap bound eIF4E and poly (A) binding protein (PABP), which associates with the 3' poly (A) mRNA tail, cause mRNA pseudo-circularisation which may provide a framework for a closed loop model of translation (Lackner and Bähler 2008; Kronja and Orr-Weaver 2011). A closed loop model would explain how known regulatory elements of translation that are found in the 3' UTR are able to effect translation despite the fact that this process starts at the far end of the mRNA (Wilkie, Dickson *et al.* 2003; Lackner and Bähler 2008). Additionally, PABP interacts with eIF4B, resulting in suggestions that this interaction leads to stimulation of the PABP - poly (A) mRNA tail interaction as well as stimulation of eIF4A's helicase activity (Wilkie, Dickson *et al.* 2003).

The interaction of eIF4F with the mRNA, as well as the possible influence of mRNA pseudo- circularisation, activates the mRNA allowing the scanning process to commence. Activation allows the binding of a preassembled 43S preinitiation complex (PIC), which in mammals is shown to interact with eIF4G (Lackner and Bähler 2008), to the mRNA near the m⁷G cap structure (Sonenberg and Hinnebusch 2009). The PIC, which consists of several eIFs including eIF2 and eIF3 and the 40S ribosomal subunit loaded with an initiator methionyl-tRNA (Met-tRNA_i), then scans the 5' UTR, in the 5' to 3' direction, searching for the AUG start codon that matches the loaded Met-tRNA_i (Mendez and Richter 2001). Upon the discovery of this start codon the release of several eIFs is triggered which allow the recruitment of the large 60S subunit and subsequent formation of the 80S ribosome. This 80S ribosome formation signals the end of the initiation process prompting the start of the elongation phase (Sonenberg and Hinnebusch 2009).

1.2.3 Cap independent Initiation

Although cap dependent translation initiation is the most common for cellular mRNAs, there are a subset of eukaryotic mRNA which can circumvent the scanning process, initiating translation through the use of internal ribosomal entry sites (IRES) (Lackner and Bähler 2008; Sonenberg and Hinnebusch 2009). IRES have been described as cis-regulatory elements which contain no obvious consensus sequence. Present in the 5' UTR these elements achieve translation initiation independently of the interactions of eIF4F with mRNA, hence the term cap independent translation (Kronja and Orr-Weaver 2011). IRES elements are common in viral mRNAs allowing the continuation of their translation when the eIFs required for cap dependent translation are inhibited (Sonenberg and Hinnebusch 2009). In contrast, yeast IRES elements are far less prevalent and it is unknown whether this mechanism is commonly used to initiate translation in any yeast mRNA's (Zhou, Edelman et al. 2001). One such circumstance where cap independent initiation has proven crucial is in the observed function of the *URE2* gene. The *URE2* IRES element is able to produce a second, shorter form of Ure2p than that which is produced by cap dependent initiation. This IRES element is necessary as the two protein forms act with different functions; the shorter protein in nitrogen assimilation and the longer in the formation of prion like aggregates (Reineke and Merrick 2009).

1.2.4 Translation initiation and eIF4A in *Saccharomyces cerevisiae*

Saccharomyces cerevisiae, brewer's or baker's yeast, is considered to be a model eukaryotic organism because of its small size, rapid doubling time, haploid/diploid life cycle and economically favourable conditions for growth (Gershon and Gershon 2000). Furthermore, *S. cerevisiae*'s comparatively simple genome, genetic tractability and range of unparalleled genetic tools which can be applied, such as the yeast genome deletion sets (YGDS) and yeast GFP clone collection (Winzeler, Shoemaker et al. 1999; Tong, Evangelista et al. 2001; Huh, Falvo et al. 2003; Tong and Boone 2006), make it a model organism for use throughout this study.

1.2.4.1 Translation Initiation and eIF4A

Translation initiation is a highly homologous process between the mammalian and yeast systems. However similar the proposed function of eIF4A during both derivatives of initiation, there are subtle differences in how eIF4A is believed to achieve its function. In

order for eIF4A to contribute its helicase activity in either system, an association between eIF4A and mRNA needs to be established. However, this is not the only association eIF4A is able to make and during translation initiation eIF4A can be considered in two forms when associated with mRNA; either free of further interactions or bound in the eIF4F complex following interaction with eIF4E and eIF4G (Lanker, Müller et al. 1992).

In the free form of eIF4A, the helicase activity proceeds bi-directionally, from either 5' to 3' or 3' to 5', with no bias (Rogers, Komar et al. 2002). In addition, whilst in this form eIF4A is considered to be a non processive helicase meaning the average number of base pairs unwound per helicase binding event is considered low (Rogers, Richter et al. 2001; Betterton and Jülicher 2005). As part of the eIF4F complex, the helicase activity of eIF4A, and its subsequent function in translation initiation, improves, becoming processive (higher average number of base pairs unwound per helicase binding event) as well as proceeding uni-directionally in the 5' to 3' direction (Rogers, Richter et al. 2001; Linder and Jankowsky 2011).

Until recently, it was believed that unlike mammalian eIF4A, yeast eIF4A was only found in the free form as eIF4A could not be co-isolated with the remaining components of eIF4F: eIF4E and eIF4G (Rogers, Komar et al. 2002; Schutz, Bumann et al. 2008). There is low homology between mammalian and yeast eIF4G, of particular interest is a domain present in the C terminal region of mammalian eIF4G that yeast eIF4G lacks. In mammalian systems this region includes a second eIF4G-eIF4A binding site which has been suggested as a possible requirement for stable inclusion of eIF4A in the eIF4F complex (Dominguez, Kislig et al. 2001). However, crystal structures of core regions of yeast eIF4G interacting with full length yeast eIF4A have since been obtained illustrating both free and bound forms are crucial for translation initiation in both the yeast and mammalian processes (Schutz, Bumann et al. 2008).

As previously stated, mammalian eIF4A activity can be stimulated by the presence of eIF4B and eIF4H (Rogers, Richter et al. 2001). However in yeast, this increase in processivity, as a result of stimulation, is not as well defined. The yeast eIF4B homolog (encoded by *TIF3*), is non-essential for cell viability and its role in translation initiation is

unclear, furthermore there is no known homolog of eIF4H seen in yeast (Altmann and Linder 2010).

1.2.4.2 Homology of eIF4A

The yeast DEAD box protein family, in contrast to the mammalian, is much smaller, 26 and 37 members respectively (Linder and Jankowsky 2011). However, as testament to their required functions, most yeast DEAD box proteins have conserved counterparts in higher eukaryotes (Linder 2006). Mammalian systems encode for three isoforms of eIF4A: eIF4AI and eIF4AII which are involved in translation initiation and eIF4AIII which is involved in the non-sense mediated decay (NMD) pathway. The yeast homolog of mammalian eIF4A, first isolated as a suppressor of a mitochondrial missense mutation, is encoded by the paralogous genes *TIF1* and *TIF2* (Linder and Slonimski 1989). These paralogs show an almost identical sequence in their coding region (barring six synonymous changes in the gene sequence) and both encode for an identical 395 amino acid protein (Linder and Slonimski 1988). At the protein level, yeast eIF4A shares a 66% overall identity and 82% amino acid sequence similarity with its mammalian homolog (Dominguez, Kislig et al. 2001). Despite this high homology mammalian eIF4A cannot substitute for the yeast factor *in vivo* and is not functional in a yeast *in vitro* translation system (Dominguez, Kislig et al. 2001). Furthermore, although mammalian eIF4A cannot rescue yeast in the absence of yeast eIF4A, rescue has been observed between eIF4A of the murine and drosophila systems (Altmann and Linder 2010).

Whilst the sequence identity across all eukaryotes of eIF4AIII to the other isoforms is 65%, the sequence identity of eIF4AI and eIF4AII is as high as 90-95%. However similar these two proteins are, they remain functionally distinguishable as illustrated by the existence of differing expression patterns between tissues (Rogers, Komar et al. 2002). In contrast to other eukaryotes, currently *TIF1* and *TIF2* have not been proven to be functionally distinguishable (Linder and Slonimski 1989). This retention of protein sequence and apparent retention of function between a pair of duplicates is unusual, and understanding its importance in yeast eIF4A is an underlying motivation for the research presented in this thesis.

1.3 Duplicate genes

1.3.1 The need for duplication

The theory of evolution by gene duplication, although its technicalities are debated by many evolutionary geneticists, is considered a general principle of biological evolution (Zhang 2004). Gene duplication provides raw genetic material which can be adapted for functional innovation, implying new genes are not constructed *de novo* but are co-opted from existing genes (Conant and Wolfe 2008). A functional bias for the retention of duplicated genes which function as transcription factors, kinases, particular enzymes and transporters has been identified in a variety of organisms including humans and yeast (Conant and Wolfe 2008). In *S. Cerevisiae* the retention rate of ribosomal proteins highlights this functional bias as 59 of the 78 duplicate gene pairs have retained two copies (Komili, Farny et al. 2007). Amongst these are components of the yeast translation initiation machinery other than eIF4A, which is produced by the paralogous genes *TIF1* and *TIF2*, such as the homolog of eIF4G which is produced by the duplicate genes *TIF4631* and *TIF4632*.

1.3.2 Duplication methods

Duplication events can occur on two scales; small scale duplication (SSD), which is a continuous process involving a single gene or a small pool of genes; or whole genome duplication (WGD) (Davis and Petrov 2005; Conant and Wolfe 2008). Although rare, WGD is postulated to have played a major role in the evolution of different species such as the vertebrate lineage and has been attributed to shaping the facultative anaerobic lifestyle of the *Saccharomyces* lineage (Cliften, Fulton et al. 2006).

Although both these duplication processes involve gene duplication, the likelihood of fixation of the newly duplicated gene or genes in the genome and subsequent preservation of both gene duplicates by divergence of gene function, differs between them (see Figure 1.3). While SSD requires an independent mutation event followed by selective pressure to fix the gene in the genome, WGD is characterised not by independent mutations but duplication with the entire genome. Subsequently, genes deemed to be advantageous are not immediately fixed in the genome, like after SSD events, but must survive the period of genome shrinkage that follows WGD to eliminate

functional redundancy (Davis and Petrov 2005). As a result of these procedural differences, gene duplication via SSD or WGD presents differing evolutionary opportunities (Conant and Wolfe 2008).

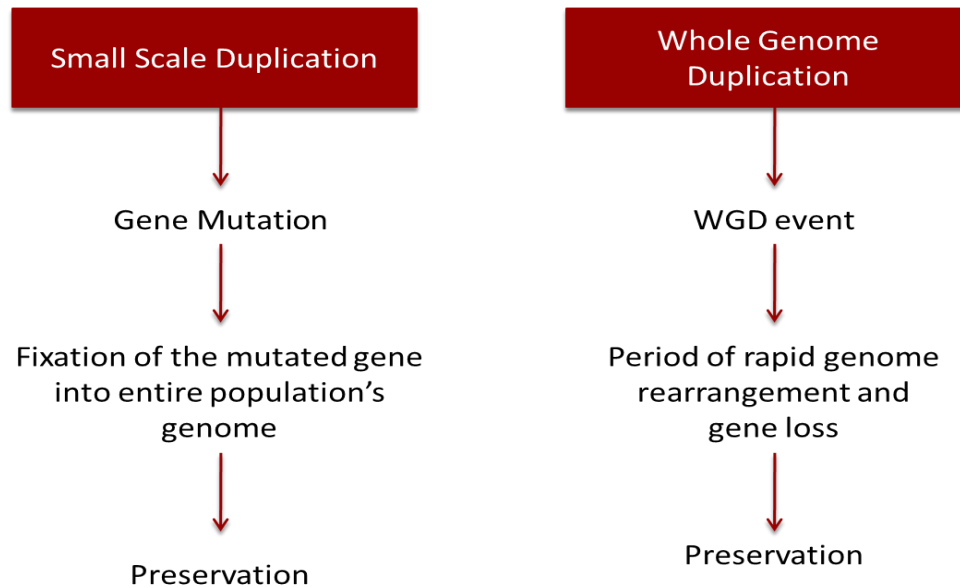


Figure 1.3: The process leading to long-term duplicate survival. Figure adapted from Davis and Petrov (2005).

1.3.3 Gene functional bias

Genes duplicated by either SSD or WGD, at the time of duplication, present similar levels of codon bias to each other, which are in favour of higher levels of gene expression, as well as a tendency to arise from slowly evolving genes, however, the two duplicate sets are enriched for different functional classes of genes (Davis and Petrov 2005; Conant and Wolfe 2008). Compared with proteins retained after SSD, those proteins retained after WGD events more strongly retain the aforementioned bias towards higher-expression favouring codon bias (Conant and Wolfe 2008), show lower frequency of essential genes and have a higher synthetic lethality rate thus indicating that the duplication is not contributing to robustness through direct back-up of functionality (Guan, Dunham et al. 2007). Moreover, WGD events result in proteins that diverge more in their expression pattern and upstream regulatory region than duplicates retained from SSD, implying their role in fine tuning expression levels (Guan, Dunham et al. 2007). Independent of sequence similarity, paralogous genes from WGD share more protein interactions and

biological functions than SSD duplicates (Guan, Dunham et al. 2007; Conant and Wolfe 2008).

As previously stated, a functional bias for the retention of certain duplicated genes exists. Davis and Petrov (2005) were among the first researchers to investigate gene expression of duplicates generated by SSD or WGD and were able to propose a cause for some of this functional bias. They reported WGD as showing enrichment in ribosomal proteins in contrast to SSD which showed a lack of transcriptional regulator proteins and overabundance of enzyme regulator proteins. The '*dosage balance hypothesis*', supported by the types of genes and gene functions resulting from both WGD and SSD, has been postulated as a possible explanation for this duplication scale-dependent functional divergence (Conant and Wolfe 2008).

1.3.3.1 Dosage Balance Hypothesis

Evolution via duplication may provide genes with novel functions adaptable to the genome; however, these adaptations can become precluded if the initial duplication event is selectively disadvantageous to the genome. This notion that duplication might interfere with highly constrained cellular systems and hence be selectively unfavourable is termed the '*dosage balance hypothesis*' (Conant and Wolfe 2008). In order to retain these so called duplication-resistant genes post duplication, the environmental stoichiometry must be maintained.

One way of achieving this is to regulate the stoichiometry at the genomic level. The stoichiometric environment of the genome is changed drastically in SSD compared with WGD which maintains the stoichiometry ratio as the size of the genome doubles. The stoichiometric environment becomes critically important when genes are inherently expressed in a specific ratio with other loci; either as a result of their involvement in a multi-protein complex or function in a biochemical pathway, and a deviation from this stoichiometry may be catastrophic (Davis and Petrov 2005). As such duplication-resistant genes are unlikely to become fixed in the genome after SSD but are likely to become fixed after WGD because of the deleterious effects on the genome if gene function is not maintained (Davis and Petrov 2005; Conant and Wolfe 2008). The other alternative is to

compensate for the imbalance in gene dosage by regulation of protein expression through transcriptional / translational regulation, mRNA degradation and the like.

1.3.4 Surviving duplication – Gene fixation

Complete redundancy between duplicated genes is evolutionarily unfavourable and there are many possible and more favourable evolutionary fates these genes can take (Musso, Costanzo et al. 2008).

One of the copies may be redundant, accumulate deleterious mutations leading to development of a non-functional pseudogene. If the ancestral gene will benefit at a higher gene dosage natural selection will allow for functional preservation of both duplication genes. Upon preservation the paralogous genes may functionally diverge via two commonly accepted pathways: Neofunctionalisation, considered the classic scenario, where the ancestral gene function is maintained in one of the paralogous genes whilst the other evolves a new biochemical function; Subfunctionalisation, the partitioning of ancestral gene functions between the paralogous genes so that their joint function/activity is equal to that of the ancestral gene (Marques, Vinckenbosch et al. 2008).

1.3.5 Genetic redundancy

When subjected to genetic perturbations or external stimuli the cellular constituents of an organism may change. It is the fundamental property of all living systems to attempt survival of such changes. The lack of change to phenotypic variations is termed '*genetic robustness*' (Borenstein and Ruppin 2006). In certain cases, an organism's ability to demonstrate genetic robustness has proven to reduce its susceptibility to mutations such as protein tolerance to amino acid substitutions, gene dispensability in yeast, and the error tolerance of complex biological networks. Subsequently, if the consequences attributed to such mutations are left unresolved, the result could be disastrous for cell viability (Borenstein and Ruppin 2006).

Duplicate gene pairs can provide an organism with an added level of genetic robustness as a consequence of the duplication process. As previously stated (section 1.3.4), duplicate genes can endure many different evolutionary fates. Of these, fates which result in the partitioning of ancestral function between paralogous genes may provide

the genome with genetic redundancy (Louis 2007). Genetic redundancy is the ability of gene A to “back-up” the gene function usually attributed to gene B, or with reference to duplicate genes, the ability of one gene, in the absence/inactivation of its paralogous gene, to provide the function usually generated by the paralogous gene (Meyer 2003).

Accordingly, genetic redundancy following duplication is only possible because paralogous genes evolve from a single gene with a shared function and potentially still retain some level of functional overlap. For example one third of all surveyable WGD duplicates have retained the ability to buffer for the loss of their respective paralogous gene under standard laboratory growth conditions (Musso, Costanzo et al. 2008). Furthermore, the protection offered by duplicate genes is believed to gradually attenuate over evolutionary time as their sequences diverge and the functional overlap reduces (Brookfield 2003).

In contrast, there are examples of duplicate gene pairs which have retained their functional overlap over an extended evolutionary period or even speciation; such is the case for the yeast genes *ARE1/ARE2* and the mammalian homologs *ACAT1* and *ACAT2*. Subsequently it has been suggested that redundancy may form a part of a larger regulatory network (Kafri, Levy et al. 2006).

Genetic redundancy presents a major hurdle for researchers trying to elucidate the function of a particular gene as the application of standard techniques may not have a noticeable effect due to the functional substitution of the duplicate gene (Meyer 2003). An illustration of this comes from research of Rotenberg who, using knock out (KO) technologies, was investigating whether duplicate ribosomal protein genes were functionally distinct. Rotenberg concluded that duplicate ribosomal proteins were functionally redundant and the only difference observed was that the more highly expressed duplicate played a more significant cellular role (Rotenberg, Moritz et al. 1988). However, recent studies present evidence for gene specific defects among paralogous genes including defects in sporulation, actin organisation and bud site selection. Thus illustrating a much more complex relationship challenging this conclusion as well as demonstrating the flaw in directly attributing growth rate to observed function (Komili, Farny et al. 2007).

1.3.6 Duplication in *S. cerevisiae*

1.3.6.1 Ancient Gene Duplication

The *S. cerevisiae* genome is believed to have formed from an ancient round of whole genome duplication approximately 100 million years ago (Davis and Petrov 2005). After genomic restructuring, shrinkage and preservation of the functionally evolved genes, approximately 5770 genes remained in the genome of which it is thought only 10 – 11 % of the duplicated genes were maintained (Cliften, Fulton et al. 2006; Komili, Farny et al. 2007). Evidence for an ancient WGD event was first proposed by Wolfe and Shields who identified 55 duplicate regions, encapsulating 376 pairs of homolog genes with an average amino acid sequence similarity of 63% (Wolfe and Shields 1997). Regions (or blocks), were identified in 70% of the genome and Wolfe and Shields concluded the resulting pattern of duplicates identified in the genome was statistically unlikely to form via SSD rather than WGD (Wolfe and Shields 1997; Kellis, Birren et al. 2004; Cliften, Fulton et al. 2006).

Of particular interest to this study, is the 41st block defined by Wolfe and Shields which identified regions of homology between sequences of chromosomes X and XI, which bear the genes *TIF1* and *TIF2* respectively. Although unwise to assume without knowledge of the gene functions of the ancient genome, it seems plausible that *TIF1/TIF2*, as well as their three nearest chromosomal neighbours; *RPS21A/ RPS21B* (small ribosomal subunits), *GLG1/GLG2* (initiators of glycogen synthesis), and *YUR1/KTR2* (mannosyltransferases involved in N linked protein glycosylation), are all examples of gene preservation from subfunctionalisation (SGD. 2012).

1.4 Regulation of protein synthesis

The regulation of protein expression can be separated into regulation of gene expression, be it transcriptional or posttranscriptional control, or regulation of the proteins themselves by posttranslational modification and protein degradation (Lackner and Bähler 2008).

However, as previously stated, the duplicate genes *TIF1* and *TIF2* convey almost identical sequence similarity in their coding regions (six synonymous changes) and both encode for an identical protein. Subsequently, regulation at the protein level here, will

indistinguishably affect both genes' identical protein products. Furthermore, as literature suggests that complete redundancy between gene pairs is evolutionary unfavourable (Musso, Costanzo et al. 2008) it is essential, in order to avoid complete gene redundancy, that differential regulation of *TIF1* and *TIF2* be achieved at the gene level.

It is well established that gene expression is regulated at multiple levels which include transcription, RNA processing/export, translocation and translation (Lackner and Bähler 2008). Transcriptional regulation can be achieved at many stages throughout the transcript's lifetime. Instances include co-transcriptional modifications such as addition of the m7G cap, splicing of introns out of pre-mRNA to generate functional mRNA and polyadenylation. In contrast to transcriptional regulation, translation regulation permits for more rapid changes in expressed protein concentrations making it a suitable method for maintaining cellular homeostasis (Sonenberg and Hinnebusch 2009). As previously described (see section 1.2) there are three main phases of translational control, all of which have potential targets for regulation. Due to a relative plethora of proteins and initiation factors involved at the initiation phase, it is the initiation phase where most of the translation regulation occurs (Lackner and Bähler 2008). Regulatory events during translation are usually reversible as they are often mediated through reversible protein modifications such as phosphorylation of initiation factors (Sonenberg and Hinnebusch 2009).

Although *TIF1* and *TIF2* have a high sequence similarity, it is of particular interest that to note that their flanking UTRs bear almost no similarity. Furthermore, although many features of mRNA contribute to translation, it has been suggested that most control elements are located within the untranslated regions (Wilkie, Dickson et al. 2003), including transcription factors and cis regulatory elements. For example, iron regulatory proteins (IRPs) which recognise a stem loop structure known as the iron response element (IRE) are contained in the 5' UTR. In response to intracellular iron concentration this interaction, depending on the location of the IRE in the 5' UTR, can act to either inhibit the 40S ribosomal protein binding the mRNA or impede the scanning mechanism of the pre-initiation complex (Wilkie, Dickson et al. 2003). Contained in the 3' UTR are the 3' UTR transcript localisation motifs which are crucial for the development of several organisms across the evolutionary spectrum. These motifs regulate mRNA localisation to

discrete sites allowing the controlled synthesis of proteins at target locations and production of morphogen gradients (Farooq, Choi et al. 2012).

In many mRNA AU-rich elements (AREs) are present in the 3' UTR and accompanied by their specific ARE-binding proteins are able to influence gene expression through mRNA turnover and translation. One such example is the ARE-binding proteins of the AUF1 family which promote the degradation of mRNA encoding cytokines or cell cycle regulators as well as inducing translation of MYC proto-oncogene mRNA (Lackner and Bähler 2008). Studies in human RKO colorectal carcinoma cells illustrated that the ARE-binding protein TIAR, in response to ultra violet radiation, binds the 3' UTR of eIF4A and eIF4E, potentially suppressing their translation (Mazan-Mamczarz, Lal et al. 2006). A protein-protein BLAST search of TIAR indicates the closest yeast ortholog is PUB1 which is a Poly (A) and RNA binding protein that binds many mRNA and is required for mRNA stability. This, however, was not a full sequence overlap and alignment score of 26 (where alignment score is the number of identities between the two sequences, divided by the length of the alignment, and represented as a percentage) was generated by ClustalW alignments. Although, in the aligned sequences where they are most common they share a 41% identity and 60% sequence similarity.

1.5 Aims and Objectives

1.5.1 Aims

The primary aim of this study was to evaluate whether *TIF1* and *TIF2* can be individually regulated by small molecule intervention.

This aim fits into broader programmes; 1) To understand the evolutionary benefit to the retention of two genes producing an identical eIF4A protein in *S. cerevisiae*; and 2) to evaluate the potential for small molecule regulation of translation for therapeutic intervention in yeasts or potentially other eukaryotes.

The strategy employed to identify differential regulation between *TIF1* and *TIF2* was to construct a dual fluorescent probe where either a red or green fluorescent protein (RFP or GFP) marker was fused to the coding region of *TIF1* or *TIF2*. This creates a functional reporter strain with distinguishable *TIF1* and *TIF2* markers that can be used to identify

their respective gene functions in a variety of cellular conditions. To encompass as many different variables as possible the dual fluorescent reporter was subjected to a morphological high throughput screening (HTS) utilising the LOPAC¹²⁸⁰ Library of Pharmacologically-Active Compounds, chosen as a library of environmental variables rather than as a library of bio-actives of interest.

We hypothesise that retention of the *TIF1* and *TIF2* duplicates arises from the need to respond to a wide variety of cellular needs in *S. cerevisiae*. These may include environmental and nutrient stresses or cell cycle requirements. These may be responded to by the up regulation or suppression of total eIF4A through one of the *TIF* genes or through regulation of the stability or localisation of a specific *TIF* mRNA. Discovery of small molecule regulators of TIF1p or TIF2p will provide molecular tools for further understanding of this system.

1.5.2 Experimental Objectives

- 1) Creation of a TIF1-RFP fusion protein
- 2) Utilisation of commercially available TIF1-GFP and TIF2-GFP yeast strains and the generated TIF1-RFP fusion protein in the construction of a dual fluorescent reporter strain capable of distinguishing between the duplicate genes *TIF1* and *TIF2*
- 3) Developing the methodology for the use of these fluorescent reporters in high throughput screening
- 4) Screening of the LOPAC¹²⁸⁰ Library of Pharmacologically-Active Compounds with the commercially available single TIF1-GFP and TIF2-GFP strains using high throughput confocal microscopy
- 5) Screening of the LOPAC¹²⁸⁰ Library of Pharmacologically-Active Compounds with the dual fluorescent reporter strain using high throughput confocal microscopy
- 6) Cheminformatic analysis of the compounds identified in the high throughput screens.

These objectives will address the aim through identification of lead compounds for further study. Extensive validation and determination of the pathways through which TIF1 and TIF2 are regulated is beyond the scope of this Masters project.

2 Methods

2.1 Materials and Equipment

2.1.1 Reagents

The amino acids including monosodium glutamate (MSG, L-glutamic acid sodium salt hydrate) were purchased from Sigma-Aldrich (Auckland, New Zealand); Bacto Peptone, Bacto Triptone, Bacto Yeast Extract, and Bacto Yeast Nitrogen Base were purchased from DIFCO (Detroit, MI, USA). D-Glucose (Dextrose) was purchased from Fisher Scientific (Auckland, New Zealand) and was dissolved in water to a concentration of 40% (w/v), then autoclaved and stored at room temperature. As required, D-Glucose was added to medium to a final glucose concentration of 2% (w/v).

Unless stated otherwise, all other reagents were sourced from Sigma-Aldrich (Auckland, New Zealand).

2.1.2 Chemical libraries and individual chemicals

The LOPAC¹²⁸⁰ Library of Pharmacologically-Active Compounds (LOPAC library) was purchased from Sigma-Aldrich (Auckland, New Zealand). The LOPAC library was originally arranged in 96-well, one compound per well format covering 16 plates. However, a working stock set of the LOPAC library which had previously been diluted in DMSO, from 10 mM to 1 mM, was employed in this study.

In addition to the LOPAC library, two research compounds also diluted in DMSO to 1 mM, DD1 and DD2 were employed in this study. DD1 and DD2 were kindly gifted by Dylan Davies from the School of Biological Sciences, Victoria University of Wellington. DD1 and DD2 were screened in tandem with the LOPAC library as both were added to the library at the expense of two separate internal control wells, maintaining each plate's border control.

2.1.3 Yeast strains

All deletion strains were purchased from Thermo Scientific-Open Biosystems (Huntsville, AL, USA) as a part of the YKO MATa Strain Collection and have the genotype *XXXΔ::KANMX4, his3Δ1, leu2Δ0, met15Δ0* and *ura3Δ0* (Winzler, Shoemaker et al. 1999).

Strains bearing the green fluorescent protein fusions were purchased from Invitrogen (Carlsbad, CA, USA) as part of the Yeast GFP Clone Collection (Huh, Falvo et al. 2003). Strains of this collection are of the MAT α mating type which has a genotype XXXGFP-HIS3MX6, *his3 Δ 1*, *leu2 Δ 0*, *met15 Δ 0*, and *ura3 Δ 0*.

Strains of *S. cerevisiae* utilised in the construction of the dual fluorescent probe were of the MAT α mating type of the BY4741 background which have a genotype of *his3 Δ 1*, *leu2 Δ 0*, *met15 Δ 0* and *ura3 Δ 0*.

Two dual fluorescent constructs were assembled during this study; TIF1-GFP, TIF2, TIF1-RFP which has the genotype MAT α , *tif2 Δ ::KANMX4*, *his3 Δ 1*, *leu2 Δ 0*, *met15 Δ 0*, *ura3 Δ 0*, *TIF1-GFP-his3.MX6*, *pRS316-URA3 TIF1-RFP*, as well as the construct TIF2-GFP, TIF1, TIF1-RFP which has the genotype MAT α , *tif1 Δ ::KANMX4*, *his3 Δ 1*, *leu2 Δ 0*, *met15 Δ 0*, *ura3 Δ 0*, *TIF2-GFP-his3.MX6*, *pRS316-URA3 TIF1-RFP*.

2.1.4 Growth media

Yeast strains used in this study were cultured according to standard yeast methods (Amberg and Burke 2005) in one of the following media:

2.1.4.1 Yeast Peptone Dextrose (YPD)

YP medium was prepared according the following recipe: 8 g yeast extract, 16 g Bacto Peptone, and 0.096 g adenine were mixed into 760 mL of distilled water and sterilized by autoclaving; 40 mL of 40% glucose solution was added. G418 was added at a final concentration of 200 μ g/ml when required.

2.1.4.2 Synthetic Complete (SC)

Synthetic Complete (SC) medium was prepared using the following recipe: 1.7 g of Bacto Yeast Nitrogen Base (without amino acids or ammonium sulphate), 1 g of MSG, and 2 g of “amino acid supplement powder mix” were mixed into 760 mL of distilled water and sterilized by autoclaving; 40 mL of 40% glucose solution was added. G418 was added at a final concentration of 200 μ g/ml when required.

“Amino acids supplement powder mix” contains: 3 g adenine, 2 g uracil, 2 g inositol, 0.2 g para-aminobenzoic acid, 2 g alanine, 2 g arginine, 2 g asparagine, 2 g aspartic acid, 2 g cysteine, 2 g glutamic acid, 2 g glutamine, 2 g glycine, 2 g histidine, 2 g isoleucine, 10 g

leucine, 2 g lysine, 2 g methionine, 2 g phenylalanine, 2 g proline, 2 g serine, 2 g threonine, 2 g tyrosine, 2 g tryptophan, and 2 g valine.

2.1.4.3 Synthetic Dropout (SD)

Drop-out (DO) powder mixture is a combination of the listed ingredients for SC media minus the appropriate supplement; 2 g of the DO powder mixture is used per litre of medium.

SD –URA (synthetic dropout missing uracil): as for SC medium but without uracil in the “amino acid mix”; glucose added at final concentration of 2% (w/v).

SD –HIS (synthetic dropout missing histidine): as for SC medium but without histidine in the “amino acid mix”; glucose was added at final concentration of 2% (w/v).

2.1.5 Bacterial growth media

All bacteria were cultured at 37 °C overnight in Luria Bertani (LB) media supplemented with 100 µg/ml of ampicillin.

2.1.5.1 Luria-Bertani (LB)

LB medium was prepared according to the following recipe: 10 g Bacto Tryptone, 5 g Bacto Yeast Extract, and 10 g NaCl were mixed into 800 mL of distilled water and the pH adjusted to 7.5 with 1 M NaOH solution. The volume was finally adjusted to 1 L with distilled water, the medium sterilized by autoclaving and stored at room temperature.

2.1.6 Plasmids used in this study

The transformed *E. coli* bearing the plasmid pYM43 containing the red fluorescent protein RedStar2 and the natNT2 cassette was purchased from EUROSCARF (Institute of Molecular Biosciences, Johann Wolfgang Goethe-University Frankfurt, Germany). This plasmid was created as part of the “PCR toolbox” (Janke, Magiera et al. 2004).

The transformed *E.coli* bearing the plasmid pRS316 containing the *URA3* selectable marker was kindly donated from D Bellows (School of Biological Sciences, Victoria University of Wellington). This plasmid was created as part of a series of yeast shuttle vectors allowing greater efficiency in the manipulation of DNA in *Saccharomyces cerevisiae* (Sikorski and Hieter 1989).

2.2 Methodology

2.2.1 DNA preparation and manipulation

2.2.1.1 Genomic DNA preparation

Isolation and purification of *Saccharomyces cerevisiae* genomic DNA was achieved using the MasterPure Yeast DNA Purification Kit according to the manufacturers' guidelines (Epicentre Biotechnologies, Madison, Wisconsin, USA). In brief, yeast cells were pelleted from saturated 1.5 mL cultures by centrifugation in a microcentrifuge tube at 10,000 rpm for 5 minutes. The pellet resuspended in 300 µL of Yeast Cell Lysis Solution by vortexing before samples incubated at 65 °C for 15 minutes. Samples were then placed on ice for 5 minutes before addition of 150 µL of MPC Protein Precipitation Reagent and vortexed for 10 seconds. The cellular debris were pelleted by centrifugation for 10 minutes at >10,000 rpm and the supernatant transferred to a clean microcentrifuge tube. Genomic DNA was then precipitated by isopropanol precipitation (section 2.2.1.4) and subsequent ethanol precipitation (section 2.2.1.3) and then suspended in 35 µL of Tris EDTA (TE) buffer.

In order to degrade any RNA present the purified DNA was then treated with 5 µg of RNase A and incubated at 37 °C for 30 minutes. The RNase A was then removed with phenol:chloroform (section 2.2.1.2) followed by ethanol and isopropanol precipitations (method 2.2.1.3 and 2.2.1.4) The size of the DNA was assessed by DNA electrophoresis (method 2.2.1.5) and quality assessed by DNA quantification (method 2.2.1.6).

2.2.1.2 Phenol Chloroform Isoamyl alcohol DNA extraction

Phenol chloroform isoamyl alcohol DNA extraction was performed according to standard methods (Ausubel 1988). An equal volume of phenol: chloroform: isoamyl alcohol (25:24:1) was added to the DNA sample and vortexed to mix before centrifugation at 16,000 xg for 10 minutes. The aqueous layer was removed and added to an equal volume of chloroform, mixed by a brief vortex and centrifuged at 16,000 xg for 10 minutes. The aqueous layer was then removed and kept.

2.2.1.3 Ethanol precipitation

Ethanol precipitation was performed according to standard methods (Ausubel 1988). Sample DNA was prepared by the addition of 2.5 volumes of 96% Ethanol and 1/10 volume of 3 M sodium acetate (pH 5.3), mixed by gentle inversion and incubated at -20 °C

for 25 minutes. The sample was then centrifuged in a microcentrifuge for 10 minutes at 16,000 g collecting the pelleted DNA. The pellet was then washed with 1 volume of 70% ethanol, centrifuged at 16,000 g for 5 minutes, the resulting pellet was then air dried in the fume hood for 15 minutes and resuspended in 35 μ L of TE buffer. This was then incubated at 65 °C for 10 minutes to ensure the DNA had completely dissolved.

2.2.1.4 Isopropanol precipitation

Isopropanol precipitation was performed according to standard methods (Ausubel 1988). Briefly, ammonium acetate was added at a final acetate concentration of 2 M, followed by 0.7 volumes of isopropanol. After 15 minutes incubation at room temperature the precipitated DNA was recovered by micro-centrifugation and the pellet air-dried at room temperature for a further 15 minutes before being resuspended in TE buffer (1 mM EDTA, 10 mM Tris-HCl pH 7.5).

2.2.1.5 Agarose gel electrophoresis

Agarose gel electrophoresis was performed according to standard methods (Ausubel 1988). Briefly, electrophoresis was performed in 1% agarose gels, run in Tris Borate EDTA (TBE) buffer (89 mM Boric Acid, 2 mM EDTA disodium dihydrate, 89 mM Tris Base pH 8.3) with 0.5 μ g/ml ethidium bromide. All samples were dissolved in TE buffer and mixed 5:1 with 6x sample loading buffer (30% glycerol (v/v), 0.25% (w/v) bromophenol blue, 0.25% (w/v) cyanol xylene) prior to loading onto the gel. The 1Kb plus DNA ladder, prepared for gel loading as above, was used as a DNA size control and was purchased from Invitrogen (Carlsbad, CA, USA). Electrophoresis was performed at a constant 100 V and visualized on a transilluminator at 365 nm (UVItec, Cambridge, UK).

2.2.1.6 DNA quantification

Quantification of DNA was performed with the Sigma DNA Quantification Kit, DNA-QF (Sigma-Aldrich) according to the manufacturers' instructions. In brief, a 2 μ g/mL solution of the fluorescent H33258 dye was prepared prior to use. In a microtitre plate, 200 μ L of the dye solution was added to each of a series of DNA standards (calf thymus DNA), with known concentrations ranging from 20 - 2000 ng of DNA. A standard curve of H33258 dye fluorescence at known concentrations was generated by excitation at 360 nm and measuring fluorescence on a SpectraMax Plate reader at 460 nm. The resulting standard curve was used to measure the DNA quantity of 5 μ L of DNA sample.

2.2.1.7 Plasmid purification and isolation

Isolation and purification of plasmid DNA from DH5 α *E. coli* was performed with the Zyppy Plamid Miniprep Kit (Zymo Research Corporation, Irvine, CA, USA) according to the manufacturers' instructions. In brief, 600 μ L of a bacterial culture grown in LB medium was added to a 1.5 mL microcentrifuge tube before addition of 100 μ L of 7x Lysis buffer. The solution was then mixed by inversion and within two minutes 350 μ L of cold neutralization buffer was added and thoroughly mixed to ensure complete neutralization. The solution was then pelleted by centrifugation at 16,000 g for 4 minutes before the supernatant was transferred into the provided Zymo-Spin II column. The column was then placed in a collection tube, centrifuged for 15 seconds and the flow through discarded. The column was then placed back in the collection tube and 200 μ L of Endo-wash buffer was added before the column was centrifuged for a further 15 seconds. 400 μ L of Zyppy wash buffer was added to the column before centrifugation for 30 seconds and transfer of the column into a clean 1.5 mL microcentrifuge tube. Zyppy elution buffer, 30 μ L, was added directly to the column matrix and left to stand at room temperature for one minute before the plasmid DNA was eluted via 15 s of centrifugation. The plasmid DNA was then kept at 4 $^{\circ}$ C until required for further experiments.

2.2.1.8 Restriction digest of plasmid

The purified plasmid pRS316 was subjected to restriction digest by the restriction endonucleases XbaI and HindIII (New England Biolabs, Ipswich, MA, USA) by a modified version of the digestion protocol from New England Biolabs (New England Biolabs, Ipswich, MA, USA). In brief, 42.5 μ L of pRS316 suspended in LB media was added to a 1.5 mL microcentrifuge tube. To this, 5 μ L of 10x NEBuffer #2, 0.5 μ L of BSA and 1 μ L of both XbaI and HindIII were added, all of which were purchased from New England Biolabs (New England Biolabs, Ipswich, MA, USA). The solution was then thoroughly mixed and incubated at 37 $^{\circ}$ C for 12 hours to ensure complete digestion.

2.2.1.9 Post digestion plasmid clean up

The endonuclease enzymes utilised in the restriction digest protocol were denatured and removed along with any other contamination in order for future use of digested pRS316. This was performed with the PCR clean-up protocol which is a part of the Gel/PCR DNA

fragments extraction kit (Geneaid Biotech Ltd, Agoura Hills, CA, USA). In brief, 48 μ L of the digest reaction product was transferred to a 1.5 mL microcentrifuge tube along with 5 volumes of DF buffer and the two were mixed thoroughly. This solution was transferred to a DF column in a 2 mL collection tube and centrifuged at 16,000 g for 30 seconds. After the flow through was discarded, 600 μ L of Wash Buffer (ethanol added) was added to the centre of the DF column and left to stand for 1 minute. The column was then centrifuged at 16,000 g for 30 seconds, flow through discarded and then centrifuged for a further 3 minutes at 16,000 g in order to dry the column matrix. The dried DF column was then transferred into a new 1.5 mL microcentrifuge tube and 50 μ L of elution buffer was added into the centre of the column matrix. This was allowed to stand at room temperature for 2 minutes to ensure the elution buffer was completely absorbed before a final centrifugation for 2 minutes at 16, 000 g to elute the purified DNA. The purified DNA was then stored at 4 °C until required for future use.

2.2.1.10 Polymerase Chain Reaction (PCR)

The *TIF1*-RFP component module construction PCR was performed using the Qiagen Hotstar Taq DNA polymerase Kit with the following PCR reaction conditions (Bio-Strategy Ltd, Auckland, New Zealand).

Component module construction PCR was performed using the Techne TC-5000 thermal cycler (Bibby Scientific, Staffordshire, UK) at the optimised cycling conditions: 15 min initial denaturation at 95 °C, then 10 cycles of 1 min at 97 °C (denaturation), 30 sec at 54 °C (annealing), 2 min 40 sec at 68 °C (elongation). Followed by a further 20 cycles of 1 min at 97 °C (denaturation), 30 sec at 54 °C (annealing), 2 min 40 sec at 68 °C, for the first cycle and incrementally increasing by 18 s per cycle thereafter until the final cycle of 8 min 40 s at 68 °C (elongation). Following this was a final elongation for 5 min at 72 °C. The resulting PCR product (1 μ L) was subjected to DNA electrophoresis and DNA quantification as described in sections 2.2.1.5 and 2.2.1.6 respectively.

Reagent	Volume (μL)
10X buffer (15 mM MgCl ₂)	5
Q buffer	10
dNTP (5 mM)	2
MgCl ₂ (25 mM)	0.28
ddH ₂ O	29.47
Hotstar Taq (5 units/μ l)	0.5
Fwd Primer (100 pM)	0.5
Rev Primer (100 pM)	0.25
DNA Template	2
Total	50

Table 2.1: Reaction setup for the component module construction PCR using Qiagen Hotstar Taq DNA polymerase.

The primers used in the component module construction PCR, described in Table 2.2, were ordered from Integrated DNA technologies (IDT, Leuven, Belgium) and suspended in water to a final concentration of 100 pM.

Identification Number	Primer	Sequence
#484	Module 1 Forward	CAGTCACGACGTTGTAAAACGACGGCCAGTGAATTGTAATTC AGCAACAACATCCGATGCTT
#485	Module 1 Reverse	ACCTGCACCAGCTCCAGCTCCGTTCAACAAAGTAGCGATGTCTG GATGGCAATTCTTCAATTT
#486	Module 2 Forward	AATTGAAGAATTGCCATCCGACATCGCTACTTTGTTGAACGGA GCTGGAGCTGGTGCAGG
#378	Module 2 Reverse	TAGCCTCACAAGATACTTTTTTAAGAAGTTTTGTCTCCCTTAC AAGAACAAGTGGTGTC
#487	Module 3 Forward	ACTGAAGGTAGACACCACTTGTTCTTGTAAGGGAGACAAAAA CTTCTTAAAAAAGTATCTTGTGAGGCTATCTTG
#488	Module 3 Reverse A	CGGCTCCTATGTTGTGTGGAATTGTGAGCGGATAACAATTTTT GATGTACACTTTTTCTTTTCAG
#489	Module 3 Reverse B	CGGCTCCTATGTTGTGTGGAATTGTGAGCGGATAACAATTTTT GATGTACACTTTTTCTTTT

Table 2.2: Primers used in the TIF1-RFP component module construction PCR, their laboratory identification number, sequence, and acknowledgement of which primer sets match to make all component modules.

2.2.2 Dual fluorescent strain generation

2.2.2.1 Transformation

Transformation of *S. cerevisiae* was performed using a modified version of the lithium acetate/ single stranded carrier DNA/ PEG method (Gietz and Schiestl 2007). A single colony of each strain to be transformed (see section 2.1.3) was inoculated into YPD media and incubated overnight at 30 °C with constant agitation. The following day the culture was diluted to an OD₆₀₀ of 0.05 in YPD and returned to incubate at 30 °C until OD₆₀₀ reading reached 0.7-1.2. The cell density was then measured using a haemocytometer and the volume corresponding to 1×10^8 cells were removed. The cells were then pelleted by centrifugation at 4000 rpm and washed three times with distilled water. The procedure was then continued in one of two ways depending on how many plasmids were included in the transformation.

A) For single plasmid transformations

Cells were resuspended in transformation mix (33.33% PEG 3350 (w/v), 0.27 mg/mL single stranded salmon sperm DNA, 0.1 M lithium acetate) containing 1 µg of the plasmid DNA and incubated at 42 °C for 40 min.

B) For multiple plasmid transformation

Cells were resuspended in transformation mix (33.33% PEG 3350 (w/v), 0.27 mg/mL single stranded salmon sperm DNA, 0.1 M lithium acetate) containing 1 µg each of all the plasmid DNA to be transformed. For example the dual fluorescent strain required the addition of 1 µg of each pRS316, and the aforementioned Modules 1-3. This solution was then incubated at 42 °C for 40 min.

Subsequent steps of both transformation procedures were carried out identically. Samples were centrifuged for 1 min and the supernatant discarded. The remaining pellet was resuspended in 1 mL of YPD and left to incubate at 30 °C for 1 hour. Following the sample was centrifuged for 1 min, supernatant discarded and pellet washed in 1 mL YPD. Onto an agar plate, of appropriate selection media, 150 µL of the solution was added. This was evenly distributed across the plate by a glass rod, sterilized by an ethanol soak passed through the flame of a Bunsen burner. The remaining sample was then

concentrated by pelleting the sample by centrifugation for 1 min and resuspending the sample in 200 μ L of YPD. Using the same procedure as above, 150 μ L of the concentrated solution was streaked onto a second agar plate of the same selection media. Both plates were then incubated at 30 °C for 3-4 days to allow for transformant growth.

As the presence of both fluorescent proteins used in this study cannot be attained through positive selection markers (RFP does not have a selection marker attached), the presence of the fluorescent proteins were confirmed through confocal microscopy utilising the OPERA microscope. A sample of the transformant, diluted in the appropriate selection media, was viewed under the OPERA microscope to detect for the presence and correct sub-cellular localisation of both fluorescent proteins.

Transformant samples which met this requirement were restreaked on the appropriate selection media, allowed to grow, then selected and frozen down in 25% glycerol stocks.

2.2.3 LOPAC library HTS of dual fluorescent probe

Liquid handling of the LOPAC library was performed using a CyBi-Well 96-channel simultaneous pipettor (CyBio AG, Jena, Germany) in order to generate a 384-well library available for high throughput screening (HTS). To each plate in the library, 1 μ L of each compound (using 1 μ L pins) from the 96-well format was transferred into 49 μ L of the appropriate growth media in a 384-well clear bottom plate (PerkinElmer CellCarrier). As multiple strains were being treated during each HTS screen and these fluorescent strains have differing growth requirements both the LOPAC library and the strains being screened were prepared differently. The strictly GFP strains were screened in media containing LOPAC compound and 49 μ L of SC – His media whilst the dual fluorescent strain were screened in media containing LOPAC compound and 49 μ L of SC – (His, Ura).

The pins were subjected to subsequent methanol (Scharlau Chemie, S.A)/DMSO rinses and blotting on fresh filter paper between plates in order to keep the pins clean and stop contamination. The libraries 16 plates were arrayed four times into a corresponding 384 well plate. Thus, generating a set of 16 384-well plates, where each compound of the LOPAC library was represented in the corresponding plate in quadruplicate (4 μ L of each compound total).

2.2.3.1 Growth Conditions and Image preparation

Following the preparations described above, the single TIF1p-GFP and TIF2p-GFP strains were then manually added to each alternating well in the top row of the first 384-well clear bottom plate before the dual fluorescent constructs: TIF1p-GFP TIF1-RFP and TIF2p-GFP TIF1-RFP were added in the same pattern to the second row of the same plate. This pattern was then continued ensuring that all LOPAC drugs were being exposed to all 4 strains. An example of the plating pattern described above is illustrated in Figure 2.1.

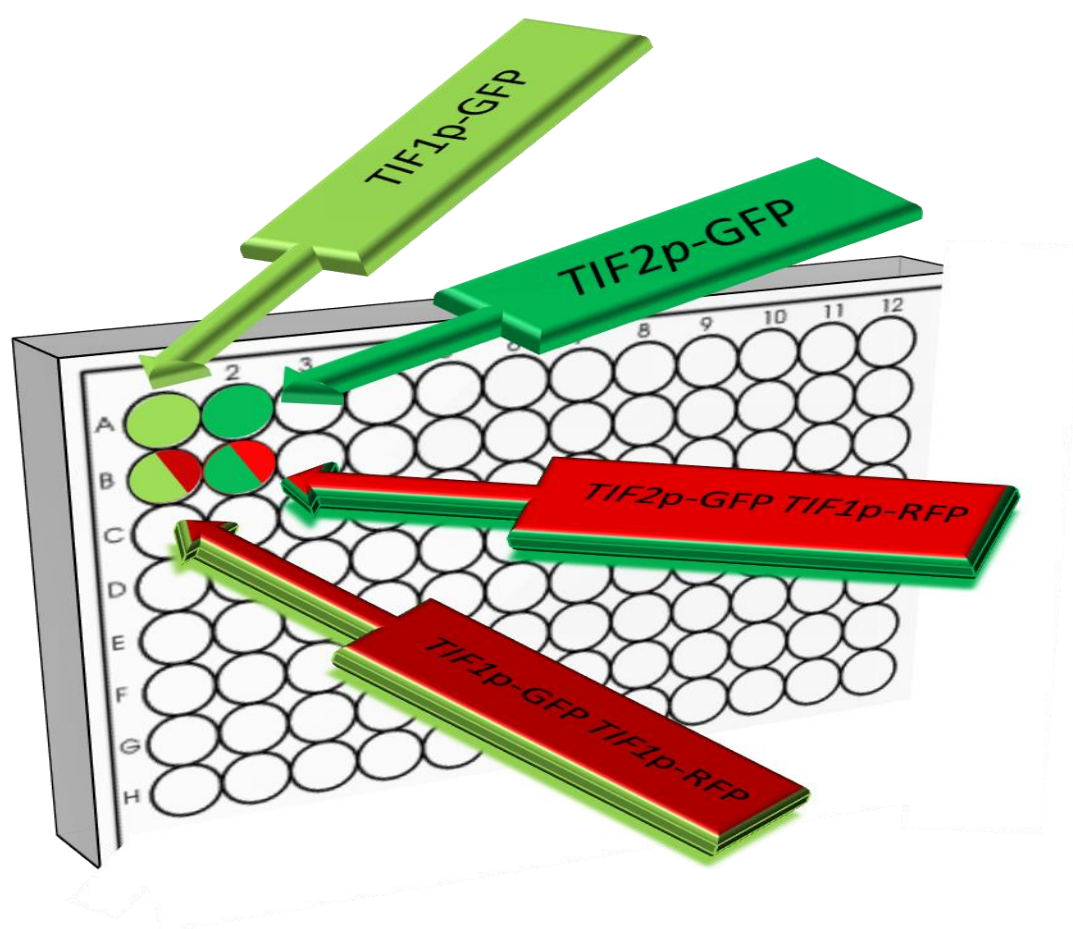


Figure 2.1: Illustration of LOPAC screen 384-well plate plan. Incorporation of four different fluorescent yeast strains with differing selection requirements into one screening plate. Starting at Row A, every other row contains SC - His media whilst starting at Row B every other row contains SC - (His, Ura) media. The pattern, continued throughout the entire plate, is colour coordinated to the fluorescent protein (or proteins) which were manual pinned in the corresponding well.

The plate was then left to settle for 10 min before image acquisition by confocal microscopy (see section 2.2.3.2). The plate was then shaken at 1000 rpm for 1 min and placed in a 30 °C incubator before a further image was acquired after 4 h. This process was then repeated for all remaining 384 clear bottom plates inoculated with the LOPAC library.

2.2.3.2 Image Acquisition

Each plate was imaged initially and after 4 h using OPERA, an automated spinning disc confocal microscope with a 60x water emersion lens. Using an exposure time of 200 ms, GFP was excited with a 488 nm laser. The emitted light was separated through a 568 nm primary dichroic mirror and filtered through 520/35 nm bandpass filter for imaging of GFP.

For the two dual fluorescent strains both GFP and RFP imaging was required. The GFP image was captured as above and the RFP was captured separately, directly after the GFP image was captured, in an attempt to limit the effects of fluorescent cross over between the channels. Using an exposure time of 200 ms, RFP was excited with a 561 nm laser. The emitted light was separated through a 568 nm primary dichroic mirror and filter through a 600/40 nm bandpass filter for imaging of RFP.

In either case for each field of view, three z-stacks were imaged each 1 µm apart covering different depths of the cell. Each plate took approximately 20 mins to image.

2.2.3.3 Image Analysis

Images were analysed using the image recognition software Acapella, which utilises fluorescent markers to identify and segment individual cells. The images generated by Acapella were analysed using a modified version of the methodologies established in our laboratory (Bircham, Maass et al. 2011).

The image recognition process developed by Bircham *et al.*, (2011) analyses yeast strains with a GFP fused to a gene of interest as well as a nuclear localisation signal (NLS) protein fused with the marker Redstar2 (NLS-RedStar2) and an mCherry cytoplasmic marker. The bright NLS marker was used to locate and identify the nucleus by thresholding and water shedding techniques before whole cells were identified in the same way by the fainter mCherry cytoplasmic marker. Reporter strains used in the current research however lack

these markers so cell objects were identified without the nuclei identification and whole cells were identified by cytoplasmic *TIF*-GFP.

In brief, the script adapted from Bircham *et al.*, (2011), employs cytoplasmic *TIF*-GFP to find all potential cells contained in each well and using thresholding and water shedding techniques defines accurate cell boundaries. Upon accurate cell boundary identification, of the three z-stacks imaged (see section 2.2.3.2), only the z-stack which represents the cell midsection were analysed. The GFP and or RFP intensity were then calculated based on the average pixel intensity within the cell objects boundaries. The fluorescence intensities were then reported for every single cell identified in the well not just an average of the entire wells' fluorescence intensity.

2.2.3.4 Statistical analysis of LOPAC screen

Sample whole cell fluorescence intensities were compared against that of control whole cell fluorescence intensities using the Wilcoxon rank sum test (Wilcoxon 1945) as implemented in the freely available software program R (R version 2.12.2 The R foundation for statistical computing). The p values generated by the Wilcoxon rank sum test were then subjected to the Bonferroni correction for multiple comparisons (Dunnett 1955), generating Q values, as implemented in the program R.

Each compound was then presented as a ratio of sample whole cell fluorescence intensities over control whole cell fluorescence intensities. The compounds were, using this ratio, ranked from greatest observed increase in GFP expression to greatest observed decrease in GFP expression. This list was then refined to only include compounds that had an associated Q value of <0.01 , such compounds were considered hits.

Where appropriate the GFP and RFP components for each strain were treated as though each were a completely different screen, hence a list of hits for both GFP and RFP components for the dual fluorescent strains were generated.

3 Results and Discussion

3.1 Creation of a *TIF1*p-RFP fusion protein

In order to dissect eIF4A's role as a part of the yeast translational machinery, the genes *TIF1* and *TIF2* were evaluated. As previously stated, the potential for genetic redundancy amongst duplicate genes hampers the efficiency of many standard research techniques, and subsequently to circumvent this issue a fluorescent based methodology was adopted. The ultimate goal of this research is to generate a set of fluorescent reporter strains integrated into the genome via the chromosome instead of being dependent on exogenous plasmids. However, the plasmid borne system was an achievable intermediate goal suitable for a Masters project. In addition this system, if successfully generated, still could be used to evaluate whether *TIF1* and *TIF2* can be individually regulated by small molecule intervention. Thus, construction of the reporter strains with a plasmid based TIF1p-RFP was pursued.

3.1.1 Primer design and module creation

To create the reporter strain TIF1p-RFP TIF2p-GFP a TIF1p-RFP fusion protein needed to be composed and united with the TIF2p-GFP strain from the commercially available Yeast GFP Clone Collection. To generate the TIF1p-RFP fusion, polymerase chain reaction (PCR) – mediated module construction was employed to produce analogues of the major constituents required for a TIF1p-RFP fusion, namely; a *TIF1* Open Reading Frame (ORF) with endogenous promoter (*TIF1pr*) and endogenous terminator region (*TIF1tr*), a source of RFP as well as the shuttle vector pRS316. Furthermore, these modules had to include additional flanking regions with significant overlap between modules allowing for their union through homologous recombination. These flanking regions were incorporated into each module through the use of primers with sticky ends to the DNA source each primer would amplify. The primer design and resulting module constructs are outlined in Figure 3.1 and Table 3.1 respectively.

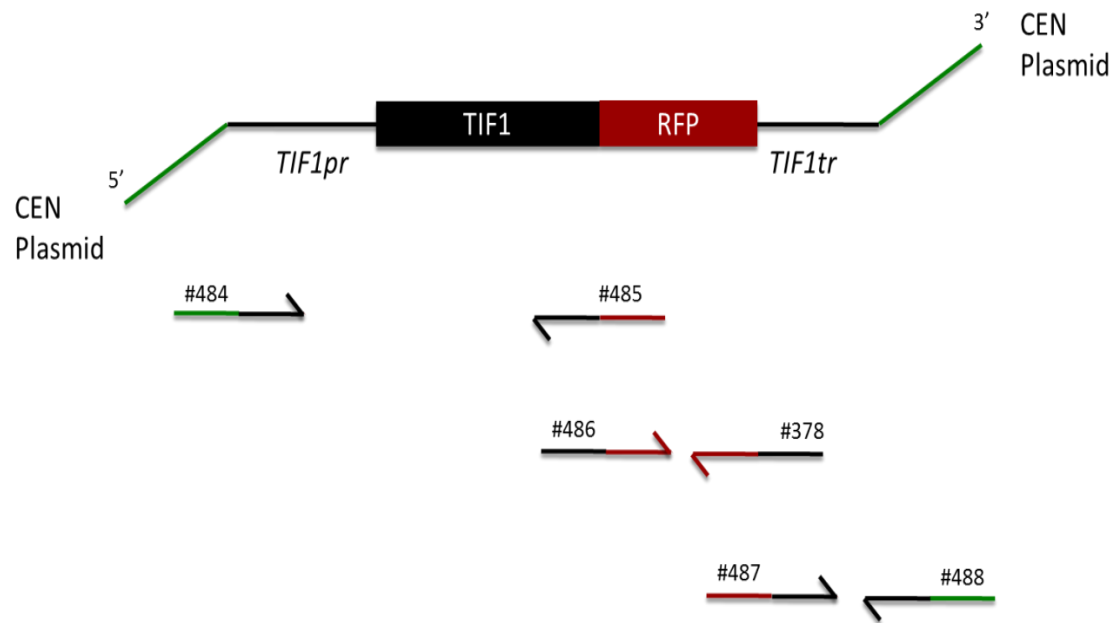


Figure 3.1: Primer design employed when creating TIF1-RFP via PCR mediated module construction. At the top is the full length construct produced whilst aligned below are the primers, with associated catalogue number, that were used. Flanking regions were incorporated in the modules through the use of primers with sticky ends to the source of DNA which the primer would amplify. The figure is colour coded to show each DNA source of the major constituents; *TIF1* ORF, *TIF1pr* and *TIF1tr* = black, CEN plasmid pRS316 = green, and RFP = red. The primers use this same colour code to show the sticky end parts of each primer.




Module Number	Construct	Forward Primer	Reverse Primer	Module Size (bp)
1		#484	#485	1493
2		#486	#378	791
3		#487	#488	392

Table 3.1: The construct design of the TIF1p-RFP modules as assembled through PCR mediated module construction. The size (in base pairs) of each module, and the primer sets required to generate them. In modules where a fraction of a component is present an arrow head is shown as opposed to a full block illustrating a complete component.

In order to limit the ambiguity with which the modules were allowed to form, specific strains and plasmids were utilised in their construction. In addition to the primer sets described, module 1 required a source of genomic *TIF1* for successful amplification. Therefore *tif2Δ::KANMX4*, *his3Δ1*, *leu2Δ0*, *met15Δ0* and *ura3Δ0* from the yeast knockout MATa strain collection was selected in an endeavour to limit any possibility of *TIF2* supplanting *TIF1* and fusing to the RFP in its place. As module 3 also required a source of genomic *TIF1* present to amplify the endogenous *TIF1* terminator (TIF1tr) the aforementioned strain was again utilised. RedStar2, the RFP source that module 2 utilised for amplification, was sourced from the plasmid pYM43.

Following PCR-mediated component module construction (see methods 2.2.1.10) each module was confirmed electrophoretically as shown in Figure 3.2. That bands corresponding to each module were determined by comparing the DNA markers and unknown bands with the expected module sizes, as reported in Table 3.1.

During the primer design phase of this research, two different reverse primers for module 3 were devised (See Table 2.2). The two were created with varying lengths of sequence homology to the *TIF1* terminator, an area that had potential to hinder the success of any subsequent PCR, promoting the success of a specific step in the creation of TIF1p-RFP. The longer “reverse primer A” endowed greater provision for the success of homologous recombination whilst the shorter “reverse primer B” was expected to be better equipped for the PCR phase of TIF1p-RFP construction. However, since both primers resulted in successful amplification of module 3, the module constructed with reverse primer A was applied in the subsequent steps of TIF1p-RFP construction.

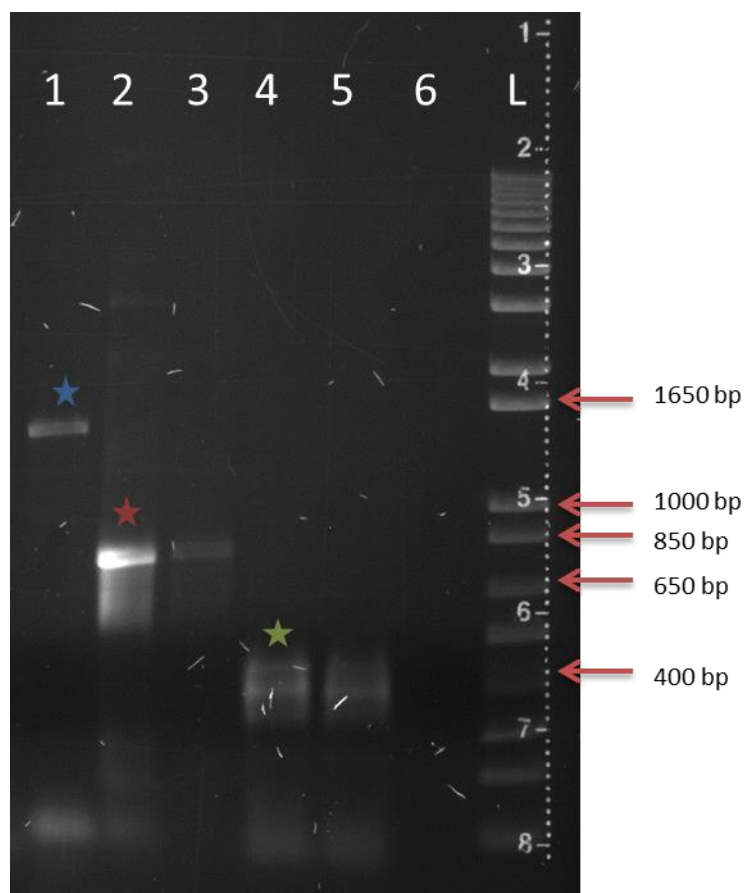


Figure 3.2: PCR mediated component module construction. Lanes: 1 – Module #1, 2- Module #2, 3- 1:10 dilution of Module #2, 4- Module #3 using primer #488 as reverse primer, 5- Module #3 using primer #489 as reverse primer, 6- Negative PCR control, L - 1Kb plus DNA ladder. The coloured stars represent the versions of each module which were utilised in subsequent experiments.

3.1.2 Introduction of modules into a functional host

In order for yeast strains to integrate and express the desired TIF1p-RFP, these modules need to be integrated into a suitable vector which will be expressed inside the parental yeast strain alongside normal gene expression. As previously stated, the use of sticky ended primers created regions of sequence overlap between the three modules granting the potential for homologous recombination to link the module into one linear segment. In addition this methodology facilitated the incorporation of the modules into the CEN plasmid pRS316 (Figure 3.3).

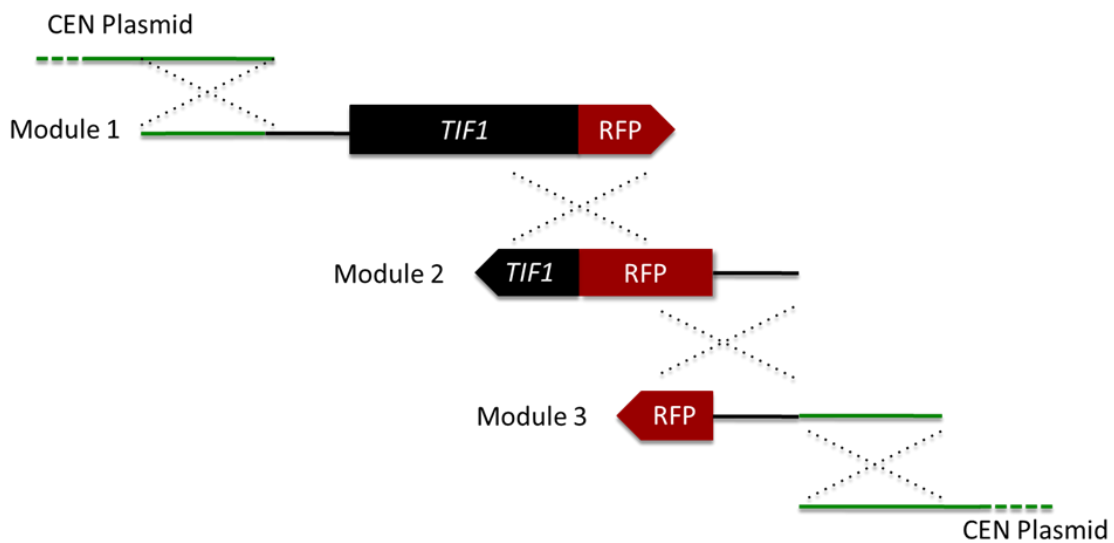


Figure 3.3: Sites of sequence homology between modules and plasmid. These sites allow the facilitation of homologous recombination between each component of the *TIF1*-RFP construct. In modules where a fraction of a component is present an arrow head is observed as opposed to a full block illustrating a complete component. Colour key – green = Cen plasmid, black = *TIF1* (ORF, *TIF1pr* and *TIF1tr*), red = RFP, black dash = sequence homology allowing homologous recombination.

The incorporation of the *TIF1*-RFP construct into a circular plasmid would not naturally occur hence the plasmid must first be made linear. As pRS316 is selected for under the same selection pressure whether or not the *TIF1*-RFP construct is successfully incorporated in the plasmid it is necessary that the religation of a construct free pRS316 be extremely unfavourable. Therefore, a 42 bp cut in the multiple cloning site of pRS316

was generated using the restriction enzymes XbaI and HindIII. This allowed for the incorporation of each module in pRS316 rather than the religation of a construct free pRS316.

3.2 Construction of a dual fluorescent reporter strain

To facilitate the homologous recombination and introduction of the *TIF1*-RFP modules into a functional strain of yeast, the transformation procedure described in section 2.2.2.1 was performed (Figure 3.4). This process could have been achieved by many different methods, however transformation was preferred. The parental yeast strain, under standard transformation conditions, is able to facilitate homologous recombination and replication of the plasmid, as well driving transcription and translation of both plasmid borne and chromosomal gene copies. Thus no other external steps are required to create a functional fluorescent protein that will be expressed under normal cellular conditions.

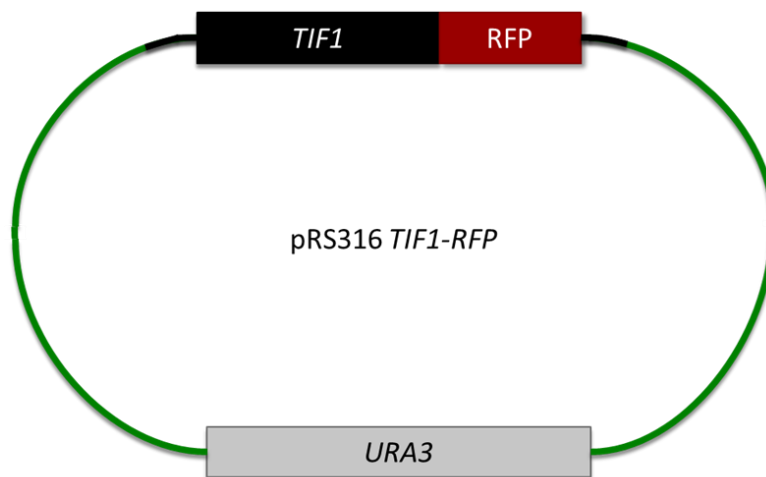


Figure 3.4: Plasmid borne *TIF1*-RFP. The CEN plasmid pRS316 with the *TIF1*-RFP construct inserted via transformation mediated homologous recombination. In order to make the *TIF1*-RFP respond in a manner as close to unlabelled *TIF1* as possible, effort was made to incorporate both the endogenous promoter and terminator regions of *TIF1* at the 5' and 3' ends of the construct respectively. The selection marker for plasmid pRS316, *URA3* is also shown.

In order to attain a one to one ratio between the number of copies of *TIF1*-RFP that were expressed in each cell, and hence a way of measuring cellular levels of *TIF1*, the shuttle vector pRS316 was selected in this study. The main advantage of shuttle vectors are their ability to propagate in two different host species for example, pRS316 contains components that allow for the replication and selection in both *S. cerevisiae* and *E.coli*. This shuttle vector is one of a series of CEN-based plasmid (CEN plasmid) which are characterised by the presence of a centromere sequence as well as a normal yeast origin of replication. The CEN sequence is recognised by the host strains which are able to replicate the plasmid as though it was a small chromosome at a rate of one copy per cell. Of this plasmid set, pRS316 was utilised as it has a strong selection marker, *URA3*, which importantly is not being utilised as a selection marker for any other method employed during this research.

Two parental strains were utilised during the transformation procedure to produce differing fluorescent reporter genes of interest. The most pertinent reporter gene to this study was TIF1p-RFP TIF2p-GFP which during this transformation step was created in the parental strain *TIF2*-GFP from the Yeast GFP clone collection (section 2.1.3).

It is also important to note that satisfactory transmission of plasmid between parental and daughter generations was observed when allowed to grow on fresh media for five generations. However, the plasmid fluorescence intensity was more variable than the chromosomally expressed GFP. Confocal microscopy Images of the four strains utilised during LOPAC library screening are shown in Figure 3.5.

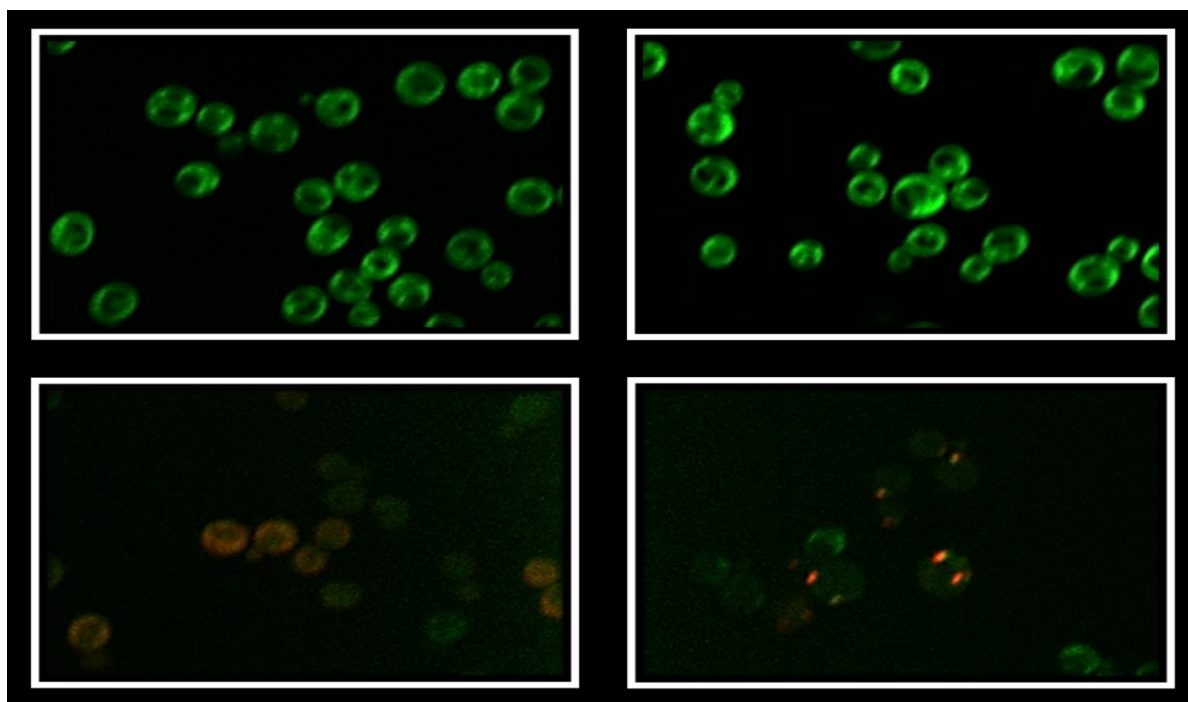


Figure 3.5: Confocal microscopy images of the 4 strains utilised in the HTS of the LOPAC library. Top Left – *TIF1p-GFP*, top right – *TIF2p-GFP*, bottom left – *TIF1p-RFP TIF1p-GFP TIF2*, bottom right – *TIF1p-RFP TIF1 TIF2p-GFP*. Images were converted from 16 bit to 8 bit copies and therefore the brightness/expression levels cannot be used as direct comparisons between the four strains.

Although these images cannot be directly used to correlate the brightness of each fluorescence protein to the gene expression levels in each strain (see caption above), it is important to note that the differences illustrated in these images reflect the underlying optical data and are not an artefact arising from either the image acquisition or image processing phases.

Figure 3.5 clearly illustrates that the RFP signal from *TIF1-RFP* is weaker than that observed for the GFP equivalent. This is found consistently and does not arise from inherent fluorescence efficiencies of the red and green fluorescent proteins. Literature records that S65T GFP has an extinction coefficient of $65,000 \text{ M}^{-1}\text{cm}^{-1}$ and a quantum yield of 64% (Patterson, Knobel et al. 1997). RedStar2 has been shown to have an intensity 2-4 times greater than that of DsRed, coupled with a lower tendency to aggregate (Janke, Magiera et al. 2004). DsRed has an extinction coefficient of

approximately $75,000 \text{ M}^{-1}\text{cm}^{-1}$ and a quantum yield of approximately 0.75 (Heikal, Hess et al. 2000). This shows that the RFP RedStar2 should be at least as bright as the S65T GFP.

In addition to the overall brightness levels of the two fluorescent proteins, these images highlight three trends observed throughout the dataset. Firstly that, in the case of GFP, the GFP expressed is more intense before the introduction of the second fluorescence source. Secondly, the chromosomally expressed GFP is more intense than their plasmid RFP counterparts. Finally, that the RFP present in both the dual fluorescent strains, even though fused to the same protein, is being expressed in two unsimilar ways. The RFP signal generated from TIF1p-RFP TIF1p-GFP is evenly expressed in the cytoplasm, whilst the RFP signal generated from TIF1p-RFP TIF2p-GFP, is expressed in a more punctate pattern throughout the cytoplasm.

A possible explanation for the first and second trends may arise from the approach with which the two dual fluorescent strains were generated. At the beginning of this research there was no established precedent that suggested it was possible to individually identify both *TIF1* and *TIF2* in the single environment. Therefore it was expedient in this initial study to work with the untransformed commercial *TIF2*-GFP strain. Consequently a third source of the eIF4A protein was retained in both the dual fluorescent strains resulting in the genotypes *TIF1*-RFP *TIF1*-GFP *TIF2* and *TIF1*-RFP *TIF1* *TIF2*-GFP (herein these strains will be referred to without the third *TIF* gene unless required). As such, the low apparent signal of *TIF1*-RFP as well as the higher signal intensity of the GFP in the single fluorescent strains as opposed to the dual strains maybe, to some extent, a consequence of the cellular demand for eIF4A being supplied by a greater number of genes that can only compete for a limited number of transcription factors.

With regard specifically to the low apparent signal of *TIF1*-RFP, the second highlighted trend, the explanation may also lie in the UTR and location of the gene. It is plausible that some transcription factor associations are lost or otherwise affected in a way unique to a plasmid based expression system. Such potential differences between the genomic architecture of chromosomal and plasmid borne expression systems are a largely unknown quantity at play in this intermediary CEN plasmid construct.

In regard to the third trend, it is possible that the punctate RFP pattern, observed only in TIF1p-RFP TIF2p-GFP, was caused by a number of factors unique to the genetic interaction of the plasmid borne *TIF1*-RFP with the parental host strain. For example, it might be that there is a previously unrecognised mRNA localisation motif in the region of the 3' UTR that has not been captured in this CEN plasmid construct. This could potentially cause misslocalisation of *TIF1* and a loss of function. However, normal function would be retained regardless as a consequence of the third unlabelled source of eIF4A. Alternatively, TIF1p-RFP may be more accurately reflecting the impact that the 3' UTR is having on site directed localisation than TIF1p-GFP, since the commercially available strains have their 3' UTR disrupted by various selection markers.

It is also possible this pattern is, to some extent, caused by the combination of three eIF4A genes contributing to cellular demand. The combination of three genes increases the maximum potential eIF4A that can be synthesised which may lead to an increased total abundance of eIF4A. The result of this rise in total abundance may force an increase in the aggregation of either eIF4A or the RFP. Additionally the punctate pattern may be caused by differences in the nature of the third eIF4A source in each of the dual fluorescent reporter strains. As the RFP is located on *TIF1*, the retention of a second unlabelled copy of *TIF1*, as is found in TIF1p-RFP TIF1 TIF2p-GFP, could result in either the unlabelled *TIF1* being produced in favour of *TIF1*-RFP or an overabundance of *TIF1* in general.

If the punctuate pattern observed were a consequence of any of these suggestions then performing a dye swap experiment, where the fluorescent proteins assigned to each gene of the dual fluorescent reporters were swapped, would illustrate this by causing the RFP expression patterns observed for each dual fluorescent reporter to alternate.

Despite the aforementioned limitations it was anticipated that the *TIF1*-RFP would still effectively report on treatments that lead to changes in *TIF1* expression. Thus, these strains were still utilised to a) identify whether the abundance differences of Tif1p and Tif2p can be detected in this system and b) to evaluate the possible screening methodologies for future use. An overview of the methodologies utilised to create the TIF1p-RFP TIF2p-GFP reporter are shown in Figure 3.6.

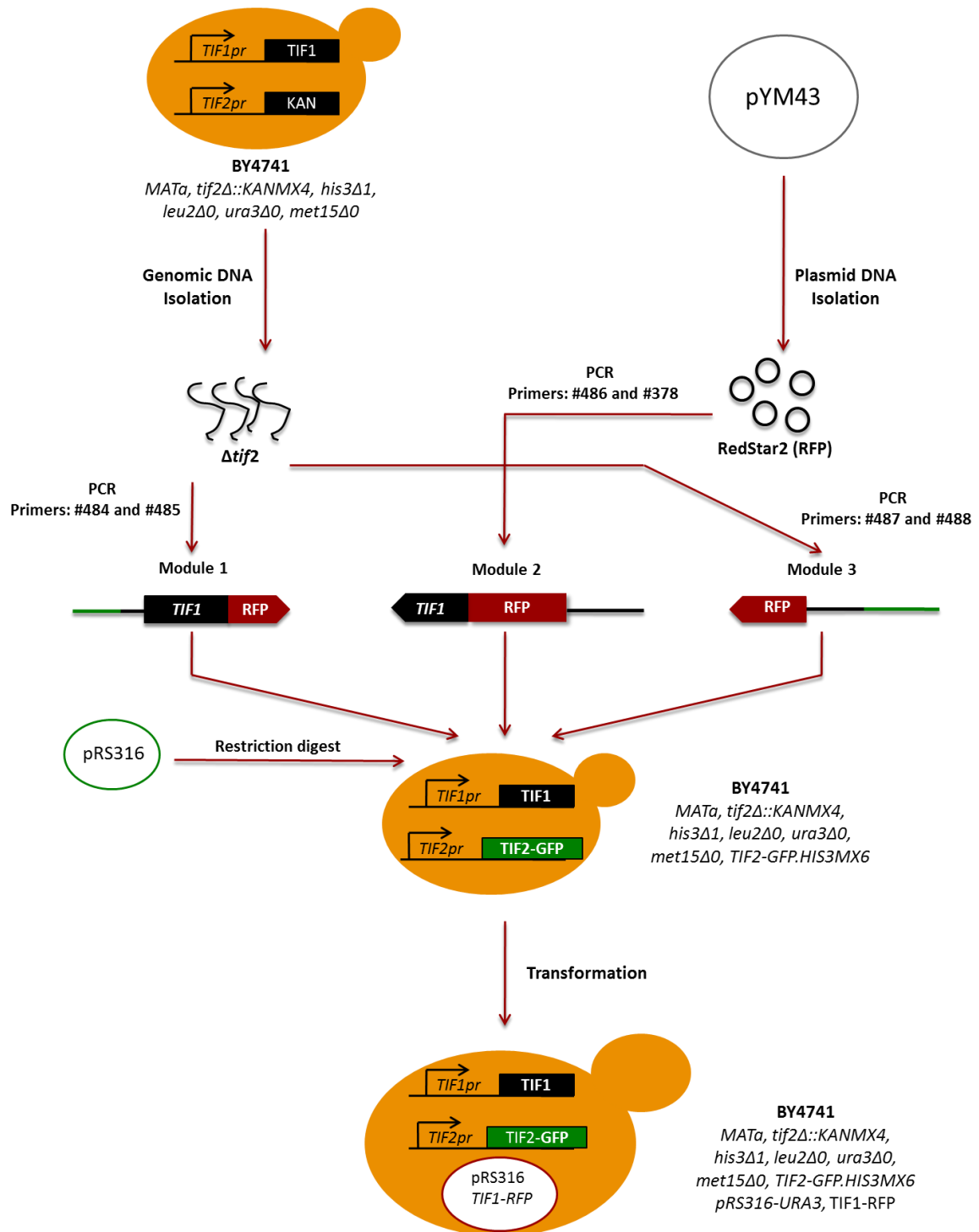


Figure 3.6: Overview of TIF1p-RFP TIF2p-GFP strain generation. All components of the three TIF1-RFP modules were individually isolated by either genomic or plasmid DNA isolation. Components were then crafted into three *TIF1*-RFP modules via PCR mediated module construction utilising the noted primer scheme. The transformation process produced functionally active *TIF1*-RFP integrated, inside processed pRS316, in the parental host strain bearing the commercially available *TIF2*-GFP.

3.2.1 Application of dual fluorescent reporter strain

A major advantage of the dual fluorescent reporter strain is its ability to be applied to a multitude of different methodologies. Throughout this study this strain had been employed in applications that utilise their fluorescent properties, however, as this reporter strain is essentially a collection of two epitope tags its use is not limited to applications with fluorescent based systems.

The inspiration for the design of this study was driven by the work outlined in Deluna *et al.*, (2010) who used fluorescent markers to investigate the regulation of protein levels, by duplicate genes in the *S. cerevisiae* genome, in response to the deletion of one their paralogous genes. In brief, the Deluna system utilised a pair of haploid yeast strains that were constructed with a duplicate gene fused to GFP in either the wild type background or in a background deleted for its paralog. In order to individually identify each strain they constitutively expressed either CFP or RFP respectively. Deluna was then able to coculture both strains, dose with the drug/drugs of choice and using three colour flow cytometry was able to measure the protein regulation by each duplicate gene in both wild type and deletion backgrounds (DeLuna, Springer et al. 2010). However, the *TIF* genes were not investigated during the aforementioned study to avoid complications of ribosomal protein genes and aneuploidy.

As a consequence, this study was originally focused on creating a dual fluorescent strain that could be examined by flow cytometry to evaluate whether *TIF1* and *TIF2* can be individually regulated by small molecule intervention. However despite months of optimisation, it became increasingly clear that the HTS approach employed in this study was not conducive to the detection of small differences in protein regulation. Under the HTS conditions used, cells were growing in log phase. Under these circumstances, unrestricted cell cycle progression creates a high demand for protein synthesis, and it is likely that eIF4A usage was near maximal. For that reason, compounds that would elicit and increase in eIF4A would have been unlikely to create notable change in fluorescence, although suppressors of eIF4A should still have been apparent. Compounding this, the HTS system seemed to suffer from large noise in the observed fluorescence intensities, and this noise would mask small changes in the signal resulting from compound treatment.

Research, utilising the growth conditions outlined in section 2.2.3.1, investigating the effects of 35 μ M atorvastatin on *PDR5*-GFP was conducted in tandem with this study. Atorvastatin, at this concentration, is known to up regulate *PDR5*, however this up regulation could not be detected via flow cytometry but was detectable by confocal microscopy (Ploi Yibmantasiri, Victoria University of Wellington, personal communication). As a goal of this research was to conduct high throughput methodologies the remaining experimentation was conducted using confocal microscopy in favour of flow cytometry.

3.3 Screening of the LOPAC¹²⁸⁰ library

Using the functional reporter strain TIF1p-RFP TIF2p-GFP, with distinguishable *TIF1* and *TIF2* markers, differentiation of the respective gene functions of *TIF1* and *TIF2* can be elucidated under a variety of cellular conditions. To encompass as many different variables as possible this strain was subjected to a morphological high throughput screen (HTS) utilising the LOPAC¹²⁸⁰ Library of Pharmacologically-Active Compounds. This library was chosen as a library of environmental variables rather than as a library of bio-actives of interest.

As indicated in section 2.2.3.2, readings from the LOPAC library HTS screens were taken at time zero and after a four hour incubation period. This methodology was exercised in an attempt to enrich the list of “hits” to only include compounds eliciting a response resulting from gene regulation. As the initial time point for imaging is almost instantaneously after the introduction of the LOPAC library it is not conceivable that any observed changes in protein expression were regulated via changes in mRNA expression. Instead, it is plausible that any such changes are instigated by disturbances in the GFP signal or arise from intrinsic fluorescence in the compound library and not as a consequence of gene expression changes in response to the introduction of the LOPAC library. Hence any “hits” observed at the initial time point are considered a crucial control in eliminating those compounds that are not pertinent to the primary aim of this study; elucidating differential regulation between *TIF1* and *TIF2*.

Two compounds were consistently identified within the top twenty compounds that lead to an apparent GFP increase at both time points of every screen; SB 216763 and SU 5416. Rapid increases in fluorescence intensity, although not always to a level of statistical significance, were observed for SB 216763 and SU 5416 in all the screens which incorporated GFP. Furthermore the fluorescence intensity was at its highest, approximately 100 times that of control (see Figure 5.3 and Figure 5.6) from data sets procured at time point zero. As previously discussed, such rapid increases are unlikely to be a result of gene regulation but rather a result of intrinsic compound fluorescence or compound promoting fluorescence from some other mechanism. A brief literature search highlighted that SB 216763 was itself fluorescent and subsequently omitted in the analysis of previously conducted work with fluorescent proteins (Bayliss, Bellavance et al.

2006). Contrastingly SU 5416, with a similarly sizeable increase in GFP expression levels, was not identified as intrinsically fluorescent (Nuutinen, Ropponen et al. 2009). However, the mechanism by which SU 5416 increases the fluorescence intensity will not be further pursued in this study.

The results of the 4 hour readings are presented and discussed below whilst the time zero readings are presented in Appendix A and discussed where necessary.

3.3.1 Rational of statistical analysis

The Wilcoxon rank sum test and Bonferroni correction for multiple comparisons are standard statistical analysis for non-parametric data sets such as those generated by the LOPAC library screening. However, it was expedient in this initial study to identify “hits” at a more stringent threshold than normally required so as to assess the system’s capabilities for identify compounds that caused differential regulation between the paralogous *TIF* genes. The threshold employed had a cut-off point of Q value <0.01, in other words the threshold filtered hits to maintain a false discovery rate <1 %.

In addition to a low false discovery rate it was also pertinent to filter results on the strength of their associated ratio of sample whole cell fluorescence intensities over control whole cell fluorescence intensities. This streamlined the focus of this study to concentrate on the success of the fluorescent based system to identify substantial regulation differences between *TIF1* and *TIF2* whilst limiting the effect of experimentally introduced noise. This was achieved by assigning an arbitrary threshold on the aforementioned ratio of above 1.2 and below 0.8.

For each of the screens, graphical representations for the following are presented; A) the ratio of sample whole cell fluorescence intensities over control whole cell fluorescence intensities for the entire library, B) The top twenty compounds (if possible), with Q values <0.01, that elicited an increase in fluorescence intensity C) The top twenty compounds (if possible), with Q values <0.01, that elicited a decrease in fluorescence intensity.

The compounds highlighted in B) and C) were tabulated and presented with their associated ratio of sample whole cell fluorescence intensities over control whole cell fluorescence intensities and levels of statistical significance.

3.3.2 Single GFP screening – Reading at 4 Hours

The central dogma of gene expression proposes that a single gene encodes for a single protein, however, duplicate gene pairs are able to encode for the same protein. Thus, from an experimental vantage point, it is possible to generate an impeded perception of a protein if not all of the protein manifestations are considered. Consequently, compounds which instigate differential regulation of the duplicate genes *TIF1* and *TIF2* may draw varying conclusions if the genes were examined individually as opposed to co-examination. In contrast to the single GFP reporters, the dual fluorescent reporter is able to evaluate all manifestations of eIF4A in a single environment whilst acknowledging the role that both genes perform.

Conversely the single GFP reporter screens are not redundant as they are used to validate both the dual fluorescent reporter as well as the choice of methodologies employed in deciphering the role of eIF4A as a part of yeast translation machinery. Screening these strains in conjunction with the dual fluorescent reporter allows suitable comparisons for the evaluation of the impact that genetic redundancy may have on *TIF1* and *TIF2*.

It is important to note that the set of singular GFP control screens applied to validate these methodologies, although necessary and appropriate in the ideal genomically integrated dual reporter system, are not a perfect validation of this system. As previously stated, an extra source of eIF4A was retained in the dual fluorescent reporter in the form of an unlabelled copy of *TIF1* (see section 3.2). Subsequently not all manifestations of eIF4A are being monitored in a single environment. In addition the GFP set used as validation is integrated into the chromosome whilst the *TIF1*-RFP is a plasmid borne copy. Nevertheless although changes in *TIF1* expression might be distributed between the two copies of *TIF1* in these reporter strains it was anticipated that the *TIF1*-RFP would still effectively report on treatments that lead to changes in *TIF1* expression.

Although the primary objective of this project was to create a dual fluorescent reporter strain, the individual *TIF1*p-GFP and *TIF2*p-GFP strains were also screened to provide a point of reference. The commercially available Yeast GFP Clone Collection allowed high throughput screening of the LOPAC library against these singular GFP strains. The ensuing

screens are outlined in Figure 3.7/ Table 3.2 and Figure 3.8/Table 3.3 respectively and discussed below.

3.3.2.1 TIF1p-GFP Screen

Of the compounds screened against TIF1p-GFP 154 compounds, as ascertained by the Wilcoxon rank sum test ($\alpha = 0.01$), elicited a significant increase in GFP expression when compared with control GFP. However, 83 compounds were unearthed in the same screen at the initial time point. Conversely, 155 compounds elicited a significant decrease in GFP expression when compared with control GFP with 48 compounds identified at the initial time point. Using an arbitrary established threshold of compound-treated GFP: control GFP ratio to constitute positive hits; above 1.2 and below 0.8, 3 compounds increased GFP expression whilst 19 decreased GFP expression. Although these thresholds are arbitrarily chosen, they reflect the aims of this study in identifying notable differentiation between the *TIF* genes.

Three compounds; Myricetin, Mainserin hydrochloride and Orphenadrine hydrochloride were identified as decreasing GFP expression below a score of 0.8 in both the zero and four hour time points of this TIF1p-GFP screen. As already discussed any such changes are potentially instigated by disturbances in the GFP signal and not as a consequence of the introduction of the LOPAC library. It is possible that such instances of GFP signal interruption are consequences of compounds which instigate fluorescence masking or miss folding of the GFP protein or some other mechanism which modulates GFP directly. In the case of Myricetin, which has a highly conjugated structure, it seems plausible that this modulation be attributed to quenching of the GFP signal owing to a high extinction coefficient. A high extinction coefficient could lead to the absorption of light at one or both of the GFP excitation or emission wavelengths.

Although identified as statistically significant by the Wilcoxon rank sum test ($\alpha = 0.01$), these outliers are removed from detailed consideration here for reasons outlined in section 3.3. However, in future studies, they will be deserving of consideration.

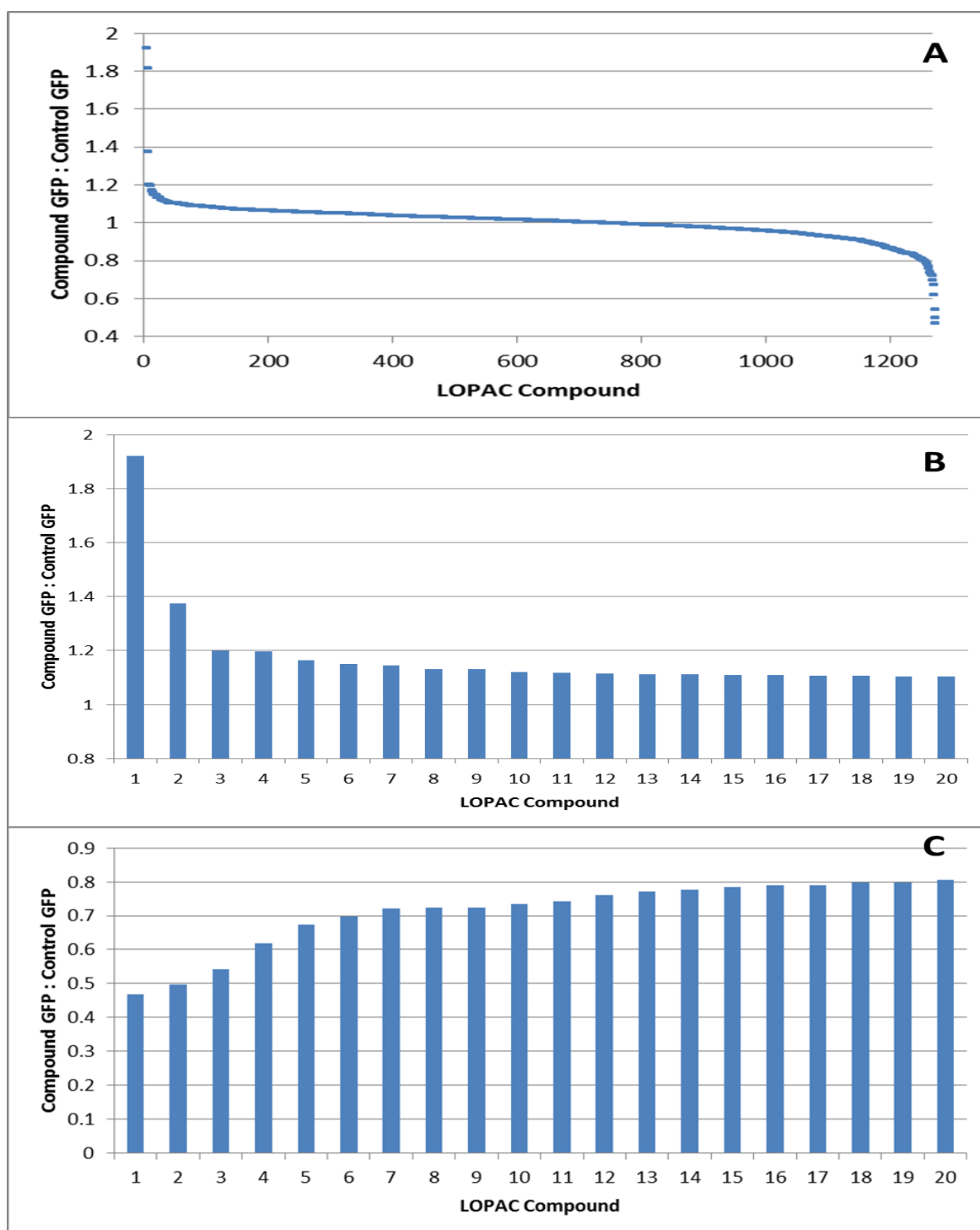


Figure 3.7: Whole cell fluorescence intensities of TIF1p-GFP treated with the LOPAC library (compounds 1 – 1280) compared against the pooled collection of whole cell fluorescence intensities of TIF1p-GFP treated with DMSO (as described in section 2.2.3.4). (A) Compound-treated GFP:control GFP for the entire LOPAC library ranging from highest to lowest. The top 20 compounds that (B) elicited an increase in or (C) elicited a decrease in the ratio of compound-treated GFP compared to control. Compounds highlighted in B and C are presented in Table 3.2.

#	Compound	Compound GFP : Control GFP	Q value		Compound	Compound GFP : Control GFP	Q value
1	SU 5416 *	1.922	6.47E-31		Myricetin *	0.469	5.315E-39
2	4-Amino-1,8-naphthalimide	1.374	6.34E-30		Mianserin hydrochloride *	0.4961	1.598E-31
3	Alloxazine	1.201	0.008513		8-Cyclopentyl-1,3-dimethylxanthine	0.5426	1.834E-44
4	Emodin	1.197	1.06E-07		DNQX	0.6196	9.033E-90
5	S-Methyl-L-thiocitrulline acetate	1.163	5.85E-05		Muscimol hydrobromide	0.673	6.271E-80
6	GW9662	1.149	1.76E-05		Orphenadrine hydrochloride *	0.6982	1.663E-60
7	(±)-Nipecotic acid	1.145	0.005389		Histamine, R(-)-alpha-methyl-, dihydrochloride	0.72	4.963E-79
8	1,1-Dimethyl-4-phenyl-piperazinium iodide	1.133	4.09E-05		E-64	0.7235	1.345E-13
9	Flumazenil	1.132	4.36E-21		Etodolac	0.7236	2.335E-17
10	S-Nitrosoglutathione	1.119	6.14E-10		Phenoxybenzamine hydrochloride	0.7352	8.048E-05
11	(±)-2-Amino-7-phosphonoheptanoic acid	1.119	4.56E-06		Dihydro-beta-erythroidine hydrobromide	0.7413	2.183E-17
12	Quazinine	1.116	7.08E-15		1-(m-Chlorophenyl)-biguanide hydrochloride	0.7604	4.252E-43
13	(±)-Chlorpheniramine maleate	1.112	0.000212		R(-)-2,10,11-Trihydroxy-N-propylnoraporphine hydrobromide	0.7703	1.451E-09
14	Chelerythrine chloride	1.111	2.83E-09		Doxazosin mesylate	0.7769	5.588E-10
15	R(+)-3PPP hydrochloride	1.111	1.8E-12		(+)-Butaclamol hydrochloride	0.7843	7.77E-09
16	AA-861	1.11	1.17E-05		T-1032	0.7898	3.692E-38
17	erythro-9-(2-Hydroxy-3-nonyl)adenine hydrochloride	1.107	1.31E-15		(±)-threo-1-Phenyl-2-decanoylamino-3-morpholino-1-propanol hydrochloride	0.7908	6.033E-49
18	NS 2028	1.105	4.2E-09		Phosphonoacetic acid	0.7972	4.414E-50
19	Cefmetazole sodium	1.105	3.41E-09		Dipropyl dopamine hydrobromide	0.7978	4.345E-08
20	Ellipticine	1.104	0.000367		Emetine dihydrochloride hydrate	0.8047	5.004E-08

Table 3.2: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF1p-GFP screen (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in sections 2.2.3.4 and 3.3.1). Compounds marked with an asterisk denote those compounds which appear in both t=0 and t=4 hours.

3.3.2.2 TIF2p-GFP Screen

Of the compounds screened against TIF2p-GFP 147 compounds, as ascertained by the Wilcoxon rank sum test ($\alpha = 0.01$), elicited a significant increase in GFP expression when compared with control GFP. However, 27 compounds were unearthed in the same screen at the initial time point. Conversely, 125 compounds elicited a significant decrease in GFP expression when compared with control GFP with 23 compounds identified at the initial time point.

Using an arbitrarily chosen threshold of compound-treated GFP: control GFP ratio to constitute positive hits; above 1.2 and below 0.8, 4 compounds increased GFP expression whilst 36 decreased GFP expression. Of these 40 compounds only Ro 16-6491 hydrochloride, shown to down regulate fluorescence intensity, was identified at both time points. Although identified as statistically significant by the Wilcoxon rank sum test ($\alpha = 0.01$), Ro 16-6491 hydrochloride was removed from detailed consideration here for the reasons outlined in section 3.3. However, in future studies, Ro 16-6491 hydrochloride will be deserving of consideration.

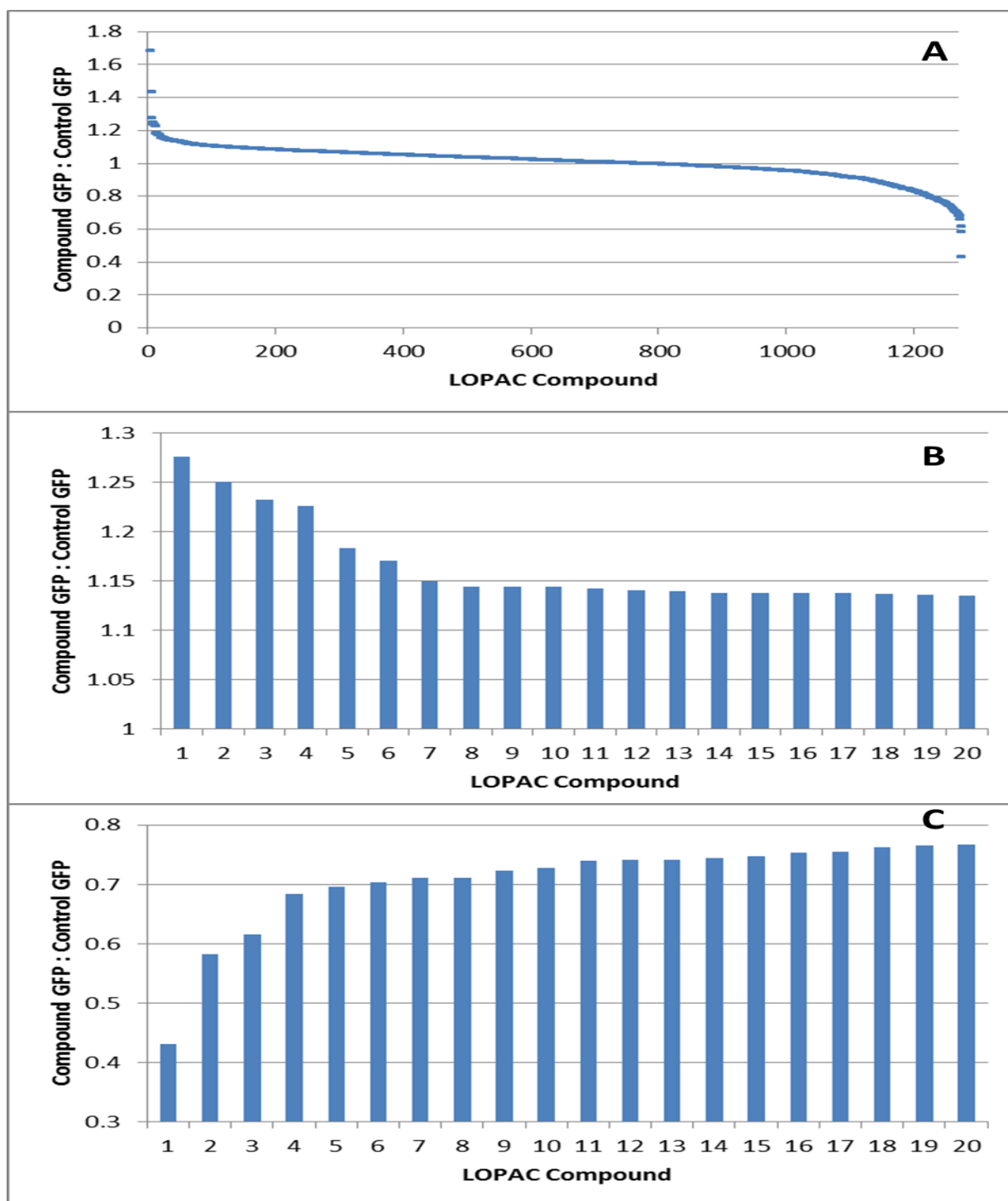


Figure 3.8: Whole cell fluorescence intensities of TIF2p-GFP treated with the LOPAC library (compounds 1 – 1280) compared against the pooled collection of whole cell fluorescence intensities of TIF2p-GFP treated with DMSO (as described in section 2.2.3.4). (A) Compound-treated GFP:control GFP for the entire LOPAC library ranging from highest to lowest. The top 20 compounds that (B) elicited an increase in or (C) elicited a decrease in the ratio of compound-treated GFP compared to control. Compounds highlighted in B and C are presented in Table 3.3.

#	Compound	Compound GFP : Control GFP	Q value		Compound	Compound GFP : Control GFP	Q value
1	AIDA	1.275	0.000193		8-Cyclopentyl-1,3-dimethylxanthine	0.4309	7.473E-27
2	(-)-Amethopterin	1.25	5.08E-06		N6-Benzyl-5'-N-ethylcarboxamidoadenosine	0.5824	0.00448
3	Enoximone	1.232	9.3E-39		DD1	0.6153	0.0002132
4	1-benzoyl-5-methoxy-2-methylindole-3-acetic acid	1.226	0.002597		5,7-Dichlorokynurenic acid	0.6839	1.865E-09
5	Se-(methyl)selenocysteine hydrochloride	1.183	2.05E-05		Dipyridamole	0.6958	2.191E-09
6	Adenosine amine congener	1.17	0.00647		Muscimol hydrobromide	0.7033	7.403E-20
7	(±)-Pindobind	1.149	2.78E-13		DBO-83	0.7111	3.882E-08
8	alpha-Lobeline hydrochloride	1.144	1.74E-08		DNQX	0.7111	1.504E-19
9	Chelerythrine chloride	1.144	9.02E-08		Ro 16-6491 hydrochloride *	0.7237	3.181E-39
10	2-Phenylaminoadenosine	1.144	5.61E-11		6-Hydroxy-DL-DOPA	0.7282	2.029E-13
11	Quinelorane dihydrochloride	1.142	3.42E-12		Dihydro-beta-erythroidine hydrobromide	0.7404	5.575E-10
12	Isonipetric acid	1.141	5.88E-07		N,N-Dihexyl-2-(4-fluorophenyl)indole-3-acetamide	0.7412	2.334E-05
13	5-fluoro-5'-deoxyuridine	1.139	2.02E-11		R(-)-2,10,11-Trihydroxy-N-propylnoraporphine hydrobromide	0.7419	1.246E-09
14	Flumazenil	1.138	1.31E-13		Lidocaine N-ethyl bromide quaternary salt	0.7449	5.537E-31
15	Famotidine	1.138	4.22E-11		L-3,4-Dihydroxyphenylalanine	0.7478	3.461E-08
16	Fenoterol hydrobromide	1.137	3.3E-08		Etodolac	0.7538	7.702E-06
17	Pyrimilamine maleate	1.137	4.62E-07		1-Aminocyclopropanecarboxylic acid hydrochloride	0.7555	5.385E-22
18	Dopamine hydrochloride	1.137	2.38E-07		1-(m-Chlorophenyl)-biguanide hydrochloride	0.7618	1.375E-12
19	PPADS	1.136	8.56E-08		SCH-28080	0.7649	1.315E-08
20	Putrescine dihydrochloride	1.135	2.14E-11		Doxazosin mesylate	0.7668	7.075E-06

Table 3.3: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF2p-GFP screen (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in sections 2.2.3.4 and 3.3.1). Compounds marked with an asterisk denote those compounds which appear in both t=0 and t=4 hours.

3.3.2.3 Comparison of TIF1p-GFP and TIF2p-GFP screens

As ascertained by the Wilcoxon rank sum test ($\alpha = 0.01$) eight compounds were identified to give fluorescence intensity outside of the arbitrarily chosen threshold of compound-treated GFP: control GFP; above 1.2 and below 0.8, for both the *TIF1* and *TIF2* screens. All eight compounds were identified below the 0.8 threshold for both *TIF1* and *TIF2* (see Table 3.4): no compounds were found to be above the threshold of 1.2 in both screens and no compounds were identified as causing an up regulation in one screen and down regulation in the other.

#	Label	GFP/Ave	Qvalue		Label	GFP/Ave	Qvalue
1	8-Cyclopentyl-1,3-dimethylxanthine	0.5426	1.83E-44		8-Cyclopentyl-1,3-dimethylxanthine	0.4309	7.473E-27
2	DNQX	0.6196	9.03E-90		DNQX	0.7111	1.504E-19
3	Muscimol hydrobromide	0.673	6.27E-80		Muscimol hydrobromide	0.7033	7.403E-20
4	Etodolac	0.7236	2.34E-17		Etodolac	0.7538	7.702E-06
5	Dihydro-beta-erythroidine hydrobromide	0.7413	2.18E-17		Dihydro-beta-erythroidine hydrobromide	0.7404	5.575E-10
6	1-(m-Chlorophenyl)-biguanide hydrochloride	0.7604	4.25E-43		1-(m-Chlorophenyl)-biguanide hydrochloride	0.7618	1.375E-12
7	R(-)-2,10,11-Trihydroxy-N-propylnoraporphine hydrobromide	0.7703	1.45E-09		R(-)-2,10,11-Trihydroxy-N-propylnoraporphine hydrobromide	0.7419	1.246E-09
8	Doxazosin mesylate	0.7769	5.59E-10		Doxazosin mesylate	0.7668	7.075E-06

Table 3.4: Compounds that elicited a decrease in the observed ratio of compound-treated GFP compared to control GFP in both the single *TIF1* and *TIF2* GFP screens (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in sections 2.2.3.4 and 3.3.1).

It is interesting to note that although 8-cyclopentyl-1,3-dimethylxanthine was shown to elicit a down regulation in GFP expression in both the TIF1p-GFP and TIF2p-GFP screens, the structurally related compound 1,3-diethyl-8-phenylxanthine did not. Furthermore, 1,3-diethyl-8-phenylxanthine was shown to elicit an increase in GFP expression of TIF2p-GFP, albeit at the arbitrary threshold not to a statistically significant level. A brief structural comparison indicates the observed changes are possibly compound specific

and, along with the other compounds highlighted in table 4.3, are deserving of further consideration.

Figure 3.9 and Figure 3.10 illustrates the comparison between the expression levels of the GFP of TIF1p-GFP and TIF2p-GFP as they are subjected to the LOPAC library. Figure 3.9 contains all compounds of the LOPAC library regardless of the p and Q values ascertained by the Wilcoxon rank sum test and Bonferroni corrections respectively.

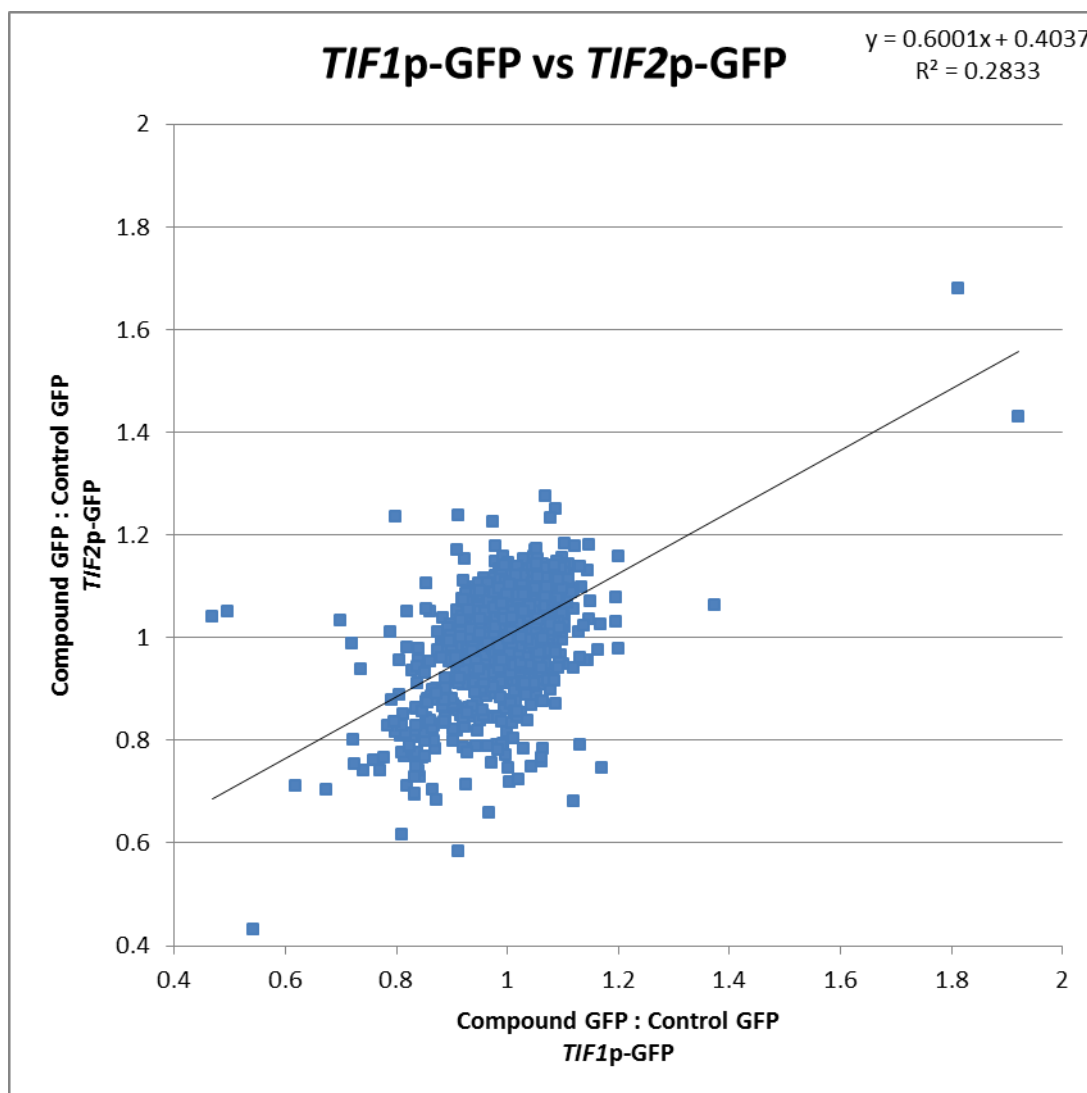


Figure 3.9: Comparison of the GFP components from the single TIF1p-GFP screen compared with that from the TIF2p-GFP screen. Each individual point represents a single compound from the LOPAC library with its associated compound-treated GFP: control GFP ratio for both the TIF1p-GFP and TIF2p-GFP screen.

As previously stated compounds are able to instigate changes in fluorescence by a plethora of mechanisms, many of which relate to modulation of the GFP signal and are not a gene regulated mechanism. An inherent level of compound fluorescence was one mechanism postulated in the cases of SB 216763 and SU 5416. For ease of comparison between the single GFP strains these compounds have been removed from the following comparison (the difference between Figure 3.9 and Figure 3.10).

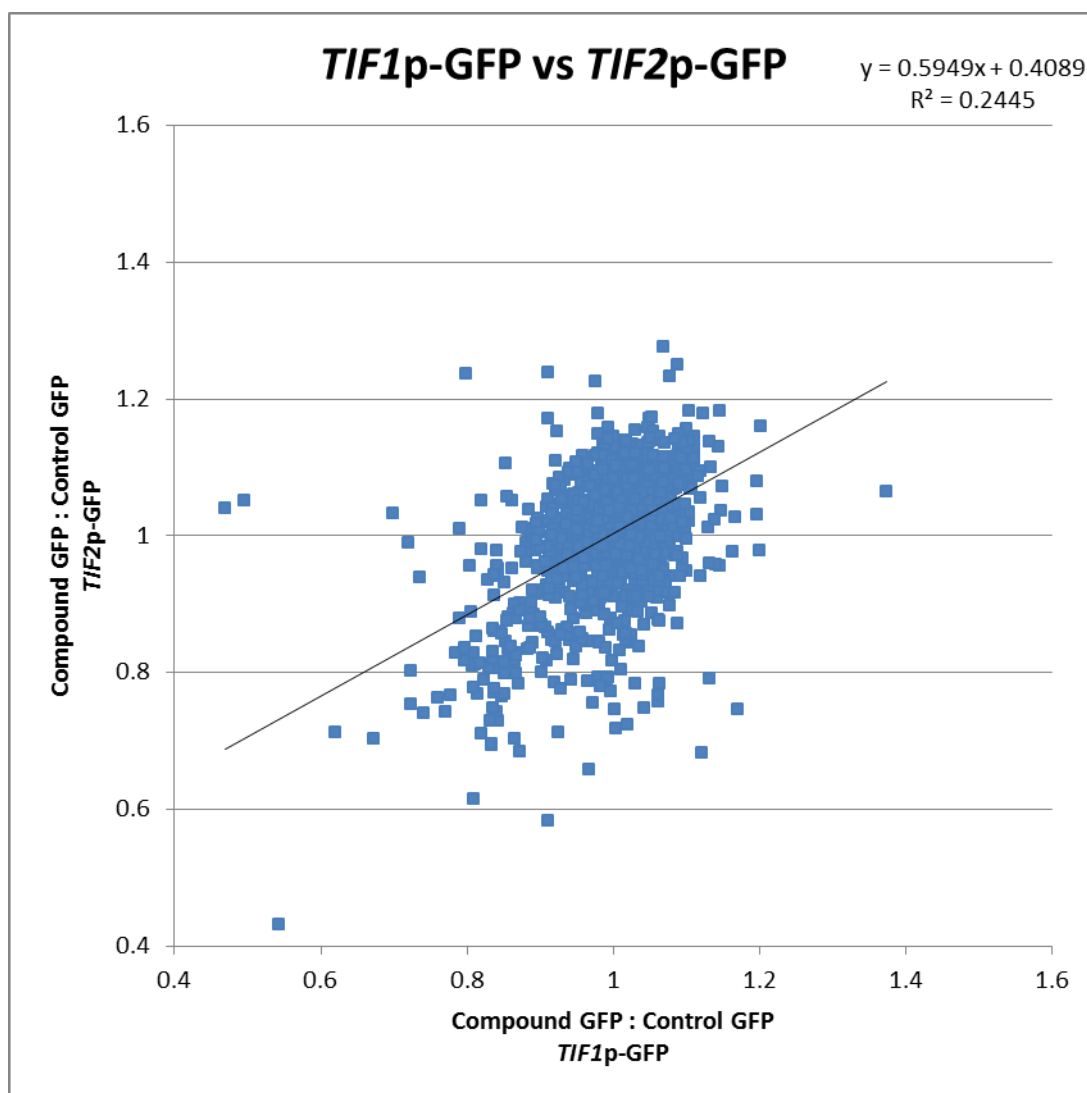


Figure 3.10: Comparison of the GFP components from the single TIF1p-GFP screen compared with that from the TIF2p-GFP screen. This figure is the same data set as Figure 3.9 but with a restricted range as both SB 216763 and SU 5416 were removed from this data set (see above). Each individual point represents a single compound from the LOPAC library with its associated compound-treated GFP: control GFP ratio for both the TIF1p-GFP and TIF2p-GFP screen.

What is not apparent from the figures above, which show the ratio of compound-treated fluorescence to control fluorescence, is that of these two duplicate genes *TIF2* is more highly expressed (Comparison from these screens indicated *TIF2* was expressed ~1.8 x the levels exhibited by *TIF1*). Subsequently Myricetin and Mainserin hydrochloride, which were observed suppressing the GFP signal in the TIF1p-GFP screen (the two data points farthest left on the above plots), were potentially not identified in both screens as the more highly expressed system is innately suited to buffer out small variations.

3.3.3 Dual Fluorescent GFP/RFP screening – Reading at 4 Hours

As previously stated, genetic redundancy between duplicate genes presents a major hurdle as the application of standard techniques may not have a noticeable effect due to the functional substitution of the duplicate gene. Therefore in order to dissect eIF4A's role as a part of the yeast translational machinery whilst maintaining total transparency, all manifestations of the genes encoding eIF4A, namely *TIF1* and *TIF2*, were evaluated in a single environment.

Although this studies primary aim, to evaluate whether *TIF1* and *TIF2* can be individually regulated by small molecule intervention, can be achieved by the development and screening of TIF1p-RFP TIF2p-GFP (presented in section 3.3.3.3), the reliance of the reporter strains on a plasmid based *TIF1*-RFP introduces the unknown disparities between the genomic architecture of chromosomal and plasmid borne expression systems. The second strain, TIF1p-RFP TIF1p-GFP, was generated to investigate this disparity; the result of which are presented in section 3.3.3.1.

3.3.3.1 TIF1p-GFP TIF1p-RFP Screen

3.3.3.1.1 TIF1p-GFP component

Of the compounds screened against TIF1p-GFP TIF1p-RFP 209 compounds, as ascertained by the Wilcoxon rank sum test ($\alpha = 0.01$), elicited a significant increase in GFP expression when compared with control GFP. However, 106 compounds were unearthed in the same screen at the initial time point. Conversely, 367 compounds elicited a significant decrease in GFP expression when compared with control GFP with 213 compounds identified at the initial time point.

Using an arbitrary established threshold of compound-treated GFP: control GFP ratio to constitute positive hits; above 1.2 and below 0.8, 13 compounds increased GFP expression whilst 1 decreased GFP expression. Of these 14 compounds, 8 compounds (ellipticine, ethosuximide, GW2974, idarubicin, L-765,314, quinacrine dihydrochloride, SB 216763, SU 5416) were identified as causing an upregulation in fluorescence intensity at both four hour and initial time points. Although identified as statistically significant by the Wilcoxon rank sum test ($\alpha = 0.01$), these compounds were removed from detailed consideration here for the reasons outlined in section 3.3. However, in future studies they will be deserving of consideration.

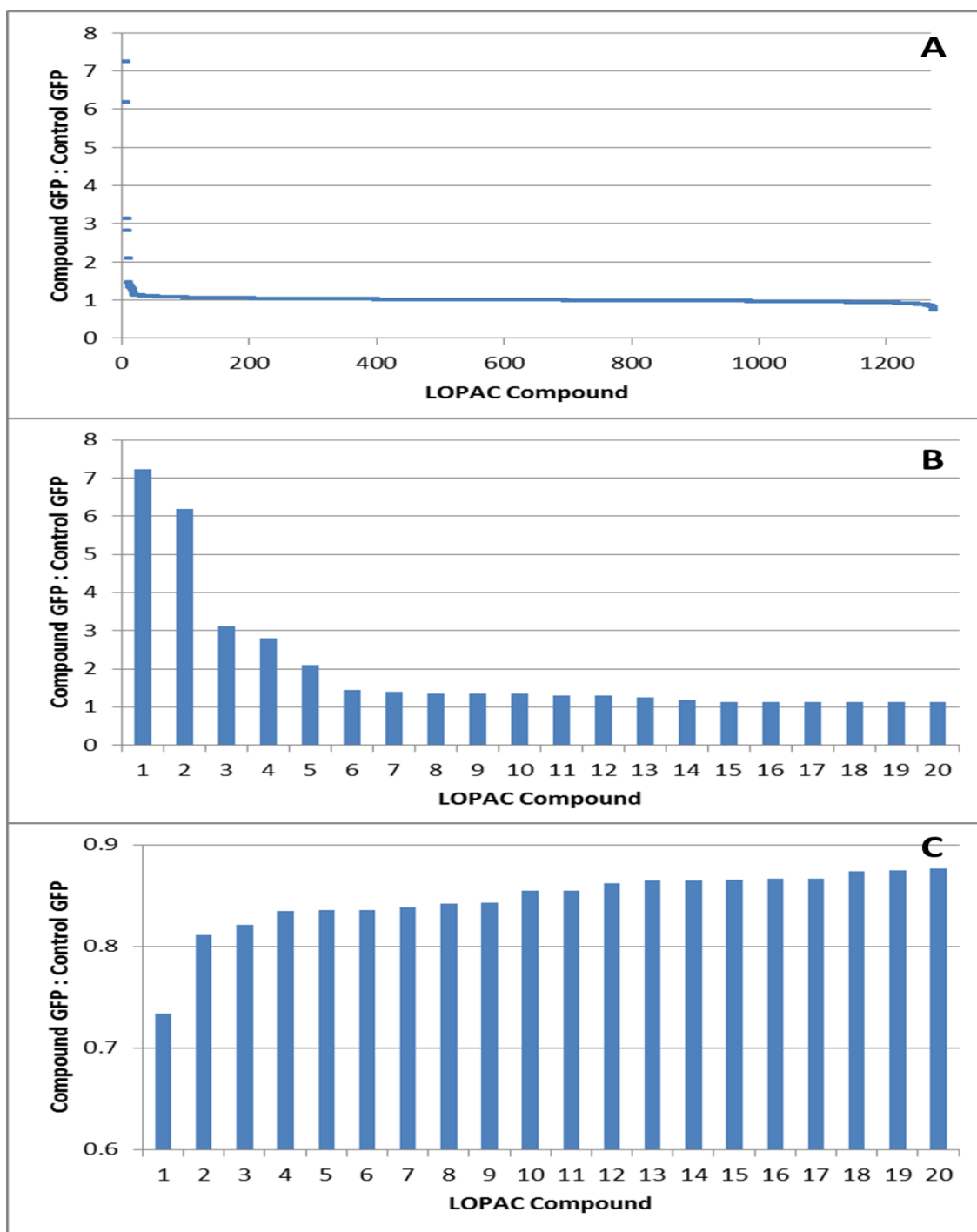


Figure 3.11: Whole cell fluorescence intensities of the GFP component of TIF1p-GFP TIF1p-RFP treated with the LOPAC library (compounds 1 – 1280) compared against the pooled collection of whole cell fluorescence intensities of TIF1p-GFP TIF1p-RFP treated with DMSO (as described in section 2.2.3.4). (A) Compound-treated GFP: control GFP for the entire LOPAC library ranging from highest to lowest. The top 20 compounds that (B) elicited an increase in or (C) elicited a decrease in the ratio of compound-treated GFP compared to control. Compounds highlighted in B and C are presented in Table 3.5.

#	Compound	Compound GFP : Control GFP	Q value	Compound	Compound GFP : Control GFP	Q value
1	SB 216763 *	7.229	3.9E-107	Ro 16-6491 hydrochloride	0.7344	3.403E-17
2	SU 5416 *	6.18	7.2E-101	CGS-12066A maleate	0.811	3.233E-62
3	Ethosuximide *	3.124	1.21E-13	DNQX	0.8212	3.688E-48
4	4-Amino-1,8-naphthalimide	2.807	1.37E-63	Droperidol	0.8348	8.963E-37
5	GW2974 *	2.086	4.8E-154	1-(m-Chlorophenyl)-biguanide hydrochloride	0.8354	2.695E-36
6	Idarubicin *	1.448	2.2E-110	N6-Cycl ohexyladenosine	0.8361	2.248E-62
7	Quinacrine dihydrochloride *	1.384	1.3E-131	Doxylamine succinate	0.8385	1.757E-48
8	Emodin	1.347	8.79E-33	Dextromethorphan hydrobromide monohydrate	0.8418	1.195E-69
9	Ellipticine *	1.34	9.8E-06	SB 203186	0.8429	2.964E-35
10	L-765,314 *	1.34	1.53E-30	CP55940	0.855	1.478E-48
11	Sanguinarine chloride	1.307	2.2E-123	Glipizide	0.8551	3.578E-35
12	SU 4312	1.296	7.91E-61	4'-Chloro-3- α -(diphenylmethoxy)tropane hydrochloride	0.862	5.913E-33
13	GW7647	1.248	5.5E-18	(\pm)-2-Amino-7-phosphonoheptanoic acid	0.8643	2.71E-14
14	Tyrphostin 47	1.17	4.03E-40	2',3'-didehydro-3'-deoxythymidine	0.8651	6.837E-54
15	Dequalinium analog, C-14 linker	1.141	1.48E-75	Diltiazem hydrochloride	0.8661	1.42E-55
16	Glibenclamide	1.129	1.41E-10	5,5-Dimethyl-1-pyrroline-N-oxide	0.8664	2.271E-43
17	2,3-Dimethoxy-1,4-naphthoquinone	1.126	2.31E-51	S(-)-Pindolol	0.8667	2.168E-31
18	T-1032	1.124	3.12E-38	Y-27632 dihydrochloride	0.8741	1.463E-42
19	LY-367,265	1.119	5.6E-50	WB-4101 hydrochloride	0.8743	1.365E-46
20	1,1-Dimethyl-4-phenyl-piperazinium iodide	1.118	7.58E-46	Doxepin hydrochloride	0.8766	2.065E-53

Table 3.5: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF1p-GFP TIF1p-RFP screen (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in sections 2.2.3.4 and 3.3.1). Compounds marked with an asterisk denote those compounds which appear in both t=0 and t=4 hours.

3.3.3.1.2 TIF1p-RFP component

Of the compounds screened against TIF1p-GFP TIF1p-RFP 104 compounds, as ascertained by the Wilcoxon rank sum test ($\alpha = 0.01$), elicited a significant increase in RFP expression when compared with control RFP. However, 15 compounds were unearthed in the same screen at the initial time point. Conversely, 95 compounds elicited a significant decrease in RFP expression when compared with control RFP with 11 compounds identified at the initial time point.

Using an arbitrary established threshold of compound-treated RFP: control RFP ratio to constitute positive hits; above 1.2 and below 0.8, 37 compounds increased RFP expression whilst 30 decreased RFP expression. Of these 67 compounds, 2 compounds; cyclophosphamide monohydrate and metrifudil, were identified as causing an upregulation in fluorescence intensity, whilst only diclofenac sodium was identified as causing a down regulation in fluorescence intensity at both four hour and initial time points. Although identified as statistically significant by the Wilcoxon rank sum test ($\alpha = 0.01$), these three compounds were removed from detailed consideration here for reasons outlined in section 3.3. However, in future studies they will be deserving of consideration.

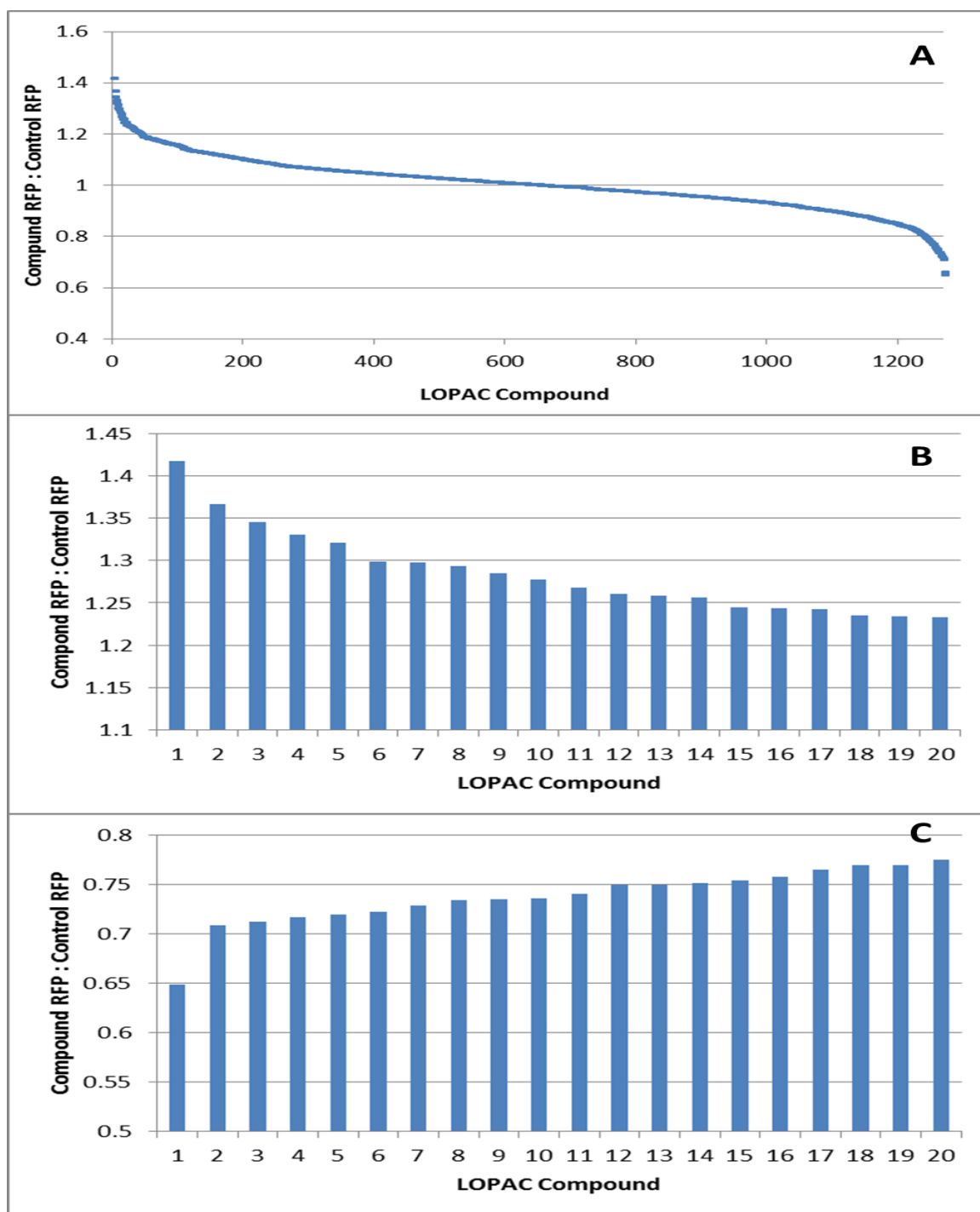


Figure 3.12: Whole cell fluorescence intensities of the RFP component of TIF1p-GFP TIF1p-RFP treated with the LOPAC library (compounds 1 – 1280) compared against the pooled collection of whole cell fluorescence intensities of TIF1p-GFP TIF1p-RFP treated with DMSO (as described in section 2.2.3.4). (A) Compound-treated RFP: control RFP for the entire LOPAC library ranging from highest to lowest. The top 20 compounds that (B) elicited an increase in or (C) elicited a decrease in the ratio of compound-treated RFP compared to control RFP. Compounds highlighted in B and C are presented in Table 3.6.

#	Compound	Compound RFP : Control RFP	Q value	Compound	Compound RFP : Control RFP	Q value
1	L-Hyoscyamine	1.418	7.27E-18	SB 218795	0.6488	6.946E-09
2	DL-threo-beta-hydroxyaspartic acid	1.366	4.56E-08	GYKI 52466 hydrochloride	0.7082	0.0001214
3	H-8 dihydrochloride	1.345	3.31E-10	Dequalinium dichloride	0.7121	5.535E-34
4	CNS-1102	1.331	2.08E-23	Doxylamine succinate	0.717	3.942E-14
5	2,6-Diamino-4-pyrimidinone	1.321	1.19E-13	Dextromethorphan hydrobromide monohydrate	0.7195	4.891E-25
6	Metrifudil *	1.299	7.72E-10	Retinoic acid p-hydroxyanilide	0.7219	0.0003869
7	N6-2-(4-Aminophenyl)ethyl adenosine	1.298	1.28E-06	N6-Cyclohexyladenosine	0.7284	1.747E-17
8	Ro 41-0960	1.293	2.46E-14	Diacylglycerol kinase inhibitor I	0.7338	8.632E-22
9	CX 546	1.285	8.96E-19	5,5-Dimethyl-1-pyrroline-N-oxide	0.7345	5.551E-19
10	RX 821002 hydrochloride	1.278	1.01E-10	SB 203186	0.7355	2.111E-10
11	(±)-Sotalol hydrochloride	1.268	2.8E-11	Diclofenac sodium *	0.7403	1.968E-13
12	2-Methylthioadenosine diphosphate trisodium	1.261	9.27E-07	IMID-4F hydrochloride	0.749	0.0002448
13	Isoguvacine hydrochloride	1.258	2.19E-05	5-fluoro-5'-deoxyuridine	0.7495	1.763E-10
14	2,3-Dimethoxy-1,4-naphthoquinone	1.257	4.63E-18	Hispidin	0.7512	2.181E-05
15	Xylometazoline hydrochloride	1.245	7.07E-09	2',3'-didehydro-3'-deoxythymidine	0.7539	3.404E-21
16	SR 57227A	1.243	2.15E-09	Cyclothiazide	0.7574	2.079E-21
17	2-Methylthioadenosine triphosphate tetrasodium	1.243	7.58E-06	R-(+)-8-Hydroxy-DPAT hydrobromide	0.7643	0.0006394
18	Amfonelic acid	1.235	3.17E-05	CP55940	0.7691	1.677E-07
19	R-(+)-7-Hydroxy-DPAT hydrobromide	1.234	1.91E-06	Doxepin hydrochloride	0.7694	1.736E-14
20	Rottlerin	1.233	1.27E-09	DNQX	0.7748	2.906E-05

Table 3.6: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated RFP compared to control RFP from the TIF1p-GFP TIF1p-RFP screen (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in section 2.2.3.4 and 3.3.1). Compounds marked with an asterisk denote those compounds which appear in both t=0 and t=4 hours.

3.3.3.2 Comparison of fluorescent components of TIF1p-GFP TIF1p-RFP

As ascertained by the Wilcoxon rank sum test ($\alpha = 0.01$) only 4-amino-1,8-naphthalimide was identified with the fluorescence intensity outside, in this case above, the arbitrarily chosen threshold of compound-treated fluorescence: control fluorescence ratio; for both the green and red fluorescently tagged *TIF1*. It is interesting to note that 4-amino-1,8-naphthalimide was also shown to upregulate *TIF1* in the single *TIF1*-GFP screen. However, due to the total number of cells in each of the quadruplicate wells dosed with 4-amino-1,8-naphthalimide being lower than the required Acapella threshold for image acquisition, data could not be collected at time zero. Data became available after 4 hours however, as each strain replicated during this incubation period and the total number of cells rose above the threshold required for image acquisition. As discussed in section 3.3, without this level of control it was impossible to experimentally determine whether the increase in fluorescence intensity, in response to 4-amino-1,8-naphthalimide, is because of a gene regulation response or from either intrinsic fluorescence of 4-amino-1,8-naphthalimide or some compound induced disturbances to fluorescent signal. However, as literature reveals that 4-amino-1,8-naphthalimide is in fact a fluorophore which emits within the green spectrum, 4-amino-1,8-naphthalimide will not be considered further (Parkesh, Clive Lee et al. 2007).

Figure 3.13 and Figure 3.14 illustrates the comparison between the expression levels of the GFP and RFP components of TIF1p-GFP TIF1p-RFP as they are subjected to the LOPAC library. Figure 3.13 contains all compounds of the LOPAC library regardless of the p and Q values ascertained by the Wilcoxon rank sum test and Bonferroni corrections respectively.

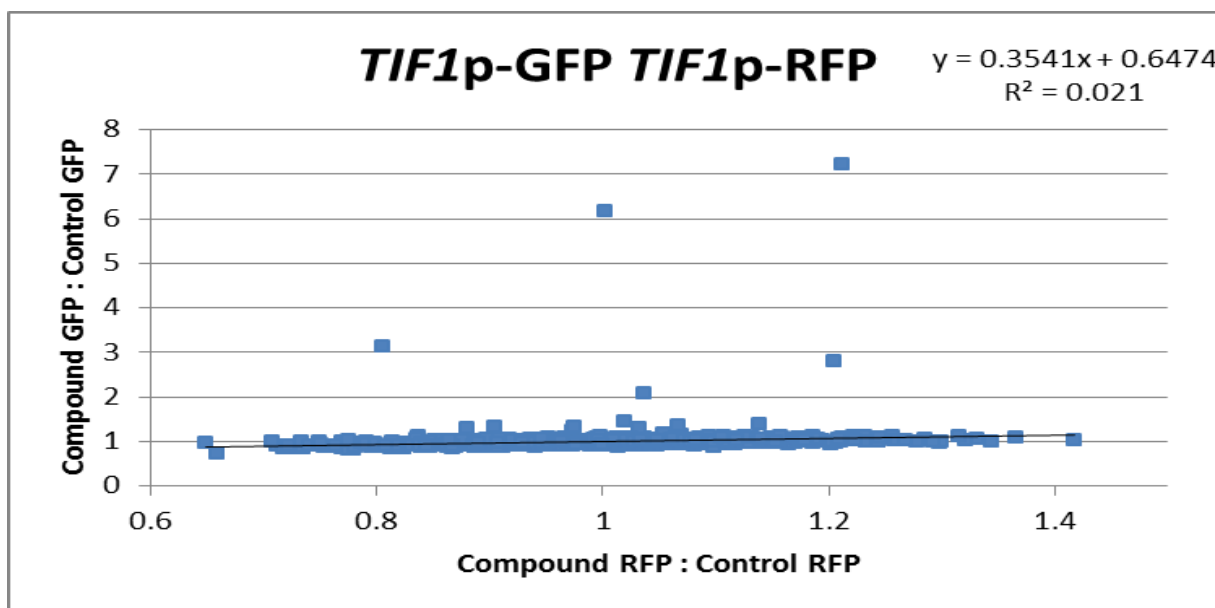


Figure 3.13: Comparison of the GFP and RFP components from the TIF1p-GFP TIF1p-RFP LOPAC library screen. Each individual point represents a single compound from the LOPAC library with its associated compound-treated GFP: control GFP ratio and compound-treated RFP: control RFP ratio for the GFP and RFP components respectively of TIF1p-GFP TIF1p-RFP.

As stated in section 3.3, compounds are able to instigate changes in fluorescence by a number of mechanisms and in regards to SB 216763 and SU 5416 it was postulated that an inherent level of compound fluorescence may be responsible. For ease of comparison these compounds have been removed from the following comparison (the difference between Figure 3.13 and Figure 3.14).

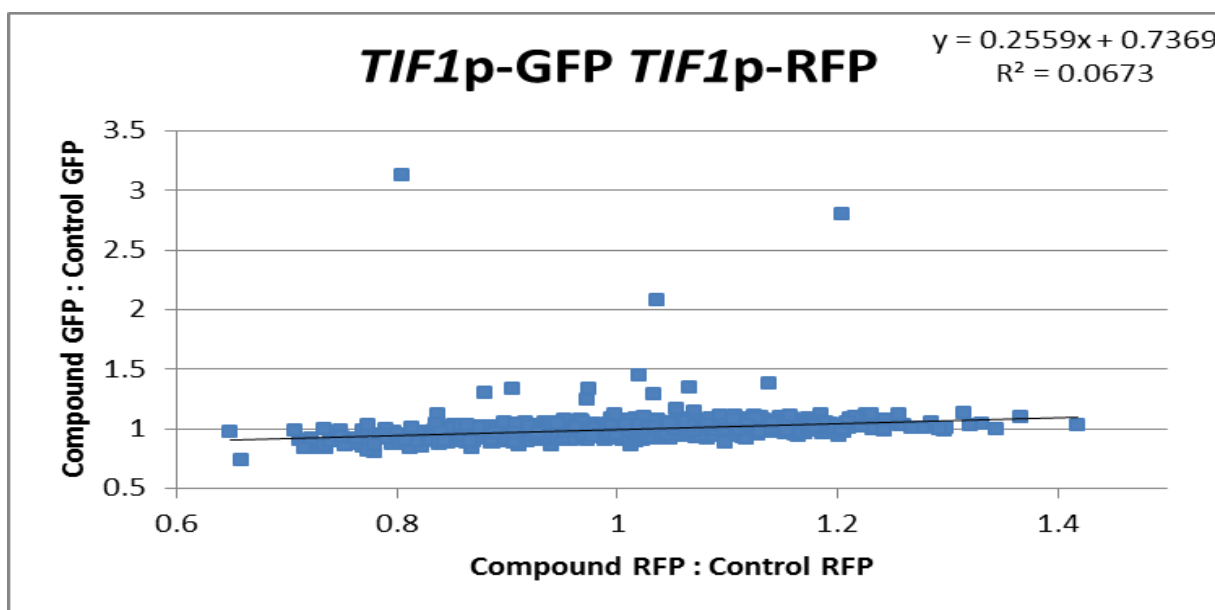


Figure 3.14: Comparison of the GFP and RFP components from the TIF1p-GFP TIF1p-RFP LOPAC library screen. This figure is the same data set as Figure 3.13 but with a restricted range as both SB 216763 and SU 5416 were removed from this data set (see above). Each individual point represents a single compound from the LOPAC library with its associated compound-treated GFP: control GFP ratio and compound-treated RFP: control RFP ratio for the GFP and RFP components respectively of TIF1p-GFP TIF1p-RFP.

From the remaining dataset the only notable compound, other than 4-amino-1,8-naphthalimide, which falls close to the arbitrarily threshold of compound-treated fluorescence: control fluorescence ratio for both the green and red fluorescently tagged *TIF1* is ethosuximide. Interestingly ethosuximide was shown to elicit an upregulation in the GFP component but a down regulation in the RFP component (compound-treated fluorescence: control fluorescence of 3.12 and 0.80 respectively). Furthermore it was found that in only the GFP component ethosuximide was highlighted as a hit at both the four and zero hour time points. A brief structural investigation suggests ethosuximide is not inherently fluorescent however; it is possible that ethosuximide may be able to modulate and enhance the GFP signal by other methods and subsequently deserving of further consideration.

It is important to note the relatively poor model fit illustrated by the parameters in Figure 3.13 and Figure 3.14. The gradient suggests that there is some proportionality, but not

close to the ideal 1:1 ratio between RFP change and GFP change, whilst the intercept suggests that there would be GFP present even if RFP was brought to zero. Additionally it appears that the removal of the fluorescent compounds SB 216763 and SU 5416, which are outliers in this system, results in a gradient even further removed from the ideal 1:1 ratio. This poor model fit makes it inadvisable to construct any decisive conclusion from these observations. However the above figures do eloquently illustrate that the RFP signal from *TIF1*-RFP is weaker than that observed for the GFP equivalent, as was stated in section 3.2.

3.3.3.3 TIF2p-GFP TIF1p-RFP Screen

3.3.3.3.1 TIF2p-GFP component

Of the compounds screened against TIF2p-GFP TIF1p-RFP 66 compounds, as ascertained by the Wilcoxon rank sum test ($\alpha = 0.01$), elicited a significant increase in GFP expression when compared with control GFP. However, 72 compounds were unearthed in the same screen at the initial time point. Conversely, 174 compounds elicited a significant decrease in GFP expression when compared with control GFP with 73 compounds identified at the initial time point.

Using an arbitrary established threshold of compound-treated GFP: control GFP ratio to constitute positive hits; above 1.2 and below 0.8, 12 compounds increased GFP expression whilst none decreased GFP expression. Of these 12 compounds only, 1-(4-fluorobenzyl)-5-methoxy-2-methylindole-3-acetic acid and quinacrine dihydrochloride were found to upregulate fluorescence intensity at the four hour time point and not time zero. The remaining 10 compounds, although identified as statistically significant by the Wilcoxon rank sum test ($\alpha = 0.01$), were removed from detailed consideration here for reasons outlined in section 3.3. However, in future studies they will be deserving of consideration.

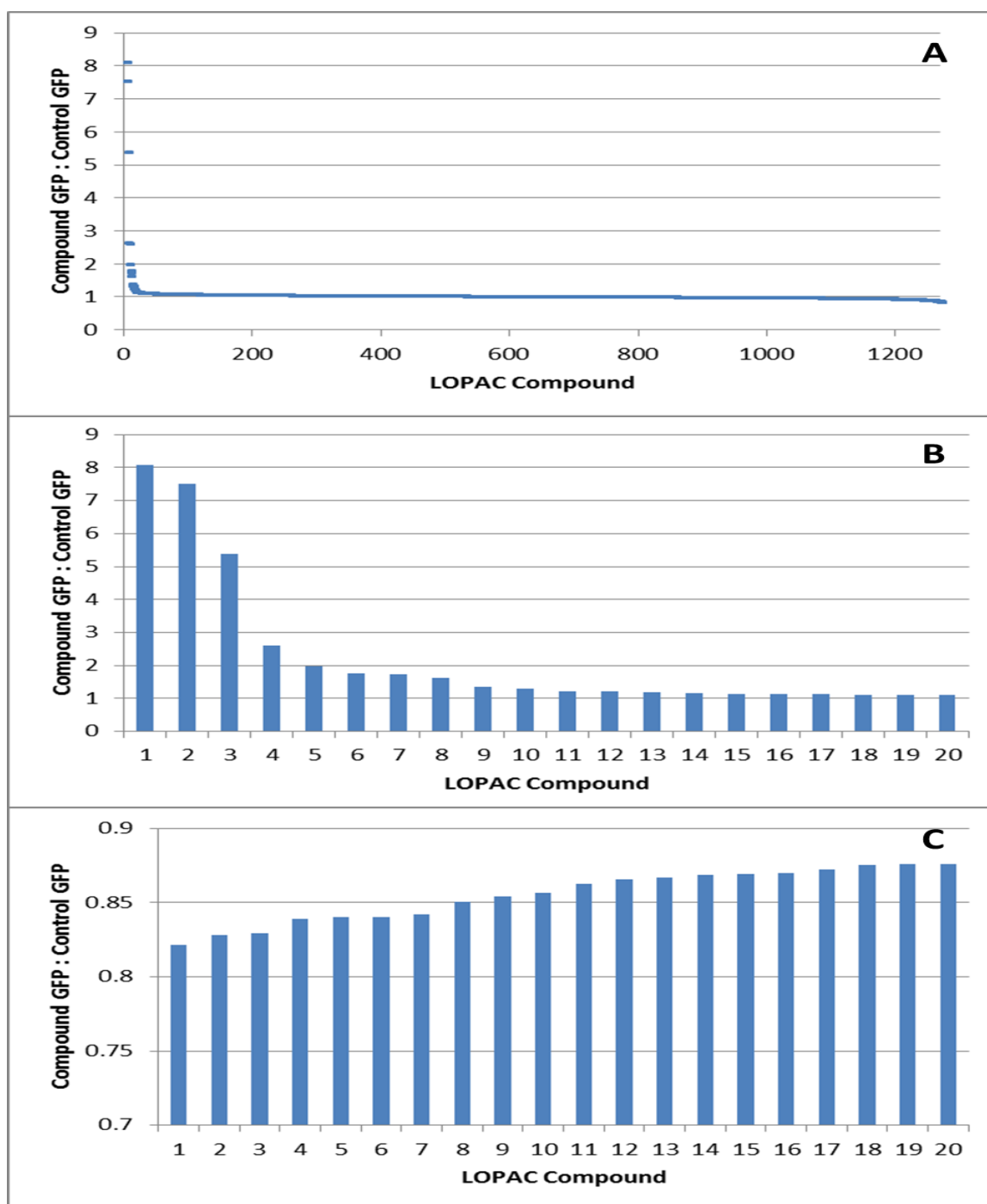


Figure 3.15: Whole cell fluorescence intensities of the GFP component of TIF2p-GFP TIF1p-RFP treated with the LOPAC library (compounds 1 – 1280) compared against the pooled collection of whole cell fluorescence intensities of TIF2p-GFP TIF1p-RFP treated with DMSO (as described in section 2.2.3.4). (A) Compound-treated GFP: control GFP for the entire LOPAC library ranging from highest to lowest. The top 20 compounds that (B) elicited an increase in or (C) elicited a decrease in the ratio of compound-treated GFP compared to control GFP. Compounds highlighted in B and C are presented in Table 3.7.

#	Compound	Compound GFP : Control GFP	Q value	Compound	Compound GFP : Control GFP	Q value
1	SU 5416 *	8.084	4.83E-91	SKF-525A hydrochloride	0.8212	9.459E-19
2	SB 216763 *	7.518	1.03E-57	DNQX	0.8278	6.194E-18
3	Emodin *	5.378	1.11E-73	1-(m- Chlorophenyl)- biguanide hydrochloride	0.8296	4.49E-18
4	4-Amino-1,8- naphthalimide *	2.606	4.61E-52	Droperidol	0.8391	2.227E-18
5	Sanguinarine chloride *	1.971	2.35E-34	Minoxidil	0.8404	7.117E-22
6	Papaverine hydrochloride *	1.773	2.69E-42	L-Ca n a vanine sulfate	0.8404	1.236E-10
7	Idarubicin *	1.733	2.13E-84	1,7- Dimethylxanthine	0.842	4.367E-20
8	Nylidrin hydrochloride *	1.618	1.07E-35	Haloperidol	0.8504	3.437E-16
9	Tyrphostin 47 *	1.364	0.000456	6(5H)- Phenanthridinone	0.8542	3.221E-20
10	SU 4312 *	1.306	2.29E-14	Phenylbutazone	0.8566	6.598E-26
11	1-(4-Fluorobenzyl)-5- methoxy-2- methylindole-3-acetic acid	1.22	5.11E-30	Prazi quantel	0.8627	5.648E-13
12	Quinacrine dihydrochloride	1.215	4.52E-42	DD1	0.8655	5.557E-09
13	GW5074	1.178	0.000108	Protriptyline hydrochloride	0.8666	7.171E-20
14	Pyridostigmine bromide	1.154	8.1E-15	(±)-Ibotenic acid	0.8687	5.456E-15
15	GW2974	1.127	1.3E-08	Danazol	0.869	1.302E-13
16	2,2'-Bipyridyl	1.123	1.49E-13	(±)-Bay K 8644	0.8699	1.396E-18
17	Agroclavine	1.121	1.47E-10	Doxylamine succinate	0.8724	2.405E-13
18	Propentofylline	1.115	1.5E-09	Protoporphyrin IX disodium	0.8755	6.379E-15
19	Tetramisole hydrochloride	1.112	2.94E-06	Nalidixic acid sodium	0.8757	1.062E-14
20	Prochlorperazine dimaleate	1.109	1.22E-11	3'-Azido-3'- deoxythymidine	0.8761	7.262E-10

Table 3.7: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF2p-GFP TIF1p-RFP screen (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in section 2.2.3.4 and 3.3.1). Compounds marked with an asterisk denote those compounds which appear in both t=0 and t=4 hours.

3.3.3.3.2 TIF1p-RFP component

Of the compounds screened against TIF2p-GFP TIF1p-RFP 40 compounds, as ascertained by the Wilcoxon rank sum test ($\alpha = 0.01$), elicited a significant increase in RFP expression when compared with control RFP. However, 16 compounds were unearthed in the same screen at the initial time point. Conversely, 20 compounds elicited a significant decrease in RFP expression when compared with control RFP with 15 compounds identified at the initial time point.

Using an arbitrary established threshold of compound treated RFP: control RFP ratio to constitute positive hits; above 1.2 and below 0.8, 28 compounds increased RFP expression whilst 17 decreased RFP expression. Of these 45 compounds only, 3-amino-1-propanesulfonic acid sodium salt and Apigenin were found to upregulate fluorescence intensity at the four and zero hour time points. These compounds, although identified as statistically significant by the Wilcoxon rank sum test ($\alpha = 0.01$), were removed from detailed consideration here for the reasons outlined in section 3.3. However, in future studies they will be deserving of consideration.

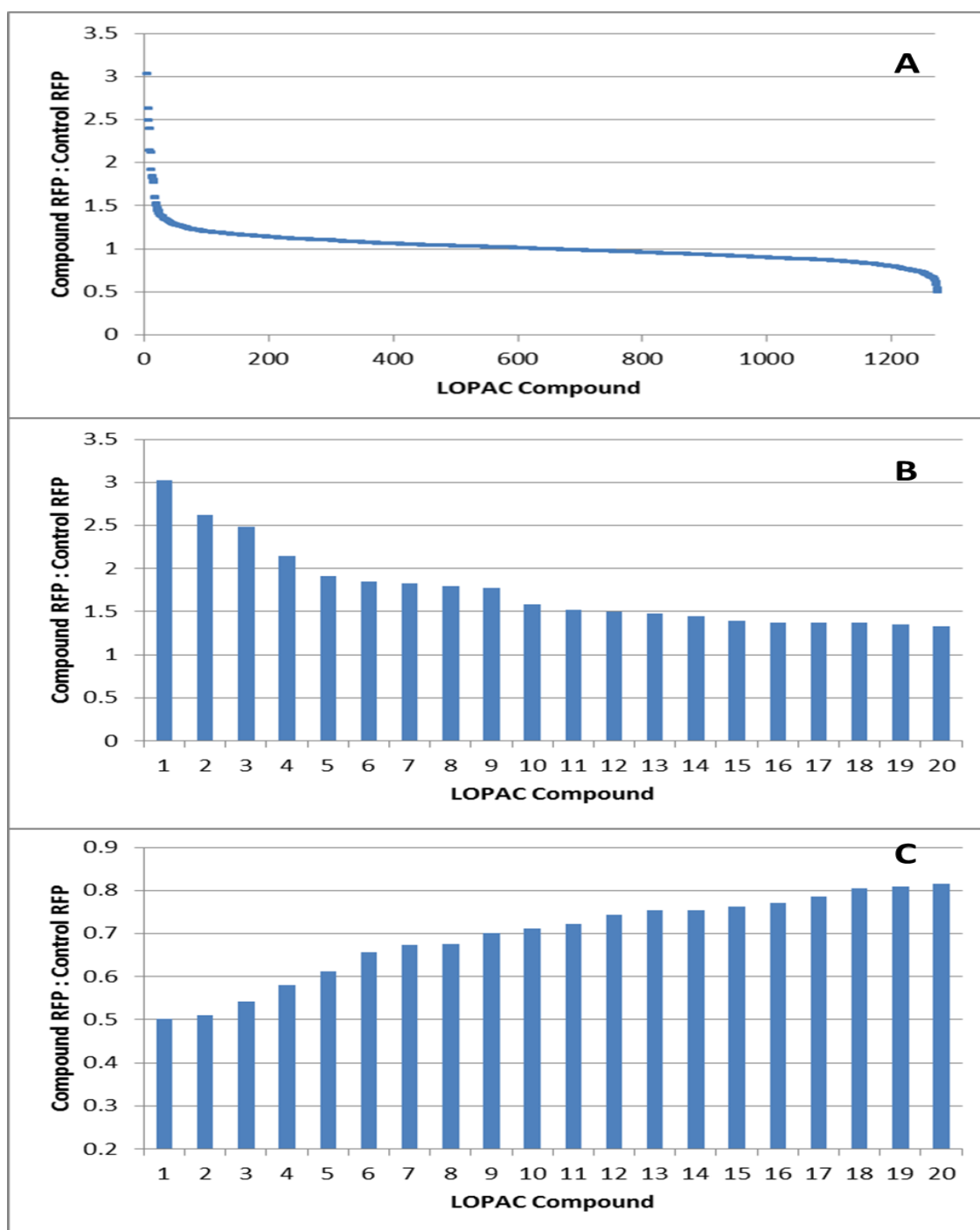


Figure 3.16: Whole cell fluorescence intensities of the RFP component of TIF2p-GFP TIF1p-RFP treated with the LOPAC library (compounds 1 – 1280) compared against the pooled collection of whole cell fluorescence intensities of TIF2p-GFP TIF1p-RFP treated with DMSO (as described in section 2.2.3.4). (A) Compound-treated RFP: control RFP for the entire LOPAC library ranging from highest to lowest. The top 20 compounds that (B) elicited an increase in or (C) elicited a decrease in the ratio of compound-treated RFP compared to control RFP. Compounds highlighted in B and C are presented in Table 3.8.

#	Compound	Compound RFP : Control RFP	Q value	Compound	Compound RFP : Control RFP	Q value
1	Cytosine-1-beta-D-arabinofuranoside hydrochloride	3.026	2.14E-17	L-Canavanine sulfate	0.5026	0.001666
2	1-Aminobenzotriazole	2.625	4.34E-32	DNQX	0.5101	1.498E-06
3	3-Amino-1-propanesulfonic acid sodium *	2.489	1.41E-19	Droperidol	0.5417	5.921E-08
4	Calcimycin	2.144	1.53E-09	Haloperidol	0.5814	0.0006219
5	R(+)-SCH-23390 hydrochloride	1.917	3.55E-13	GBR-12909 dihydrochloride	0.6118	9.836E-05
6	Emetine dihydrochloride hydrate	1.846	0.0004	2',3'-didehydro-3'-deoxythymidine	0.6568	0.002338
7	Flecainide acetate	1.823	0.000387	Dequalinium dichloride	0.673	5.227E-07
8	Emodin	1.796	9.26E-25	SKF-525A hydrochloride	0.676	0.008944
9	DM 235	1.773	0.003283	N-Methyl dopamine hydrochloride	0.7018	0.003531
10	Ouabain	1.585	1.11E-10	S(-)-3PPP hydrochloride	0.7121	0.002627
11	Apigenin *	1.523	0.003383	Chlorothiazide	0.7217	0.003485
12	Thapsigargin	1.505	0.000502	trans-Azetidine-2,4-dicarboxylic acid	0.7446	4.648E-12
13	Fluspirilene	1.474	0.000155	Phenelzine sulfate	0.7551	0.001337
14	Nylidrin hydrochloride	1.442	7.69E-07	AGN 192403 hydrochloride	0.7552	8.369E-07
15	(-)-cis-(1S,2R)-U-50488 tartrate	1.393	0.005407	6(5H)-Phenanthridinone	0.7618	0.005569
16	N-Oleylethanolamine	1.374	2.94E-07	Sulfaphenazole	0.7715	0.0002526
17	Vinblastine sulfate salt	1.374	0.001599	Acetohexamide	0.7853	0.00162
18	2,3-Dimethoxy-1,4-naphthoquinone	1.374	0.000773	Cefazolin sodium	0.8051	0.004663
19	SR-95531	1.348	0.004882	5'-N-Methyl carboxamidoadenosine	0.8088	0.005654
20	Zardaverine	1.332	0.004505	Chloroethyldonidine dihydrochloride	0.8166	0.005532

Table 3.8: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated RFP compared to control RFP from the TIF2p-GFP TIF1p-RFP screen (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in section 2.2.3.4 and 3.3.1). Compounds marked with an asterisk denote those compounds which appear in both t=0 and t=4 hours.

3.3.3.4 Comparison of fluorescent components of TIF2p-GFP TIF1p-RFP

As ascertained by the Wilcoxon rank sum test ($\alpha = 0.01$) three compounds; emodin, nylicrin hydrochloride, and papaverine hydrochloride, were identified as inducing fluorescence intensity above the arbitrarily chosen threshold of compound-treated fluorescence: control fluorescence; for both the fluorescent protein components of TIF2p-GFP TIF1p-RFP. However as stated in section 3.3.3.3.1, all three compounds were identified as causing an upregulation in fluorescence intensity of the GFP component of TIF2p-GFP TIF1p-RFP, at both the four and zero hour time points. Thus, these compounds were removed from detailed consideration here for the reasons outlined in section 3.3.

Figure 3.17 illustrates the comparison between the expression levels of the GFP and RFP components of TIF2p-GFP TIF1p-RFP once subjected to the LOPAC library. It contains all compounds of the LOPAC library regardless of the p and Q values ascertained by the Wilcoxon rank sum test and Bonferroni corrections respectively.

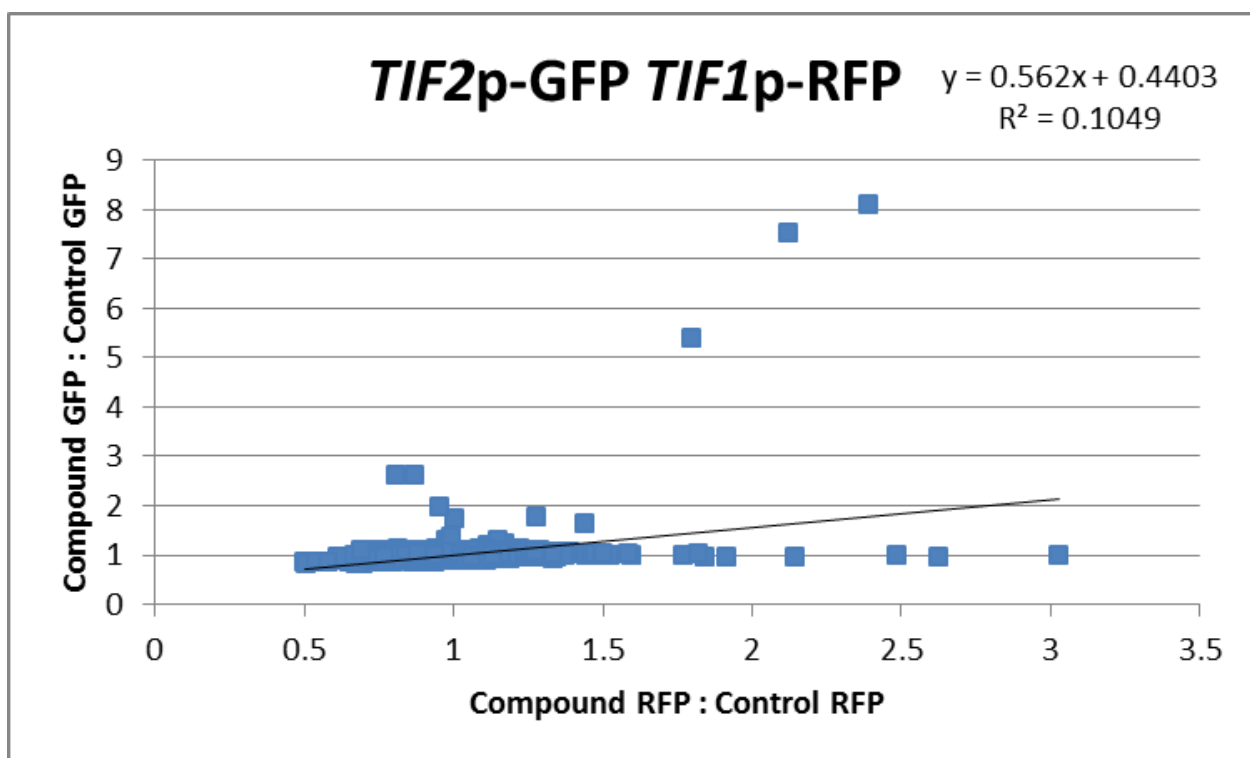


Figure 3.17: Comparison of the GFP and RFP components from the TIF2p-GFP TIF1p-RFP LOPAC library screen. Each individual point represents a single compound from the LOPAC library with its associated compound-treated GFP: control GFP ratio and compound RFP: control RFP ratio for the GFP and RFP components respectively of TIF2p-GFP TIF1p-RFP.

As stated in section 3.3, compounds are able to instigate changes in fluorescence by a number of mechanisms and in regards to SB 216763, SU 5416 and 4-amino-1,8-naphthalimide (the three highest data points on the y axis) it was postulated that an inherent level of compound fluorescence may be responsible. For ease of comparison these compounds, have been removed from the following comparison. Additionally, literature indicates that emodin is itself fluorescent, thus, for ease of comparison, Figure 3.18 was generated to also exclude emodin from its viewing range (Pecere, Gazzola et al. 2000; Hernandez, Recio et al. 2012).

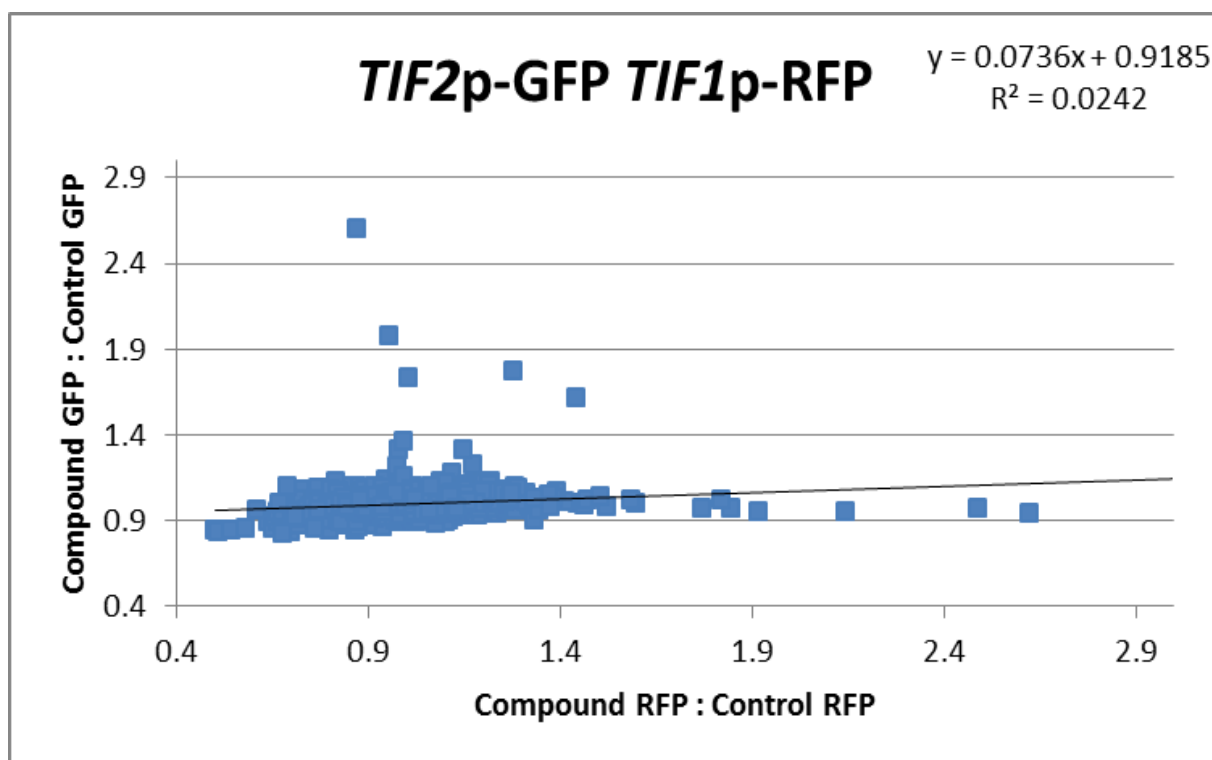


Figure 3.18: Comparison of the GFP and RFP components from the *TIF2p-GFP TIF1p-RFP* LOPAC library screen. This figure is the same data set as Figure 3.17 but with a restricted range as SB 216763, SU 5416, 4-amino-1,8-naphthalimide and emodin were removed from this data set (see above). Each individual point represents a single compound from the LOPAC library with its associated compound-treated GFP: control GFP ratio and compound-treated RFP: control RFP ratio for the GFP and RFP components respectively of *TIF2p-GFP TIF1p-RFP*.

Encouragingly, the gradient suggests that there is no strong relationship between RFP change and GFP change (Once the aforementioned fluorophores were removed).

If we consider that the RFP component of *TIF2-GFP TIF1-RFP*, *TIF1* has a far greater number of hits than the *TIF2* component, 43 and 2 compounds respectively (see section 3.3.3.3.1 and 3.3.3.3.2), then we can speculate the implications with regards to the roles of the paralogous genes. In the cell eIF4A is present at very high concentrations and as such the more highly expressed *TIF2* may be constitutively activated at level approaching its potential maximum. Upon probing with the LOPAC library, *TIF2* appears to be less able to effect change in the eIF4A levels than *TIF1* as seen by the number of hits generated for

each component. Therefore we can speculate, if we ignore all the noted problems with this system, that the responsibility for promoting change in eIF4A levels in response to chemical intervention was achieved by regulation of the generally lesser expressed *TIF1*.

This result not only illustrates that the two genes can be individually identified in the one environment but illustrates the importance of all mRNA regulatory domains. It is likely that this intervention is causing regulation differences due to uncharacterised sequence elements in the 5' and 3' UTR (an array of which were described in section 1.4) however, it would be unwise to completely rule out the role of the six synonymous changes in the two genes coding sequence. It is possible that these compounds are selectively increasing or decreasing the delivery of *TIF1*-specific tRNA based on these synonymous amino acid substitutions.

It is important to note that both of the compounds shown to upregulate the Tif2p-GFP component of *TIF2*-GFP *TIF1*-RFP; 1-(4-fluorobenzyl)-5-methoxy-2-methylindole-3-acetic acid and quinacrine dihydrochloride present only modest fluorescence increases. A brief structural analysis indicates these compounds may themselves be fluorescent and the apparent upregulation seen is possibly as result of an accumulation of these fluorescent compounds. Furthermore literature indicates that quinacrine dihydrochloride is itself fluorescent and has been used previously as a fluorescent probe (Lee 1971; Pearson, Bobrow et al. 1971).

3.3.3.5 Multiple Screen comparison

It is important to note that when comparing all other combination of the four screens described above; *TIF1*-GFP against the GFP component from *TIF1*-GFP *TIF1*-RFP, *TIF2*-GFP against the GFP component from *TIF2*-GFP *TIF1*-RFP, there is no overlap in any of the compounds that are shown to significantly increase or decrease fluorescence intensity outside the arbitrary threshold range, except for the combination described below.

3.3.3.5.1 Plasmid RFP comparison

As ascertained by the Wilcoxon rank sum test ($\alpha = 0.01$) five compounds were identified as regulating fluorescence intensity above or below the arbitrarily chosen threshold of compound-treated fluorescence: control fluorescence; for both the plasmid RFP component of *TIF2p*-GFP *TIF1p*-RFP and *TIF1p*-GFP *TIF1p*-RFP. Of these compounds, 2,3-dimethoxy-1,4-naphthoquinone and thapsigargin elicited an upregulation in fluorescence intensity, whilst DNQX, 2',3'-didehydro-3'-deoxythymidine and dequalinium dichloride elicited a down regulation in fluorescence intensity for both plasmid RFP components.

Figure 3.19 illustrates the comparison between the expression levels of the RFP components of *TIF1p*-GFP *TIF1p*-RFP and *TIF2p*-GFP *TIF1p*-RFP as they are subjected to the LOPAC library. It contains all compounds of the LOPAC library regardless of the p and Q values ascertained by the Wilcoxon rank sum test and Bonferroni corrections respectively.

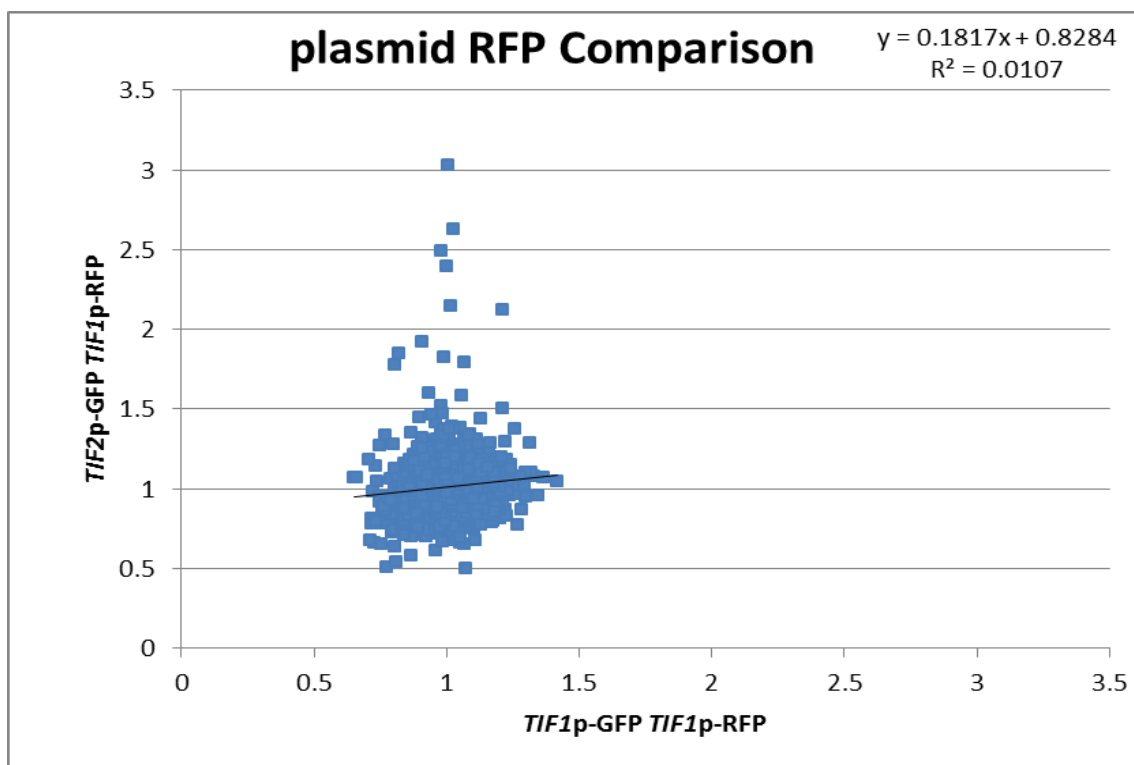


Figure 3.19 : Comparison of the RFP components from the TIF1p-GFP TIF1p-RFP screen compared with that from the TIF2p-GFP TIF1p-RFP screen. Each individual point represents a single compound from the LOPAC library with its associated compound-treated RFP: control RFP ratio for both the RFP components of TIF1p-GFP TIF1p-RFP and TIF2p-GFP TIF1p-RFP.

It is important to note that a small number of compounds appear to generate change in abundance of TIF1p-RFP in the TIF2p-GFP TIF1p-RFP strain but not in the TIF1p-GFP TIF1p-RFP equivalent when we ideally should see a clustering of hits. Since TIF1p-RFP is common to both strains, this may be an experimental artefact or may arise from underlying differences around the second source of Tif1p in these strains, with the TIF1-GFP having a disrupted 3' untranslated region and a protein product that may be less efficiently converted to a functional protein product due to the presence of the GFP tag.

3.4 Advantages and limitations of this screen

3.4.1 Advantages

The major advantages of the fluorescence based system constructed during this study is the ease with which this system can be manipulated and so is applicable to any ORF in the *S. cerevisiae* genome, under a number of interchangeable selection markers as well as providing the ability to be analysed by a diverse set of methodologies. Furthermore, this system was designed with a view to eliminating some of the problems associated with other commercially available fluorescent markers.

3.4.1.1 Universal application of plasmid borne ORF-RFP

The primer system employed in this research was designed so that in any future studies these existing primers are able to be easily modified allowing any gene in the genome to be fused with the RFP RedStar2. The specific primer system utilised in the creation of *TIF1*-RFP is outlined in Table 3.9.

Identification Number	Primer	Sequence
#484	Module 1 Forward	CAGTCACGACGTTGTA AACGACGGCCAGTGAATTGTAATTC AGCAACAACATCCGATGCTT
#485	Module 1 Reverse	ACCTGCACCAGCTCCAGCTCCGTTCAACAAAGTAGCGATGTCTG GATGGCAATTCTTCAATTT
#486	Module 2 Forward	AATTGAAGAATTGCCATCCGACATCGCTACTTTGTTGAACGGA GCTGGAGCTGGTGCAGG
#378	Module 2 Reverse	TAGCCTCACAAGATACTTTTTTAAGAAGTTTTGTCTCCCTTAC AAGAACAAGTGGTGTC
#487	Module 3 Forward	ACTGAAGGTAGACACCACTTGTTCTTGTAAGGGAGACAAAA CTTCTTAAAAAAGTATCTTGTGAGGCTATCTTG
#488	Module 3 Reverse A	CGGCTCCTATGTTGTGTGGAATTGTGAGCGGATAACAATTTT GATGTACACTTTTCTTTTCAG

Table 3.9: The primer set used in the construction of *TIF1*-RFP colour coded to highlight the origin of each component that makes up the three modules: Red – RFP RedStar2, green – pRS316, black – *TIF1*.

This primer system can be amended to suit any ORF by replacing the black (*TIF1*) sequences with sequences from the desired ORF that are equivalent to the *TIF1* ORF,

TIF1pr and *TIF1tr* regions illustrated in Figure 3.1 and outlined in section 3.1.1. Additionally as discussed in section 3.2 specific care must be taken when incorporating the upstream and downstream sections of the desired gene so as to incorporate all regions that have been previously categorised as active.

To be truly a universally interchangeable system these primers, once amended to suit any ORF in the genome have to maintain two essential functions. The first of which is creation of modules that incorporate the desired ORF, RFP and plasmid, and secondly that these modules successfully allow transformation into a parental yeast strain via homologous recombination. These two functions can roughly be achieved by selecting regions of the desired ORF, that at least with regard to gene position, mimic those utilised in this study. However, it is possible that the sections utilised are not conducive to successful module creation and homologous recombination. In these cases it is possible to amend the chosen sequences but a certain sequence length must be maintained to ensure successful homologous recombination.

Hua *et al.*, (1997) investigated the length of homology required for successful transformation of a functional plasmid. They identified the percentage of 2-micron plasmids, of ~2 kb in length bearing a Lac Z system, that following transformation were deemed to be functionally active by a beta-galactosidase activity assay. Hua equated that 30 bp of homologous sequence at each end of a DNA fragment was sufficient to integrate the fragment into a linearized plasmid in yeast (78.9% functionally active transformants). However, Hua acknowledged in order to obtain a high yield of active transformants 60 bp of sequence homology are desirable (95.8% functionally active transformants). Consequently any deviation from the sequence lengths outlined above must maintain at least 30bp of sequence homology at the end of each module to ensure successful transformation.

This primer set is not only amendable to universally fit any ORF in the genome but can be expressed on a number of different plasmids with varying selection markers. There is no need to manipulate the above sequences further as the CEN plasmid set (discussed in section 3.2) all include a multiple cloning site of identical sequence. This multiple cloning site incorporates the single copy of the motifs that the restriction enzymes XbaI and

HindIII are designated to digest, which allows for the use of a set of interchangeable plasmids. The major benefit to this universality is the ease of interchangeably available gene-RFP selection markers. This ability allows for a single fusion design to be expressed in a wider range of yeast strains as well as avoiding conflicts with selection markers rendering strains incompatible with some standard yeast techniques.

3.4.1.2 Fusion proteins and their endogenous regions

The yeast GFP clone collection is a set of *S. cerevisiae* open reading frames tagged at the carboxy terminal end using the coding sequence of Aequorea Victoria GFP (S65T GFP). This library has 75% coverage of the *S. cerevisiae* proteome and research utilising it was able to classify 4159 yeast GFP clones into one or more of 12 subcellular localisation categories (Huh, Falvo et al. 2003). However, this localisation data agrees with only 80% of the published data available on the Saccharomyces Genome Database (SGD. 2012). It has been suggested that a potential reason for this discrepancy could have been caused by the way the ~27 kDa GFP fusion protein was introduced into the library.

As illustrated in Figure 3.20, each open reading frame (ORF) was systematically GFP tagged in its chromosomal location through oligonucleotide-directed homologous recombination. These methodologies lead to the introduction of each GFP tag, and its associated *HIS3* selection marker, directly after the stop codon of each ORF. This GFP tag disrupts the wild type 3' UTR and may disrupt any transcript localisation motifs located in this region. This potential problem may have an effect the GFP components of both dual fluorescent proteins which bear this tag and may be a leading cause as to the observed discrepancies in protein localisation described by Huh (2003).

The central dogma suggests that polypeptides will only fold into functional proteins if they are allowed to fold in unique domains thus this disruption may affect normal protein folding. Furthermore, the GFP source in this library does not have a linker region which is conducive to protein folding. However, there is no implication that these issues have an effect on either of the *TIF* genes.

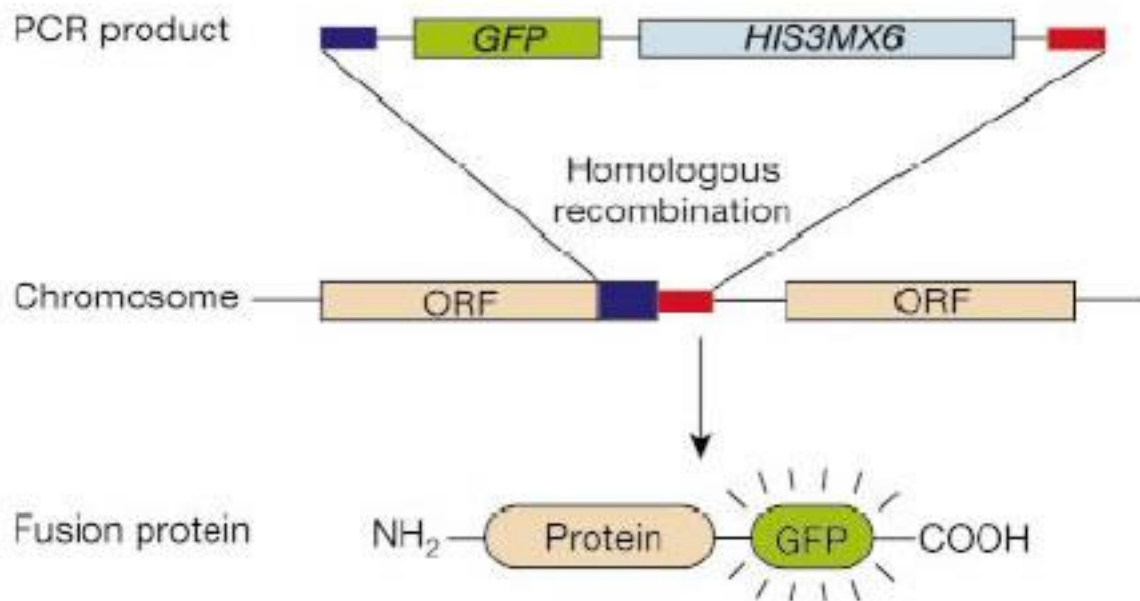


Figure 3.20: Strategy for GFP library construction. PCR products containing the GFP tag and a selectable marker gene were inserted at the C terminus of each ORF through homologous recombination, yielding a C-terminally GFP-tagged protein. Figure from Huh (2003).

In order to negate the possibility of creating a *TIF1*-RFP fusion that was not functionally equivalent to the wild type *TIF1*, as observed in a small percentage of genes in the GFP clone collection, a number of prevention measures were adopted. The RFP RedStar2 was selected specially in this research as it has an established linker region, before the start of the RFP coding sequence, of 10 residues of repeating glycine and alanine molecules. This addition gives confidence that the RFP will have a lower hindrance on the *TIF1* ORF than S65T GFP, used in the GFP clone collection.

The second prevention method can be interpreted as either an advantage or limitation. In order to limit the size of the foreign coding region inserted in the 3' UTR, the RFP was not associated with its own selection maker. Subsequently the presence of the RFP in the transformed strains could not be identified by growth under a given selection pressure but was identified visually by confocal microscopy. The initial advantage from this method may instantly become redundant if the RFP is not successfully maintained in future generations or homogenously across every cell in a strain. However this concern

was alleviated by the observation of satisfactory transmission of plasmid, and subsequent RFP expression, between parental and daughter generations.

3.4.2 Limitations

There are two major groupings by which the limitations encountered throughout this study can be classified; limitations arising from the methodologies employed in the generation of this fluorescence based system and limitations resulting from this system's application.

There are three limitations arising from the first category that, even though they significantly reduced the concluding power of this research, were all deemed expedient to the success of this initial study. Firstly, the retention of a third source of eIF4A in both dual fluorescent strains as a result of the retention of an untransformed *TIF2*-GFP strain. Secondly, the unknown disparities of how the architecture of plasmid and chromosomal copies of fluorescent proteins effects the production of eIF4A and finally the incorporation of an RFP that is not regulated by selection under a specific selection pressure. The aforementioned limitations have previously been discussed in section 3.2 and 3.4.1.2 respectively.

Additional limitations arose from issues with the screening method. Potentially the limitation with the broadest effect on this study is that, due to time constraints for this research, no repetition of any of the four screens was completed. From the data obtained in these screens there is a notable disparity between the observed levels of significance for an individual compound across multiple screens which may plausibly be attributed to the Individual fluorescence variation in a single well. Consequently a number of compounds considered to be hits that lie on the bubble of the arbitrary significance threshold employed in this study may in fact be artefacts. However, as discussed in section 3.3.1, the aforementioned threshold was intentionally set to be very stringent to limit the inclusion of false positives.

In addition to above, in some cases fluorescence intensities were deemed to be undetectable at time zero but detectable at the second time reading after 4 hours. This observation was a consequence of too few cells being delivered into a given well resulting in a failed reading by the image recognition software Acapella. As each strain replicated during the four hour incubation, during this time fluorescence intensities became

distinguishable. Subsequently this, and the observed disparities in individual variation, is likely to be rectified upon completion of a number of screen repetitions.

In terms of the chosen screening methodology a number of limitations arose due to lack of specific controls and counter screens. For example a control (or equivalent counter screen) with compound but no fluorescent protein is required in order to account for the significant changes observed in response to compounds which were hypothesised as having an inherent level of fluorescence. In addition, a dye swap experiment, where the fluorescent proteins assigned to each gene of the dual fluorescent reporters are swapped, is required in order to eliminate erroneously classified hits that are a result of variations in fluorescence intensities attributed to the fluorescent proteins themselves and not from compound interaction with *TIF1* and *TIF2*.

3.5 Conclusions

The aim of this study was to contribute to a wider programme into the *TIF1* and *TIF2* paralogs, specifically to design a system to reveal whether they could be individually regulated by small molecule intervention. To this end, a set of fluorescent reporter strains with plasmid-borne *TIF1*-RFP were generated. Although it was expected that ultimately the *TIF1*-RFP would have to be integrated into the genome, the development of a set of strains containing the *TIF1*-RFP on a CEN plasmid was chosen as a realistic target for this masters' project.

A *TIF1*-RFP construct containing both 5' and 3' UTRs was successfully generated. Transformants made with the construct were viable, stable, and expressed both red and green fluorescent proteins effectively. Extensive screening methodology development revealed that high-throughput confocal microscopy was more appropriate than high-throughput flow cytometry analysis for these strains, but also revealed limitations with the plasmid borne system. Most notable amongst these was the localisation of the RFP-labelled TIF1p, which was not as evenly distributed through the cytosol as the commercially available TIF1p-GFP. This may be due to high abundance of the protein, given that these plasmid-containing *TIF1*-RFP constructs retain the unlabelled genomic *TIF1* gene. Alternatively, TIF1p-RFP may be more accurately reflecting an impact of 3' UTR directing of localisation, since the commercial available TIF1p-GFP strain has its 3' UTR disrupted by selection markers.

A preliminary screen was undertaken with these reporter strains using the LOPAC library. Results from these screens signalled the potential of the dual fluorescent reporter system as a tool for the discovery of small molecule regulators of TIF1p or TIF2p. A number of important points emerge from this work. Firstly, it is possible to distinguish between expression from *TIF1* and *TIF2* in the same environment. Secondly, the creation of a universally applicable plasmid borne gene-RFP fusion was successful and could be more broadly applied to the study of duplicate genes. Finally, whilst there were elements of overlap between the screens, the more general outcome was a lack of coherence between the results from the screens.

Although not a specified goal of this project, a useful consequence of the approach was the ability to assess how well a plasmid containing a construct including the 5' and 3' UTRs mimicked the genomic copy of the same gene. As mentioned above, issues were noted with protein localisation of the TIF1p-RFP in some circumstances. More notably, the correlation between components of the dual fluorescent reporter, Tif1p-GFP TIF1p- was relatively low (for example, see Figure 3.14).

Chemoinformatic analysis of compounds causing large and significant apparent changes in the abundance of eIF4A from the reporter strains revealed that predominantly such changes arose from the chromophoric or fluorophoric nature of the compounds identified. Never-the-less a small number of compounds were identified that could be of interest in future study, although there were no structural elements in common between these compounds.

Overall, the study revealed the utility of the system in general terms, and provides a basis for the next step, which would be to create a reporter strain bearing a genomic *TIF1*-RFP construct, as well as a "dye-swapped" *TIF1*-GFP, *TIF2*-RFP control strain.

3.6 Future directions

There are a vast number of potential uses for the dual fluorescent system designed during the course of this study. Initially it would be wise to remedy the aforementioned limitations, by the introduction of more stringent controls, increasing the number of screen repetitions to at least three, and conducting validation experiments on those compounds identified as hits in the four screens, which due to time constraints of this research, could not be performed. However as this thesis has shown *prima-facie* evidence that the differential regulatory patterns of *TIF1* and *TIF2* can be assessed by a dual fluorescent reporter system, it is appealing to move to incorporating a genomically integrated copy of RFP and removal of the unlabelled *TIF* copy.

One potentially fruitful avenue arising from this study is an investigation into how the architecture of plasmid borne proteins differs from their genomically expressed counterparts. This could be achieved upon integration of RFP into the genome, as suggested above, and the comparison between this strain and the plasmid borne *TIF1*-RFP strain outlined in this thesis.

As previously noted, *TIF1* and *TIF2* have an almost identical sequence similarity in their coding region, barring six synonymous changes in the gene sequence, and vastly diverse sequences in both the 5' and 3' UTRs. Therefore it would be profitable to investigate how modifications in these regions effect the differential regulation of *TIF1* and *TIF2*. For example, to investigate the regulatory effect that codon bias, arising from the six synonymous changes has on the two genes ORF the current system can be amended by switching the promoter and terminator sequences of one of the *TIF* genes to mimic the other (by methodologies outlined in section 3.4.1.1). This process of region switching could be utilised to investigate the 3' and 5' UTR also.

This system is perfectly suited to validate published results that have been generated by indirect evidence, which have suggested certain transcription factors that may cause a change in the protein abundance of eIF4A. For example the interaction of the transcription factor *YAP1* with H₂O₂ has been suggested, on the freely available online resources YEASTRACT, to have an effect *TIF1* (Teixeira, Monteiro et al. 2006; Thorsen, Lagniel et al. 2007; Monteiro, Mendes et al. 2008; Abdulrehman, Monteiro et al. 2011).

However this was generated by inferences from microarray data and needs to be directly validated.

4 References

- Abdulrehman, D., P. T. Monteiro, et al. (2011). "YEASTRACT: providing a programmatic access to curated transcriptional regulatory associations in *Saccharomyces cerevisiae* through a web services interface." *Nucleic Acids Research* **39**(suppl 1): D136-D140.
- Altmann, M. and P. Linder (2010). "Power of Yeast for Analysis of Eukaryotic Translation Initiation." *Journal of Biological Chemistry* **285**(42): 31907-31912.
- Amberg, D., C. and D. Burke, J. (2005). *Methods in yeast genetics : a Cold Spring Harbour Laboratory course manual*, Cold Spring Harbour, N.Y., Cold Spring Harbour Laboratory Press.
- Ausubel, F. (1988). *Current protocols in molecular biology*, New York, Published by Greene Pub. Associates and Wiley Interscience : J Wiley.
- Bayliss, P. E., K. L. Bellavance, et al. (2006). "Chemical modulation of receptor signaling inhibits regenerative angiogenesis in adult zebrafish." *Nat Chem Biol* **2**(5): 265-273.
- Benz, J., H. Trachsel, et al. (1999). "Crystal structure of the ATPase domain of translation initiation factor 4A from *Saccharomyces cerevisiae* –the prototype of the DEAD box protein family." *Structure* **7**(6): 671-679.
- Berthelot, K., M. Muldoon, et al. (2004). "Dynamics and processivity of 40S ribosome scanning on mRNA in yeast." *Molecular Microbiology* **51**(4): 987-1001.
- Betterton, M. D. and F. Jülicher (2005). "Velocity and processivity of helicase unwinding of double-stranded nucleic acids." *Journal of Physics: Condensed Matter* **17**(47): S3851.
- Bircham, P. W., D. R. Maass, et al. (2011). "Secretory pathway genes assessed by high-throughput microscopy and synthetic genetic array analysis." *Molecular BioSystems* **7**(9): 2589-2598.
- Borenstein, E. and E. Ruppin (2006). "Direct evolution of genetic robustness in microRNA." *Proceedings of the National Academy of Sciences* **103**(17): 6593-6598.
- Brookfield, J. F. Y. (2003). "Gene Duplications: The Gradual Evolution of Functional Divergence." *Current Biology* **13**(6): R229-R230.
- Cliften, P. F., R. S. Fulton, et al. (2006). "After the Duplication: Gene Loss and Adaptation in *Saccharomyces* Genomes." *Genetics* **172**(2): 863-872.
- Conant, G. C. and K. H. Wolfe (2008). "Turning a hobby into a job: How duplicated genes find new functions." *Nat Rev Genet* **9**(12): 938-950.
- Cordin, O., J. Banroques, et al. (2006). "The DEAD-box protein family of RNA helicases." *Gene* **367**(0): 17-37.
- Davis, J. C. and D. A. Petrov (2005). "Do disparate mechanisms of duplication add similar genes to the genome?" *Trends in Genetics* **21**(10): 548-551.
- de la Cruz, J., D. Kressler, et al. (1999). "Unwinding RNA in *Saccharomyces cerevisiae*: DEAD-box proteins and related families." *Trends in Biochemical Sciences* **24**(5): 192-198.
- DeLuna, A., M. Springer, et al. (2010). "Need-Based Up-Regulation of Protein Levels in Response to Deletion of Their Duplicate Genes." *PLoS Biol* **8**(3): e1000347.
- Dominguez, D., E. Kislig, et al. (2001). "Structural and functional similarities between the central eukaryotic initiation factor (eIF)4A-binding domain of mammalian eIF4G and the eIF4A-binding domain of yeast eIF4G." *Biochemical Journal* **355**: 223-230.
- Dunnett, C. W. (1955). "A Multiple Comparison Procedure for Comparing Several Treatments with a Control." *Journal of the American Statistical Association* **50**(272): 1096-1121.
- Fairman-Williams, M. E., U.-P. Guenther, et al. (2010). "SF1 and SF2 helicases: family matters." *Current Opinion in Structural Biology* **20**(3): 313-324.
- Farooq, M., J. Choi, et al. (2012). "Identification of 3'UTR sequence elements and a teloplasm localization motif sufficient for the localization of Hro-twist mRNA to the zygotic animal and vegetal poles." *Development, Growth & Differentiation* **54**(4): 519-534.
- Gershon, H. and D. Gershon (2000). "The budding yeast, *Saccharomyces cerevisiae*, as a model for aging research: a critical review." *Mechanisms of Ageing and Development* **120**(1-3): 1-22.

- Gietz, R. D. and R. H. Schiestl (2007). "High-efficiency yeast transformation using the LiAc/SS carrier DNA/PEG method." Nat. Protocols **2**(1): 31-34.
- Gorbalenya, A. E. and E. V. Koonin (1993). "Helicases: amino acid sequence comparisons and structure-function relationships." Current Opinion in Structural Biology **3**(3): 419-429.
- Guan, Y., M. J. Dunham, et al. (2007). "Functional Analysis of Gene Duplications in *Saccharomyces cerevisiae*." Genetics **175**(2): 933-943.
- Heikal, A. A., S. T. Hess, et al. (2000). "Molecular spectroscopy and dynamics of intrinsically fluorescent proteins: Coral red (dsRed) and yellow (Citrine)." Proceedings of the National Academy of Sciences **97**(22): 11996-12001.
- Hernandez, M., G. Recio, et al. (2012). "Surface enhanced fluorescence of anti -tumoral drug emodin adsorbed on silver nanoparticles and loaded on porous silicon." Nanoscale Research Letters **7**(1): 364.
- Huh, W.-K., J. V. Falvo, et al. (2003). "Global analysis of protein localization in budding yeast." Nature **425**(6959): 686-691.
- Janke, C., M. M. Magiera, et al. (2004). "A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes." Yeast **21**(11): 947-962.
- Jankowsky, A., U.-P. Guenther, et al. (2011). "The RNA helicase database." Nucleic Acids Research **39**(suppl 1): D338-D341.
- Kafri, R., M. Levy, et al. (2006). "The regulatory utilization of genetic redundancy through responsive backup circuits." Proceedings of the National Academy of Sciences **103**(31): 11653-11658.
- Kellis, M., B. W. Birren, et al. (2004). "Proof and evolutionary analysis of ancient genome duplication in the yeast *Saccharomyces cerevisiae*." Nature **428**(6983): 617-624.
- Komili, S., N. G. Farny, et al. (2007). "Functional Specificity among Ribosomal Proteins Regulates Gene Expression." Cell **131**(3): 557-571.
- Kozak, M. (1999). "Initiation of translation in prokaryotes and eukaryotes." Gene **234**(2): 187-208.
- Kozak, M. and A. J. Shatkin (1978). "Migration of 40 S ribosomal subunits on messenger RNA in the presence of edeine." Journal of Biological Chemistry **253**(18): 6568-6577.
- Kronja, I. and T. L. Orr-Weaver (2011). "Translational regulation of the cell cycle: when, where, how and why?" Philosophical Transactions of the Royal Society B: Biological Sciences **366**(1584): 3638-3652.
- Lackner, D. H. and J. Bähler (2008). Chapter 5 Translational Control of Gene Expression: From Transcripts to Transcriptomes. International Review of Cell and Molecular Biology. W. J. Kwang, Academic Press. **Volume 271**: 199-251.
- Lanker, S., P. P. Müller, et al. (1992). "Interactions of the eIF-4F subunits in the yeast *Saccharomyces cerevisiae*." Journal of Biological Chemistry **267**(29): 21167-21171.
- Lee, C.-P. (1971). "Fluorescent probe of the hydrogen ion concentration in ethylenediaminetetraacetic acid particles of beef heart mitochondria." Biochemistry **10**(24): 4375-4381.
- Linder, P. (2006). "Dead-box proteins: a family affair - active and passive players in RNP-remodeling." Nucleic Acids Research **34**(15): 4168-4180.
- Linder, P. and E. Jankowsky (2011). "From unwinding to clamping — the DEAD box RNA helicase family." Nat Rev Mol Cell Biol **12**(8): 505-516.
- Linder, P. and P. P. Slonimski (1988). "Sequence of the genes TIF1 and TIF2 from *Saccharomyces cerevisiae* coding for a translation initiation factor." Nucl. Acids Res. **16**(21): 10359-.
- Linder, P. and P. P. Slonimski (1989). "An essential yeast protein, encoded by duplicated genes TIF1 and TIF2 and homologous to the mammalian translation initiation factor eIF-4A, can suppress a mitochondrial missense mutation." Proceedings of the National Academy of Sciences of the United States of America **86**(7): 2286-2290.

- Louis, E. J. (2007). "Evolutionary genetics: Making the most of redundancy." Nature **449**(7163): 673-674.
- Marques, A., N. Vinckenbosch, et al. (2008). "Functional diversification of duplicate genes through subcellular adaptation of encoded proteins." Genome Biology **9**(3): R54.
- Mazan-Mamczarz, K., A. Lal, et al. (2006). "Translational Repression by RNA-Binding Protein TIAR." Mol. Cell. Biol. **26**(7): 2716-2727.
- Mendez, R. and J. D. Richter (2001). "Translational control by CPEB: a means to the end." Nat Rev Mol Cell Biol **2**(7): 521-529.
- Meyer, A. (2003). "Duplication duplication." Nature **421**: 31-32.
- Monteiro, P. T., N. D. Mendes, et al. (2008). "YEASTRACT-DISCOVERER: new tools to improve the analysis of transcriptional regulatory associations in *Saccharomyces cerevisiae*." Nucleic Acids Research **36**(suppl 1): D132-D136.
- Musso, G., M. Costanzo, et al. (2008). "The extensive and condition-dependent nature of epistasis among whole-genome duplicates in yeast." Genome Research **18**(7): 1092-1099.
- Nuutinen, U., A. Ropponen, et al. (2009). "Dexamethasone-induced apoptosis and up-regulation of Bim is dependent on glycogen synthase kinase-3." Leukemia Research **33**(12): 1714-1717.
- Parkesh, R., T. Clive Lee, et al. (2007). "Highly selective 4-amino-1,8-naphthalimide based fluorescent photoinduced electron transfer (PET) chemosensors for Zn(ii) under physiological pH conditions." Organic & Biomolecular Chemistry **5**(2): 310-317.
- Patterson, G. H., S. M. Knobel, et al. (1997). "Use of the green fluorescent protein and its mutants in quantitative fluorescence microscopy." Biophysical Journal **73**(5): 2782-2790.
- Pearson, P. L., M. Bobrow, et al. (1971). "Quinacrine Fluorescence in Mammalian Chromosomes." Nature **231**(5301): 326-329.
- Pecere, T., M. V. Gazzola, et al. (2000). "Aloe-emodin Is a New Type of Anticancer Agent with Selective Activity against Neuroectodermal Tumors." Cancer Research **60**(11): 2800-2804.
- Reineke, L. C. and W. C. Merrick (2009). "Characterization of the functional role of nucleotides within the URE2 IRES element and the requirements for eIF2A-mediated repression." RNA **15**(12): 2264-2277.
- Rocak, S. and P. Linder (2004). "DEAD-box proteins: the driving forces behind RNA metabolism." Nat Rev Mol Cell Biol **5**(3): 232-241.
- Rogers, G. W., A. A. Komar, et al. (2002). eIF4A: The godfather of the DEAD box helicases. Progress in Nucleic Acid Research and Molecular Biology, Academic Press. **Volume 72**: 307-331.
- Rogers, G. W., N. J. Richter, et al. (2001). "Modulation of the Helicase Activity of eIF4A by eIF4B, eIF4H, and eIF4F." Journal of Biological Chemistry **276**(33): 30914-30922.
- Rotenberg, M. O., M. Moritz, et al. (1988). "Depletion of *Saccharomyces cerevisiae* ribosomal protein L16 causes a decrease in 60S ribosomal subunits and formation of half-mer polyribosomes." Genes & Development **2**(2): 160-172.
- Schutz, P., M. Bumann, et al. (2008). "Crystal structure of the yeast eIF4A-eIF4G complex: An RNA-helicase controlled by protein-protein interactions." Proceedings of the National Academy of Sciences **105**(28): 9564-9569.
- SGD. (2012). *Saccharomyces Genome Database*. <http://www.yeastgenome.org>.
- Sikorski, R. S. and P. Hieter (1989). "A system of shuttle vectors and yeast host strains designed for efficient manipulation of DNA in *Saccharomyces cerevisiae*." Genetics **122**(1): 19-27.
- Sonenberg, N. and A. G. Hinnebusch (2009). "Regulation of Translation Initiation in Eukaryotes: Mechanisms and Biological Targets." Cell **136**(4): 731-745.
- Teixeira, M. C., P. Monteiro, et al. (2006). "The YEASTRACT database: a tool for the analysis of transcription regulatory associations in *Saccharomyces cerevisiae*." Nucleic Acids Research **34**(suppl 1): D446-D451.

- Thorsen, M., G. Lagniel, et al. (2007). "Quantitative transcriptome, proteome, and sulfur metabolite profiling of the *Saccharomyces cerevisiae* response to arsenite." Physiological Genomics **30**(1): 35-43.
- Tong, A. H. and C. Boone (2006). "Synthetic Genetic Array Analysis in *Saccharomyces cerevisiae*." Methods in Molecular Biology **313**: 171-191.
- Tong, A. H. Y., M. Evangelista, et al. (2001). "Systematic Genetic Analysis with Ordered Arrays of Yeast Deletion Mutants." Science **294**(5550): 2364-2368.
- Wilcoxon, F. (1945). "Individual Comparisons by Ranking Methods." Biometrics Bulletin **1**(6): 80-83.
- Wilkie, G. S., K. S. Dickson, et al. (2003). "Regulation of mRNA translation by 5'- and 3'-UTR-binding factors." Trends in Biochemical Sciences **28**(4): 182-188.
- Winzeler, E. A., D. D. Shoemaker, et al. (1999). "Functional Characterization of the *S. cerevisiae* Genome by Gene Deletion and Parallel Analysis." Science **285**(5429): 901-906.
- Wolfe, K. H. and D. C. Shields (1997). "Molecular evidence for an ancient duplication of the entire yeast genome." Nature **387**(6634): 708-713.
- Zhang, J. (2004). "The infancy of duplicate genes." Heredity **92**: 479-480.
- Zhou, W., G. M. Edelman, et al. (2001). "Transcript leader regions of two *Saccharomyces cerevisiae* mRNAs contain internal ribosome entry sites that function in living cells." Proceedings of the National Academy of Sciences **98**(4): 1531-1536.

5 Appendix A – LOPAC Library HTS time zero readings

As indicated in section 2.2.3.2, readings from the LOPAC library HTS screens were taken at time zero and after a four hour incubation period. This methodology was exercised in an attempt to enrich the list of “hits” to only include compounds eliciting a response resulting from gene regulation. As the initial time point for imaging is almost instantaneously after the introduction of the LOPAC library it is not conceivable that any observed changes in protein expression were regulated via changes in mRNA expression. Instead, it is plausible that any such changes are instigated by disturbances in the GFP signal or arise from intrinsic fluorescence in the compound library and not as a consequence of gene expression changes in response to the introduction of the LOPAC library. Hence any “hits” observed at the initial time point are considered a crucial control in eliminating those compounds that are not pertinent to the primary aim of this study; elucidating differential regulation between *TIF1* and *TIF2*.

5.1 Single GFP screening

5.1.1 TIF1p-GFP screen

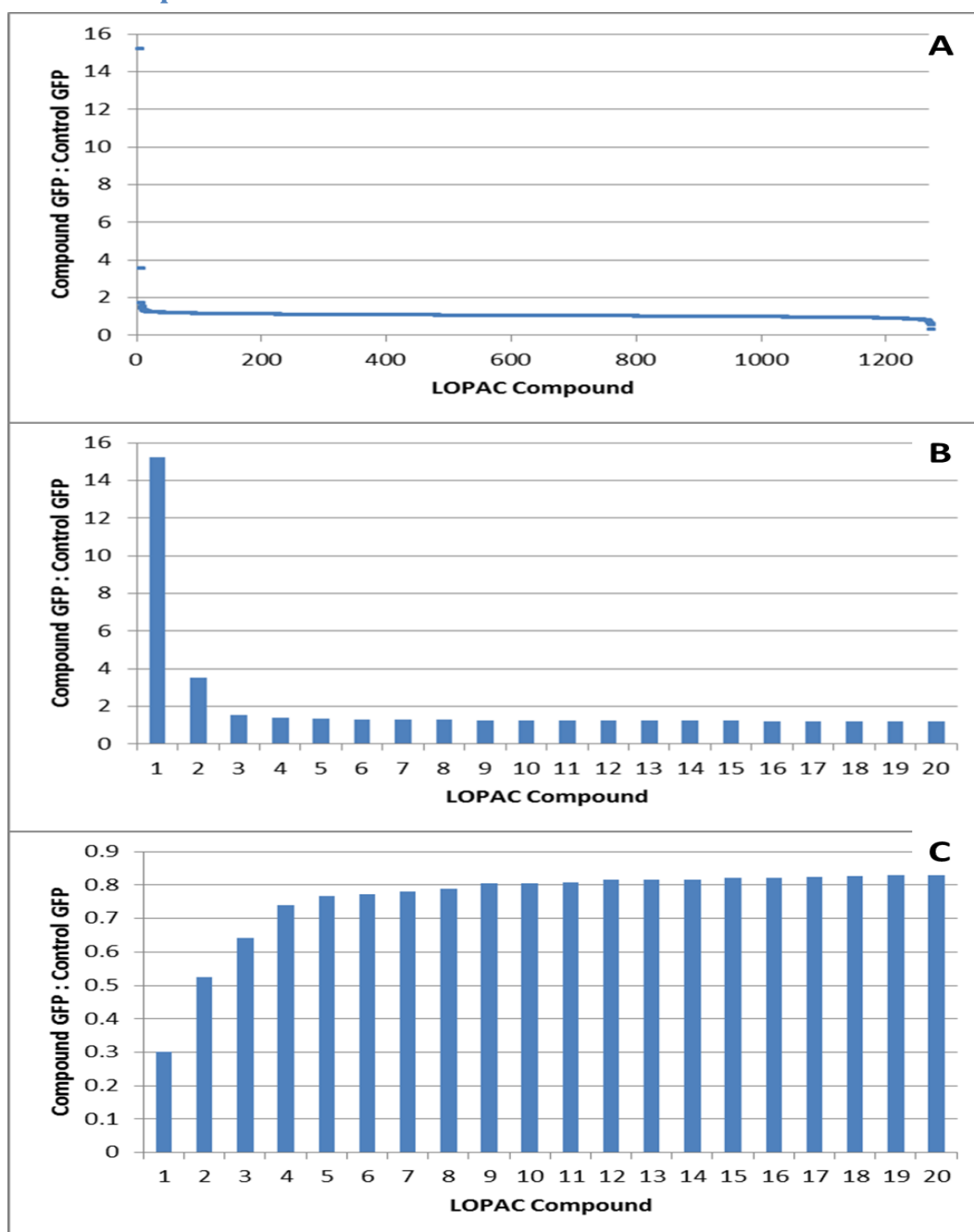


Figure 5.1: Whole cell fluorescence intensities of TIF1p-GFP treated with the LOPAC library (compounds 1 – 1280) compared against the pooled collection of whole cell fluorescence intensities of TIF1p-GFP treated with DMSO at t=0 (as described in section 2.2.3.4). (A) Compound-treated GFP: control GFP for the entire LOPAC library ranging from highest to lowest. The top 20 compounds that **(B)** elicited an increase in or **(C)** elicited a decrease in the ratio of compound-treated GFP compared to control. Compounds highlighted in **B** and **C** are presented in Table 5.1.

#	Compound	Compound GFP : Control GFP	Q value	Compound	Compound GFP : Control GFP	Q value
1	SB 216763	15.22	4.14E-06	Myricetin	0.3012	1.424E-25
2	SU 5416	3.537	6.12E-19	Mianserin hydrochloride	0.5254	3.247E-13
3	Carcinine dihydrochloride	1.52	0.005341	Arecaidine propargyl ester hydrobromide	0.6415	7.266E-08
4	SB 204070 hydrochloride	1.408	0.000583	Rp-cAMPS triethylamine	0.7396	0.0002334
5	Tyrphostin AG 34	1.32	9.19E-10	Orphenadrine hydrochloride	0.767	9.374E-08
6	(6R)-5,6,7,8-Tetrahydro-L-biopterin hydrochloride	1.301	1.24E-09	7-Cyclopentyl-5-(4-phenoxy)phenyl-7H-pyrrolo[2,3-d]pyrimidin-4-ylamine	0.773	0.0002089
7	1-(4-Chlorobenzyl)-5-methoxy-2-methylindole-3-acetic acid	1.287	0.00227	Caffeic acid phenethyl ester	0.7817	3.117E-06
8	6,7-ADTN hydrobromide	1.268	0.007032	Sulindac	0.7884	1.057E-05
9	Taurine	1.261	0.000281	Mesulergine hydrochloride	0.8048	0.0001584
10	13-cis-retinoic acid	1.235	2.53E-07	Eliprodil	0.8049	7.082E-06
11	R(-)-SCH-12679 maleate	1.233	0.000108	MK-886	0.807	0.001026
12	Tyrphostin AG 112	1.224	7.64E-07	Phorbol 12-myristate 13-acetate	0.8146	4.798E-12
13	Reactive Blue 2	1.224	0.000804	XK469	0.8154	0.00822
14	Spermidine trihydrochloride	1.221	0.000151	Nimustine hydrochloride	0.8164	2.921E-08
15	Ritodrine hydrochloride	1.217	0.001494	N6-Cyclohexyladenosine	0.8208	1.967E-05
16	4-Hydroxyphenethylamine hydrochloride	1.213	0.000352	5,7-Dichlorokynurenic acid	0.8212	0.002928
17	Ritanserin	1.209	0.003655	cis-(Z)-Flupenthixol dihydrochloride	0.8231	1.145E-14
18	Cephalothin sodium	1.206	0.005308	S(-)-DS 121 hydrochloride	0.8279	0.0003135
19	Ro 41-0960	1.203	0.007441	Hydroxyta crane maleate	0.8287	3.751E-07
20	NG-Nitro-L-arginine	1.201	0.00091	Ergocristine	0.8287	2.619E-10

Table 5.1: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF1p-GFP screen at t=0 (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in sections 2.2.3.4 and 3.3.1).

5.1.2 TIF2p-GFP screen

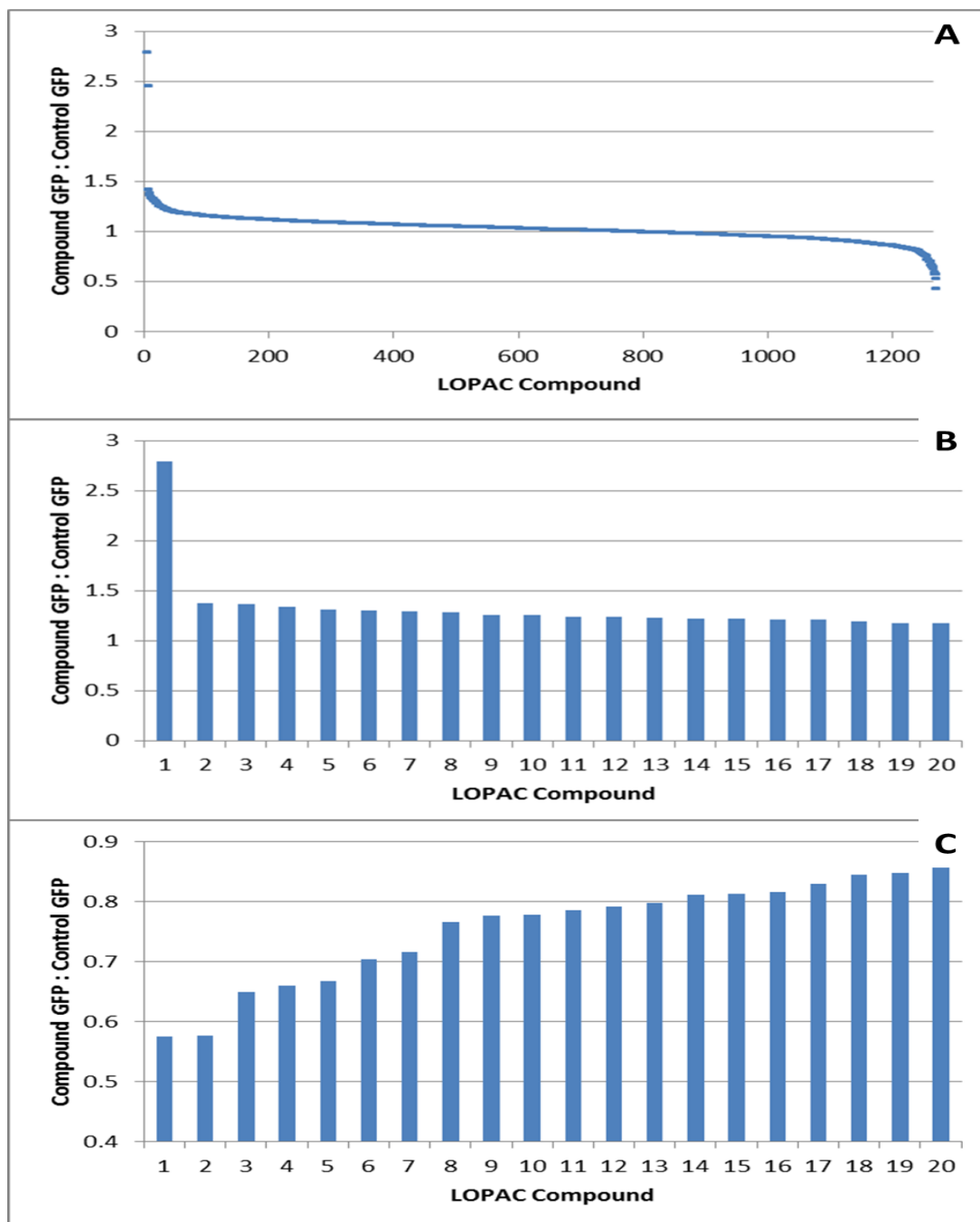


Figure 5.2: Whole cell fluorescence intensities of TIF2p-GFP treated with the LOPAC library (compounds 1 – 1280) compared against the pooled collection of whole cell fluorescence intensities of TIF2p-GFP treated with DMSO at t=0 (as described in section 2.2.3.4). (A) Compound-treated GFP: control GFP for the entire LOPAC library ranging from highest to lowest. The top 20 compounds that (B) elicited an increase in or (C) elicited a decrease in the ratio of compound-treated GFP compared to control. Compounds highlighted in B and C are presented in Table 5.2.

#	Compound	Compound GFP : Control GFP	Q value	Compound	Compound GFP : Control GFP	Q value
1	SU 5416	2.792	6.18E-06	CGP-7930	0.5756	2.946E-06
2	Naltrindole hydrochloride	1.378	0.000468	Imipramine hydrochloride	0.5758	1.997E-14
3	NAN-190 hydrobromide	1.365	2.62E-06	Dihydroergotamine methanesulfonate	0.6488	1.596E-07
4	Methiothepin mesylate	1.336	0.00098	Arecaidine propargyl ester hydrobromide	0.6602	0.001959
5	T-0156	1.312	4.05E-05	Ro 16-6491 hydrochloride	0.6667	2.587E-13
6	Nialamide	1.302	0.000753	Fluspirilene	0.7037	4.846E-10
7	(±)- Normetanephrine hydrochloride	1.297	0.000838	8-(p- Sulfophenyl)theophylline	0.7163	0.002497
8	Spermidine trihydrochloride	1.283	3.13E-05	Tiapride hydrochloride	0.7659	0.0002051
9	Neostigmine bromide	1.261	0.009829	Ergocristine	0.7754	2.706E-14
10	Mibefradil dihydrochloride	1.259	0.000581	PAPP	0.7773	0.0003945
11	Taurine	1.242	0.003962	Nitrendipine	0.7858	0.000227
12	Zonisamide sodium	1.241	0.008218	Felodipine	0.7916	3.274E-06
13	REV 5901	1.229	0.002398	Serotonin hydrochloride	0.7969	0.0003
14	SR 59230A oxalate	1.226	0.000629	(+)-Hydrastine	0.8106	0.0003053
15	Tyrphostin AG 555	1.221	1.32E-06	7-Cyclopentyl-5-(4- phenoxy)phenyl-7H- pyrrolo[2,3-d]pyrimidin- 4-ylamine	0.8126	0.0000451
16	Quazinone	1.212	0.002161	(±)-gamma-Vinyl GABA	0.8148	2.523E-08
17	4-Amino-1,8- naphthalimide	1.21	8.14E-05	JWH-015	0.8284	0.002323
18	Tetrahydrozoline hydrochloride	1.197	0.004263	1- Aminocyclopropanecarboxylic acid hydrochloride	0.844	4.398E-06
19	Xamoterol hemifumarate	1.18	0.000718	(±)-p- Chlorophenylalanine	0.8478	0.0001121
20	AC 915 oxalate	1.173	0.008797	Estrone	0.8604	0.0004657

Table 5.2: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF2p-GFP screen at t=0 (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in sections 2.2.3.4 and 3.3.1).

5.2 Dual Fluorescent GFP/RFP screening

5.2.1 TIF1p-GFP TIF1p-RFP screen

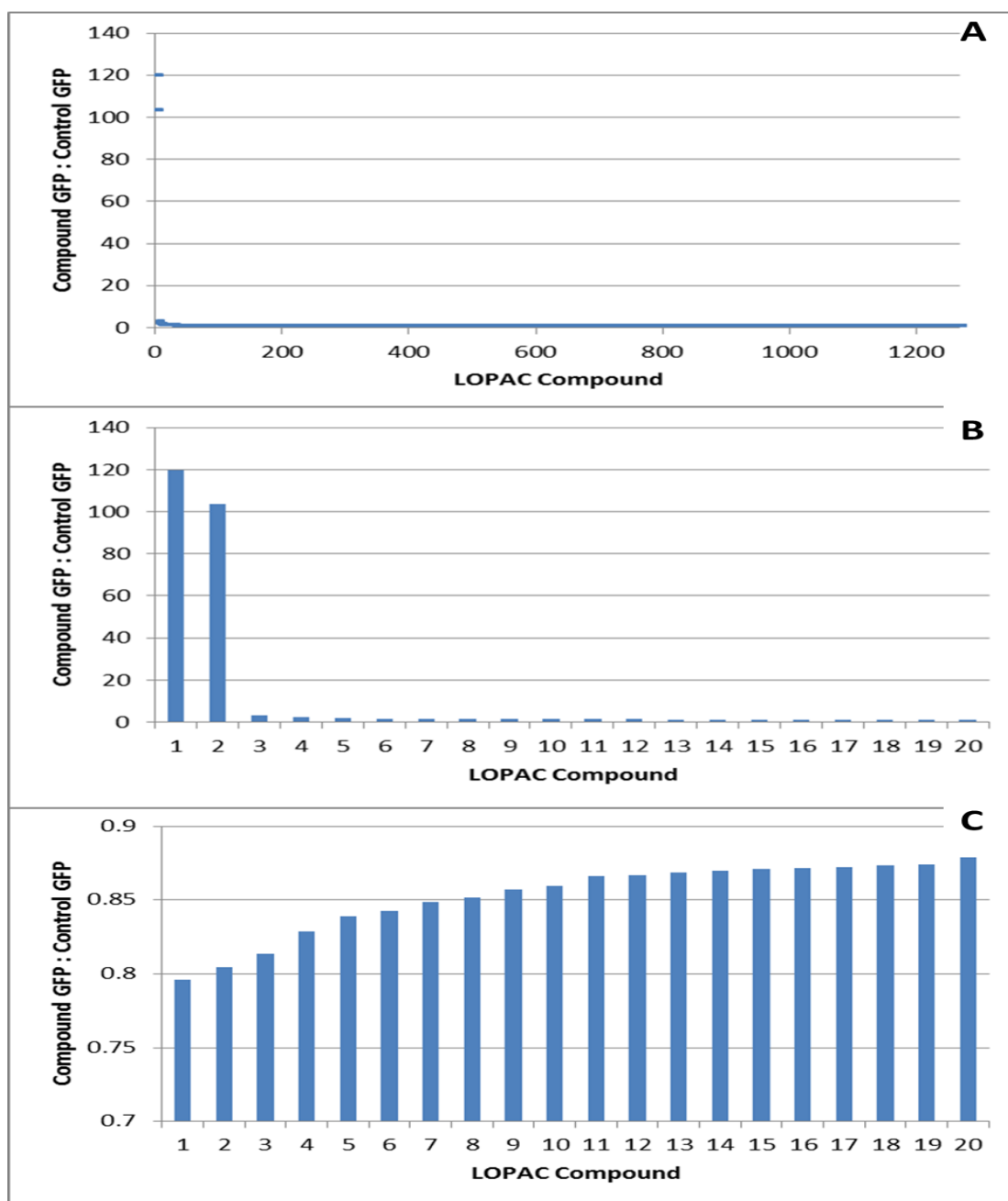


Figure 5.3: Whole cell fluorescence intensities of the GFP component of TIF1p-GFP TIF1p-RFP treated with the LOPAC library (compounds 1 – 1280) compared against the pooled collection of whole cell fluorescence intensities of TIF1p-GFP TIF1p-RFP treated with DMSO at t=0 (as described in section 2.2.3.4). (A) Compound-treated GFP: control GFP for the entire LOPAC library ranging from highest to lowest. The top 20 compounds that (B) elicited an increase in or (C) elicited a decrease in the ratio of compound-treated GFP compared to control. Compounds highlighted in B and C are presented in Table 5.3.

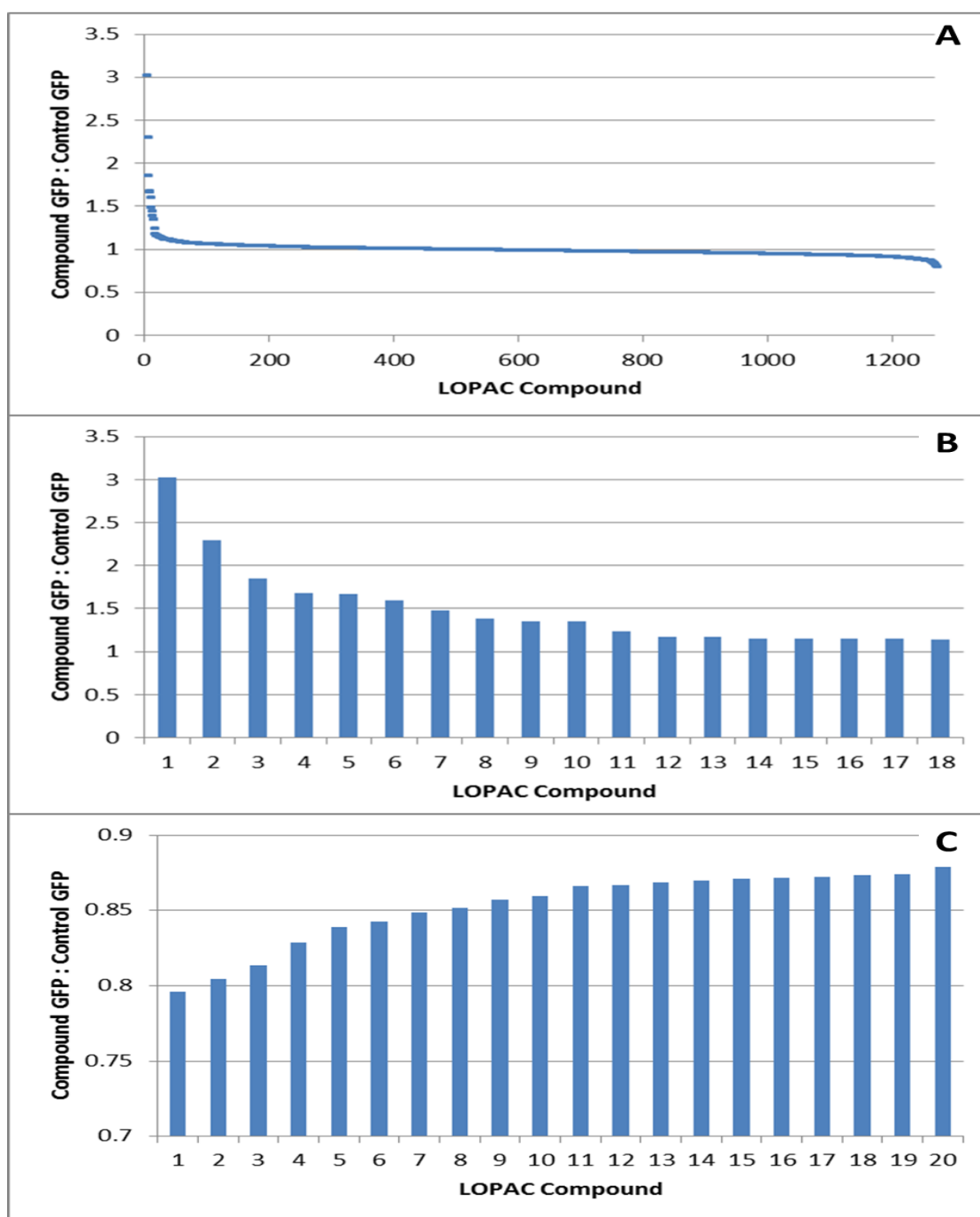


Figure 5.4: Whole cell fluorescence intensities of the GFP component of TIF1p-GFP TIF1p-RFP treated with the LOPAC library (compounds 1 – 1280) compared against the pooled collection of whole cell fluorescence intensities of TIF1p-GFP TIF1p-RFP treated with DMSO at t=0 (as described in section 2.2.3.4). This figure is the same data set as Figure 5.3 but with a restricted range as SB 216763, SU 5416, were removed from this data set. **(A)** Compound-treated GFP: control GFP for the entire LOPAC library ranging from highest to lowest. The top responding compounds that **(B)** elicited an increase in or **(C)** elicited a decrease in the ratio of compound-treated GFP compared to control. Compounds highlighted in **B** and **C** are presented in Table 5.3.

#	Compound	Compound GFP : Control GFP	Q value	Compound	Compound GFP : Control GFP	Q value
1	SU 5416	119.7	0.000385	HA-100	0.7959	6.682E-18
2	SB 216763	103.4	7.98E-05	MHPG piperazine	0.8045	3.18E-07
3	Sodium Oxamate	3.023	1.04E-85	p-MPPF dihydrochloride	0.8136	1.894E-21
4	GW2974	2.298	5.09E-39	GR 125487 sulfamate salt	0.8286	3.403E-09
5	Idarubicin	1.855	1E-10	CR 2249	0.8391	1.155E-08
6	Quinacrine dihydrochloride	1.675	6.7E-43	Na lidixic acid sodium	0.8425	1.989E-09
7	U0126	1.666	0.000237	NCS-382	0.8486	0.0006286
8	Glibenclamide	1.596	0.000935	9-cyclopentyladenine	0.8516	2.973E-36
9	L-765,314	1.483	6.2E-31	erythro-9-(2-Hydroxy-3-nonyl)adenine hydrochloride	0.8569	4.512E-25
10	GYKI 52466 hydrochloride	1.387	1.18E-06	CL 316,243	0.8594	6.188E-54
11	Ethosuximide	1.351	0.006546	R(+)-SCH-23390 hydrochloride	0.8659	1.223E-23
12	NF449 octasodium salt	1.348	3.9E-05	1-(4-Hydroxybenzyl)imidazole-2-thiol	0.8669	9.788E-07
13	Ellipticine	1.237	7.46E-16	R(+)-UH-301 hydrochloride	0.8687	8.877E-24
14	JWH-015	1.176	0.000109	beta-Estradiol	0.8696	9.423E-16
15	3-Hydroxybenzylhydrazine dihydrochloride	1.168	8.75E-07	(±)-7-Hydroxy-DPAT hydrobromide	0.8713	0.003167
16	Sanguinarine chloride	1.155	6.94E-24	HA-1004 hydrochloride	0.8715	0.0008731
17	Emodin	1.153	8.69E-28	Clofibrate	0.8724	1.829E-35
18	SU 4312	1.151	8.5E-22	2-Chloro-2-deoxy-D-glucose	0.8734	4.585E-22
19	3',4'-Dichlorobenzamil	1.149	3.52E-09	A-77636 hydrochloride	0.874	1.948E-06
20	Dubinidine	1.139	2.53E-14	O6-benzylguanine	0.8787	9.128E-10

Table 5.3: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF1p-GFP TIF1p-RFP screen at t=0 (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in sections 2.2.3.4 and 3.3.1).

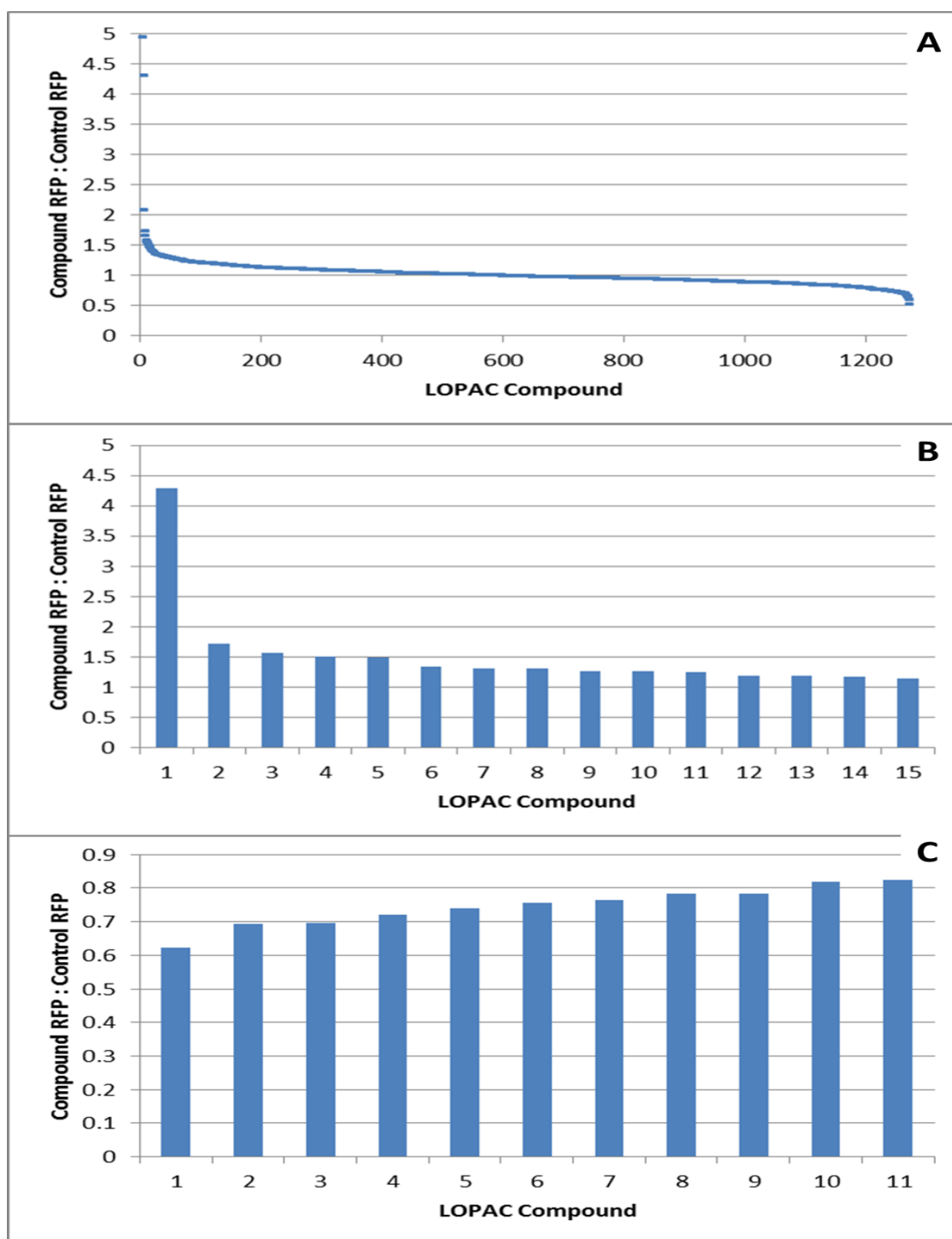


Figure 5.5: Whole cell fluorescence intensities of the RFP component of TIF1p-GFP TIF1p-RFP treated with the LOPAC library (compounds 1 – 1280) compared against the pooled collection of whole cell fluorescence intensities of TIF1p-GFP TIF1p-RFP treated with DMSO at t=0 (as described in section 2.2.3.4). (A) Compound-treated RFP: control RFP for the entire LOPAC library ranging from highest to lowest. The top responding compounds that (B) elicited an increase in or (C) elicited a decrease in the ratio of compound-treated RFP compared to control. Compounds highlighted in B and C are presented in Table 5.4.

#	Compound	Compound RFP : Control RFP	Q value	Compound	Compound RFP : Control RFP	Q value
1	SB 216763	4.298	0.000494	A-77636 hydrochloride	0.6235	0.006794
2	Metrifudil	1.726	0.002057	Phenamil methanesulfonate	0.6946	0.002572
3	(±)-Synephrine	1.568	2.21E-07	Diclofenac sodium	0.6967	0.003666
4	Salmeterol xinafoate	1.509	1.73E-06	9- cyclopentyladenine	0.7202	4.213E-05
5	Tulobuterol hydrochloride	1.498	5.78E-05	CGS-15943	0.7392	5.599E-05
6	Cyclophosphamide monohydrate	1.348	0.000318	SB 204741	0.7549	0.005417
7	B-HT 933 dihydrochloride	1.316	0.000393	(-)-Cotinine	0.765	0.0008432
8	Cortisone	1.315	3.58E-07	Oxymetazoline hydrochloride	0.7822	0.0004806
9	S-5-Iodowillardiine	1.269	0.000972	Pimozide	0.7843	0.002197
10	Mevastatin	1.26	0.000167	Cephalexin hydrate	0.8174	0.0006399
11	Bepidil hydrochloride	1.252	0.000934	Clemizole hydrochloride	0.8247	0.008955
12	(+)- Brompheniramine maleate	1.193	6.46E-05			
13	SIB 1893	1.185	0.009565			
14	Chlorpropamide	1.172	0.000175			
15	CNS-1102	1.152	0.002102			

Table 5.4: The top compounds that elicited an increase or a decrease in the observed ratio of compound-treated RFP compared to control RFP from the TIF1p-GFP TIF1p-RFP screen at t=0 (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in sections 2.2.3.4 and 3.3.1).

5.2.2 TIF2p-GFP TIF1p-RFP screen

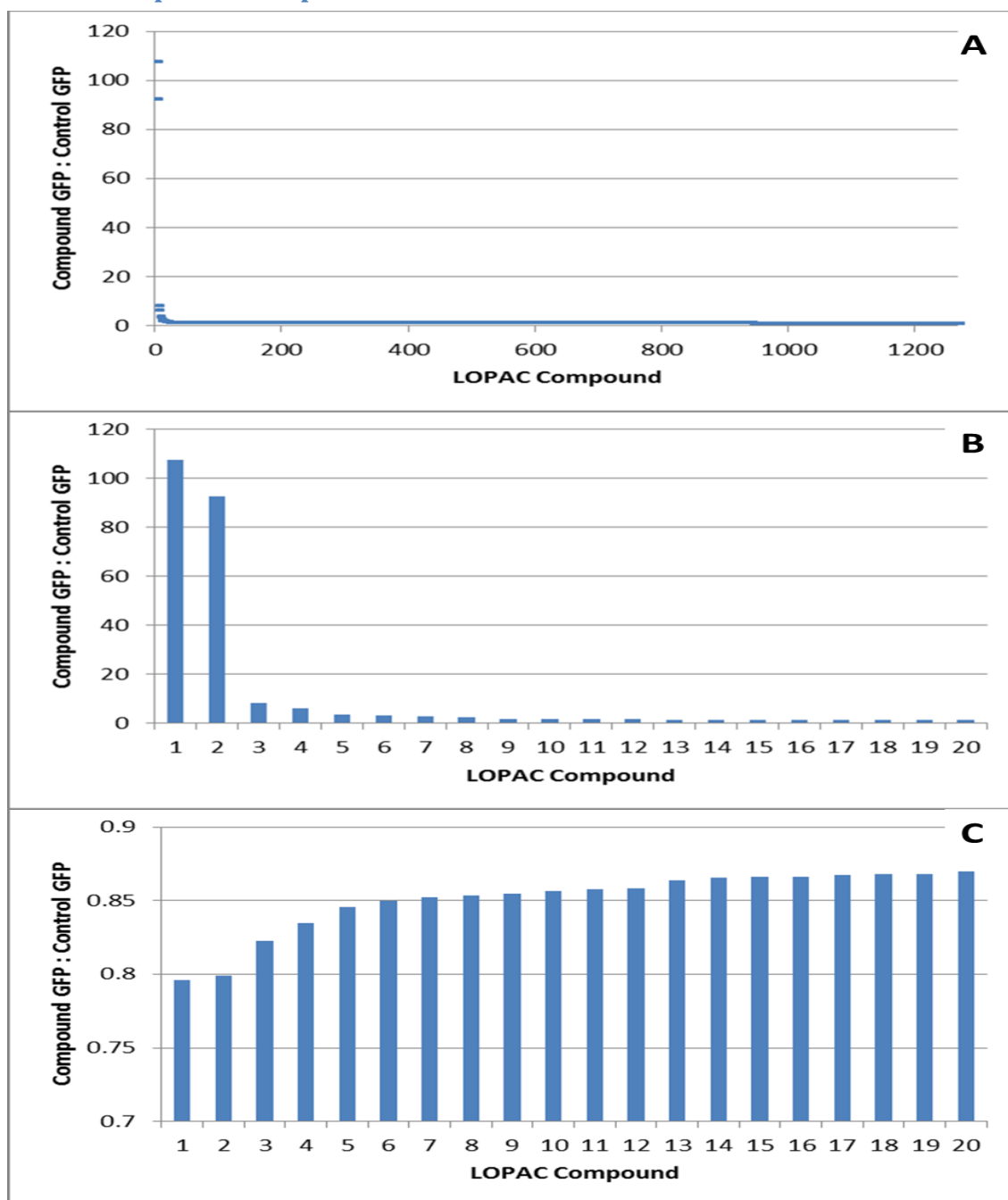


Figure 5.6: Whole cell fluorescence intensities of the GFP component of TIF2p-GFP TIF1p-RFP treated with the LOPAC library (compounds 1 – 1280) compared against the pooled collection of whole cell fluorescence intensities of TIF2p-GFP TIF1p-RFP treated with DMSO at t=0 (as described in section 2.2.3.4). (A) Compound-treated GFP: control GFP for the entire LOPAC library ranging from highest to lowest. The top 20 compounds that (B) elicited an increase in or (C) elicited a decrease in the ratio of compound-treated GFP compared to control. Compounds highlighted in B and C are presented in Table 5.5.

#	Compound	Compound GFP : Control GFP	Q value	Compound	Compound GFP : Control GFP	Q value
1	SU 5416	107.5	8.1E-05	Nimustine hydrochloride	0.7959	0.001991
2	SB 216763	92.35	8.18E-07	8-Bromo-cGMP sodium	0.7993	1.932E-15
3	Sanguinarine chloride	8.212	1.93E-10	Cefsulodin sodium salt hydrate	0.8224	1.799E-05
4	4-Amino-1,8-naphthalimide	6.065	9.58E-07	Rp-cAMPS triethylamine	0.8346	8.788E-08
5	6,7-ADTN hydrobromide	3.508	3.5E-07	Chlorambucil	0.8456	4.263E-12
6	Kenpaulone	3.284	0.000892	A-77636 hydrochloride	0.85	1.503E-08
7	Nylidrin hydrochloride	2.593	2.09E-08	N,N-Dihexyl-2-(4-fluorophenyl)indole-3-acetamide	0.8523	3.417E-10
8	Idarubicin	2.357	1.06E-06	Cephapirin sodium	0.8533	0.009893
9	U0126	1.842	5E-06	Tyrphostin 51	0.8549	1.281E-11
10	Centrophoxine hydrochloride	1.727	5.81E-06	H-8 dihydrochloride	0.8568	0.002193
11	Emodin	1.724	2.54E-80	beta-Chloro-L-alanine hydrochloride	0.8577	4.586E-06
12	Papaverine hydrochloride	1.543	1.8E-10	NS 521 oxalate	0.8584	0.009567
13	Glibenclamide	1.413	8.23E-06	Minoxidil	0.8638	3.425E-12
14	N-(p-Isothiocyanatophenethyl)piperone hydrochloride	1.394	0.000143	CB 1954	0.8654	0.0007087
15	LY-367,265	1.356	3.56E-10	L-Tryptophan	0.866	1.619E-13
16	U-74389G maleate	1.353	1.63E-10	Bromoacetylcholine bromide	0.8663	0.00002
17	1,3-Dipropyl-7-methylxanthine	1.313	9.13E-05	Methapyrilene hydrochloride	0.8674	1.715E-05
18	SU 4312	1.311	4.93E-13	L-Carnavanine sulfate	0.8681	0.002367
19	GW5074	1.295	2E-12	HA-1004 hydrochloride	0.8681	7.045E-05
20	Tyrphostin 47	1.237	5.04E-14	Oxymetazoline hydrochloride	0.8697	4.962E-06

Table 5.5: The top 20 compounds that elicited an increase or a decrease in the observed ratio of compound-treated GFP compared to control GFP from the TIF2p-GFP TIF1p-RFP screen at t=0 (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in sections 2.2.3.4 and 3.3.1).

#	Compound	Compound RFP : Control RFP	Q value		Compound	Compound RFP : Control RFP	Q value
1	SB 216763	11.87	7.1E-06		5,7-Dichlorokynurenic acid	0.4047	0.009752
2	3-Amino-1-propanesulfonic acid sodium	2.792	1.99E-13		DL-threo-beta-hydroxyaspartic acid	0.4762	0.0008282
3	4-Methylpyrazole hydrochloride	1.75	0.001367		Daidzein	0.4804	0.003666
4	Apigenin	1.714	0.004848		S(-)-p-Bromotetramisole oxalate	0.5185	0.0005587
5	K 185	1.559	0.003201		Desipramine hydrochloride	0.5364	0.0008234
6	S(-)-UH-301 hydrochloride	1.531	0.001613		Cephalothin sodium	0.5431	0.006875
7	Nialamide	1.509	0.007601		Ammonium pyrrolidinedithiocarbamate	0.575	5.414E-07
8	Buspirone hydrochloride	1.502	3.91E-06		L-Glutamine	0.6081	0.002113
9	SKF 95282 dimaleate	1.493	0.0042		Amantadine hydrochloride	0.6169	0.001555
10	Mianserin hydrochloride	1.459	0.000147		Spironolactone	0.6308	0.0003674
11	L-733,060 hydrochloride	1.459	0.000244		Ranolazine dihydrochloride	0.6431	0.0002837
12	8-Cyclopentyl-1,3-dimethylxanthine	1.345	7.01E-05		IC 261	0.644	7.879E-05
13	MJ33	1.295	0.001931		L-Cysteinesulfinic Acid	0.6713	0.008618
14	Papaverine hydrochloride	1.192	0.000127		(-)-Scopolamine hydrobromide	0.6766	0.008552
15	(±)-Taxifolin	1.188	0.00263		Fenofibrate	0.7524	0.003196
16	Tyrphostin 25	1.091	0.001409				

Table 5.6: The top compounds that elicited an increase or a decrease in the observed ratio of compound-treated RFP compared to control RFP from the TIF2p-GFP TIF1p-RFP screen at t=0 (LHS and RHS respectively). Compounds were only deemed significant if their associated Q values were <0.01 (Q value calculation as described in sections 2.2.3.4 and 3.3.1).