DEVELOPMENT OF A SYNTHETIC BIOLOGY TOOL TO QUANTIFY THE CONTRIBUTION OF PROTEIN COMPONENTS AND PROTEIN INTERACTIONS WITHIN TRANSCRIPTIONAL REGULATORY COMPLEXES

By

Kateland Simmons

March 2019

A Thesis submitted to the

Faculty of Graduate and Postdoctoral Studies

in partial fulfillment of the requirements

for the degree of

Master of Science in Cellular and Molecular Medicine

specialization Bioinformatics

Department of Cellular and Molecular Medicine

Faculty of Medicine

University of Ottawa

Abstract

We have formed the basis for a synthetic gene regulatory network in *S. cerevisiae* to assess the relative contribution of proteins and protein interactions to the function of transcriptional regulatory complexes. This gene network contains two synthetic transcriptional regulator (TR) fusion proteins that control transcriptional activation and repression of a reporter expressed by a synthetic promoter. This system can be modulated and quantified specifically by our manipulations since the transcriptional regulators (TRs) are only active when drug induced. In this thesis, I characterized and optimized my gene network to enable accurate measurements of transcriptional regulation. I demonstrated that my synthetic repressor interacts with a selected native transcriptional regulatory complex. Future mutation experiments can be performed to investigate the relative importance that the selected protein interaction plays on transcriptional regulation.

New synthetic repressor fusion proteins can be created to study the relative importance of protein interactions between other TRs, making this tool versatile. This tool will allow us to quantify interactions and gain a deeper understanding of the complicated mechanisms controlling transcriptional regulation.

Acknowledgements

Throughout my Master's degree, I have been extremely lucky to have the best mentors, the most supportive friends and colleagues, and the most loving family.

I would first like to thank Dr. Kaern for being my rock, my mentor, and my guiding light. You have helped me grow not only as a scientist, but also as a person.

I would also like to thank Dr. Downey for his unwavering support, mentorship, and for graciously allowing me to work in his lab on part of my project.

My lab mates have definitely played a positive role in my life. Thank you Brendan Camellato, Danny Salem, Hilary Phenix, Ian Roney, and Setti Belhouari for passing down your wisdom to me and making my days brighter.

I am very appreciative to John Haddad for taking me under his wing at the beginning of my Master's. You made my transition into graduate school fun and easy with your support and friendship.

I would like to thank Mila Tepliakova for her unconditional support and friendship in and out of the lab. I've learned so much from you about science and life.

I would like to thank the Downey and Rudner labs, especially Emma Bondy-Chorney, Amanda Bentley-DeSousa, and Anthony Rossl, for teaching me how to do protein work.

I am very grateful to John Girgis for his friendship and mentorship and for teaching me how to use LaTeX with his templates.

I would also like to thank Dr. Paul Joyce for giving me the chance during my undergraduate degree at Concordia University to gain my first experience in molecular biology research. You sparked my love for research and I am forever thankful for that.

I am thankful to have the wonderful Thesis Advisory Committee that I have. Thank you Dr. Downey, Dr. Perkins, and Dr. Blais for guiding me and supporting me with my project.

Finally, I would like to thank my family and friends for loving me, supporting me and always lifting me up with encouragement throughout my degree. I specifically want to thank my mother and sister, Carrie and Karley Simmons, my father, Greg Simmons, and my soulmate, Richard Gilliatt, for always pushing me to strive harder and stay positive when times are hard. I couldn't have done it without y'all.

Dedication

This thesis is dedicated to my father, Greg Simmons, who instilled a curiosity for science within me from a young age. Although he is no longer on this earth, his passion for life, science, and all of the small beautiful things in our world will continue to live on within me.

Contents

Glossary

Co-IP Co-Immunoprecipitation 3

G418 Geneticin 13

GST glutathione S-transferase 22

GTFs general transcription factors 1

HCR highly conserved region 22

HDAC histone deacetylase 1, 4, 21, 40

HID histone deacetylase interaction domains 22, 40

HTP High throughput 3, 4, 5

ID interaction domain 7

IP Immunoprecipitation 18, 19, 39

LCS Light Chain Specific 19

LiAc lithium acetate 11, 12

LTP low throughput 3, 4, 5

mRNA messenger RNA 1

MS mass spectrometry 3

NAT Nourseothricin 13, 24

PAH paired amphipathic alpha-helix 22

PCR Polymerase Chain Reaction 11, 13, 24

PEG polyethlyene glycol 12

PIC pre-initiation complex 1, 23, 37, 38, 40

PVDF polyvinylidene fluoride 19

RNA Pol II RNA Polymerase II 1, 2, 21

Rpd3L Rpd3 Large 1

Rpd3S Rpd3 Small 1, 2

rTetR reverse tetracycline repressor vi, vii, x, 6, 7, 8, 9, 10, 11, 12, 13, 18, 19, 21, 22, 24, 33, 34, 37, 38, 39

SC Synthetic Complete 14, 15

ssDNA single-stranded carrier DNA 12

TBP TATA-binding protein 1

TFs transcription factors 22

TR transcriptional regulator ii

TRs transcriptional regulators ii, 7

TSS transcription start site 9, 23

UAS upstream activating sequences 23

UASG *GAL* upstream activating sequence 7

URS upstream repressing sequences 23

WCE Whole Cell Extract 17, 18

Y2H yeast two-hybrid 3, 7, 8

List of Figures

List of Tables

Chapter 1

Introduction

1.1 Protein Interactions

Protein complexes have roles in most fundamental cellular processes. For example, RNA Polymerase II (RNA Pol II) is the enzyme responsible for all eukaryotic messenger RNA (mRNA) synthesis (Kornberg, 2001). It is conserved amongst all eukaryotes and forms the pre-initiation complex pre-initiation complex (PIC) through protein interactions with the TATA-binding protein (TBP) and general transcription factors (GTFs) TFIIA, -B, -D, -E, -F, and -H (Figure 1) (Hantsche and Cramer, 2017; Louder et al., 2016; Thomas and Chiang, 2006; Hahn, 2004). Once this promoter-bound complex is formed, basal level transcription can occur (Thomas and Chiang, 2006).

RNA Pol II activity can also be regulated by highly intricate protein complexes. For example, in yeast, Rpd3 forms two histone deacetylase (HDAC) complexes with Sin3, designated Rpd3 Small (Rpd3S) and Rpd3 Large (Rpd3L). These complexes contain the same core subunits Rpd3, Sin3, and Ume1. The Rpd3S-Sin3 HDAC complex is responsible for deacetylating core histones within transcribed sequences and consists of the core subunits plus Eaf3 and Rco1 (Figure 2A) (Carrozza et al., 2005). The Rpd3L-Sin3 HDAC complex is involved in heat stress response and consists of the core subunits plus Pho23, Sds3, Rxt2,

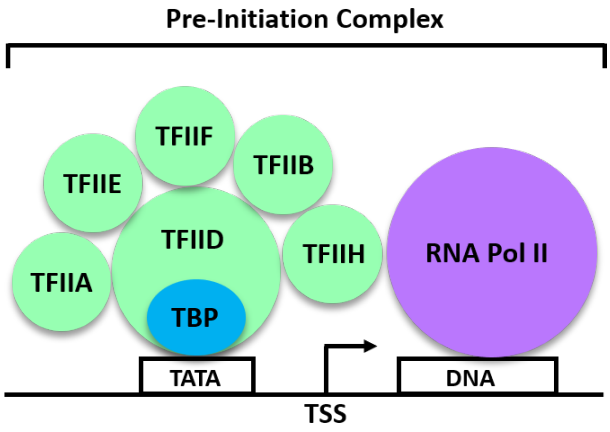

Figure 1: **Schematic of the eukaryotic pre-Initiation complex at a TATA-containing promoter.**

The TBP and TFIID bind at the TATA box and recruit TFIIA, -B, -D, -E, -F, -H and RNA Pol II to initiate transcription. Adapted from (Roney, 2016).

Rxt3, Dep1, and Cti6 (Figure 2B) (Carrozza et al., 2005). When Rpd3S deacetylates core histones, the histone tails become positively charged and bind DNA to return to a closed heterochromatin state (Bannister and Kouzarides, 2011). This represses transcription since RNA Pol II cannot bind at the promoter.

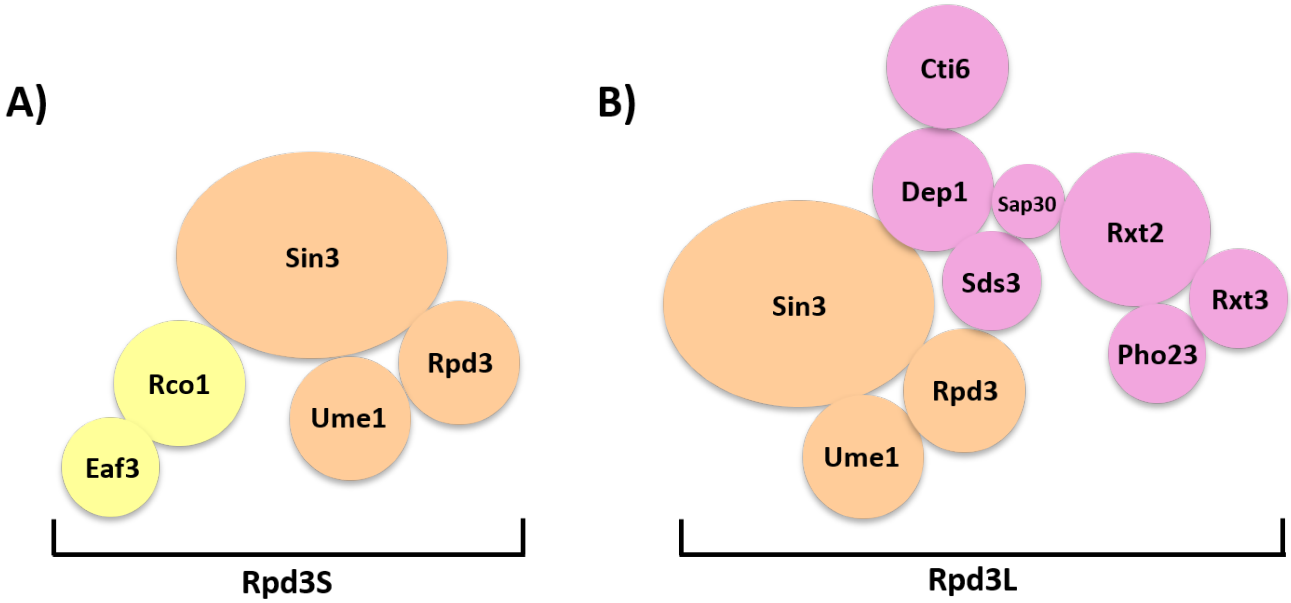

Figure 2: **Representation of the Rpd3S and Rpd3L HDAC complex.**
A) The Rpd3S complex and B) the Rpd3L complex. Orange proteins correspond to the core Rpd3 proteins that are shared between both complexes. Yellow proteins are present in the Rpd3S complex and pink proteins are present in the Rpd3L complex. Figure adapted from (McDaniel and Strahl, 2013; Sardiu et al., 2009).

Because of their often critical roles, significant efforts have been devoted to determining protein and protein complex structure and composition. Many classical biochemistry methods such as Co-Immunoprecipitation (Co-IP) and yeast two-hybrid (Y2H) assays have been used to identify protein interactions. These classical methods were considered time consuming and low-throughput in the sense that only a few proteins could be screened at a time.

Advancements to classical biochemistry methods have been made and high-throughput methods were created. These methods allow us to simultaneously screen thousands of proteins at a time at a faster pace. The most commonly used high-throughput methods are large-scale Y2H systems and protein-complex purification paired with mass spectrometry (MS) analysis (Mann et al., 2001; Titz et al., 2004).

High throughput (HTP) and low throughput (LTP) methods have been used to create massive amounts of protein interaction data that is stored in large

databases such as the Biological General Repository for Interaction Datasets (BioGRID), the Universal Protein Resource (UniProt), and the Database of Interacting Proteins (DIP). These databases provide information about a protein's function, its interactors, experimental methods used, references for its datasets, and note whether HTP or LTP methods were used. Although this information gives us a foundation of knowledge about whether proteins interact or not, it does not allow us to prioritize protein interactors based on their interaction strength with a specific protein.

Data compiled from protein interaction databases are often used to create protein interaction networks as models of protein complexes. The problem is that they can become very large, like Sin3's protein interaction network in Figure 3, making it difficult to obtain useful information about the protein. For example, based on the network in Figure 3, it is not clear that Sin3 often interacts with Rpd3 and Ume1, circled in red, to form the core of the Rpd3 HDAC complex. Additionally, these networks do not provide details about what the protein's function actually is or which interactors impact its function.

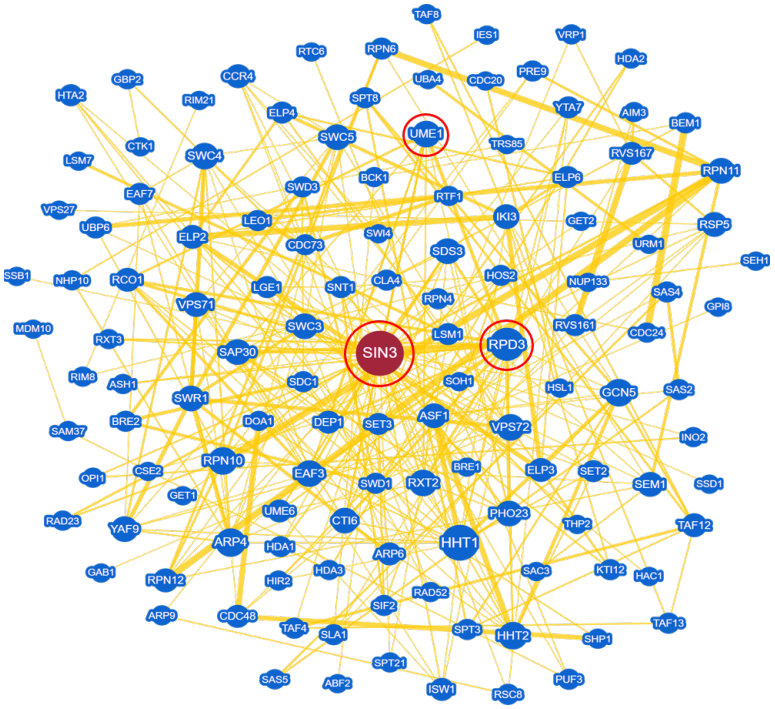

Figure 3: **Protein interaction network for Sin3.**

Shows physical interactions between Sin3 and other proteins. The core Rpd3 complex components are circled in red. Obtained from BioGrid.

Despite the wealth of information accumulated within these databases, determining the exact function and importance of individual protein complex components can be elusive (Cusick et al., 2005). False positive and false negative protein interactions may be identified in protein interaction screens, which can be misleading when trying to assign function and relative importance of protein complex components. Moreover, even if a reported interaction is real, we still do not know if the interaction will occur in other experimental settings or naturally within an organism. These HTP and LTP methods are useful to identify that a potential protein interaction occurs, but there is a need for a tool that is capable of assigning relative importance to the protein interactions so we can learn more about the function of protein complexes.

1.2 Main Goal and Objectives

The goal of my research is to contribute to the creation of a platform that can be used to validate and quantify the relative importance of protein interactions in transcriptional regulatory complexes. Achieving this goal is important because it will yield a deeper understanding of protein function and how it affects transcriptional regulation.

To address this main goal, there are two objectives I set out to accomplish:

Objective 1: Characterize a synthetic gene regulatory network to enable accurate measurements of transcriptional regulation.

Objective 2: Demonstrate our synthetic gene regulatory network interacts with a selected native transcriptional regulatory complex.

To reach these objectives, I will use a synthetic regulatory system (Figure 4). This system has three components: a transcriptional activator constitutively expressed by the *MYO2* promoter, and a transcriptional repressor constitutively expressed by the *TDH3* promoter, and a *yEGFP* reporter protein expressed by a synthetic promoter that is controlled by the activator and repressor.

The transcriptional activator is the GEV transactivator. GEV is a fusion protein made of the *GAL4* DNA binding domain, an estrogen receptor, and the electronegative portion of the herpes simplex virus protein *VP16* (Louvion et al., 1993; Sadowski et al., 1988). When induced with β-Estradiol, GEV can localize to the nucleus and bind the *GAL4* DNA binding sites present in the synthetic promoter (McIsaac et al., 2011). Once bound, it can activate *yEGFP* expression. It is advantageous to use GEV as our activator because its components make it a very strong transcriptional activator and it is only active in the presence of β-Estradiol (McIsaac et al., 2011).

The transcriptional repressor is a fusion protein containing rTetR tethered

to the interaction domain (ID) of a transcriptional repressor. rTetR was created when four amino acid substitutions were made in the Tet repressor (TetR). This caused the reverse effect of TetR and allows rTetR to bind DNA in the presence of doxycycline (M Gossen and Bujard, 1995). When induced with doxycycline, rTetR-ID can bind the *tet* operators in the synthetic promoter. Once bound, it can repress *yEGFP* expression through steric hindrance and/or histone deacetylation. We used a new rTetR variant with more mutations that improve DNA binding and enhance doxycycline sensitivity (Roney et al., 2016).

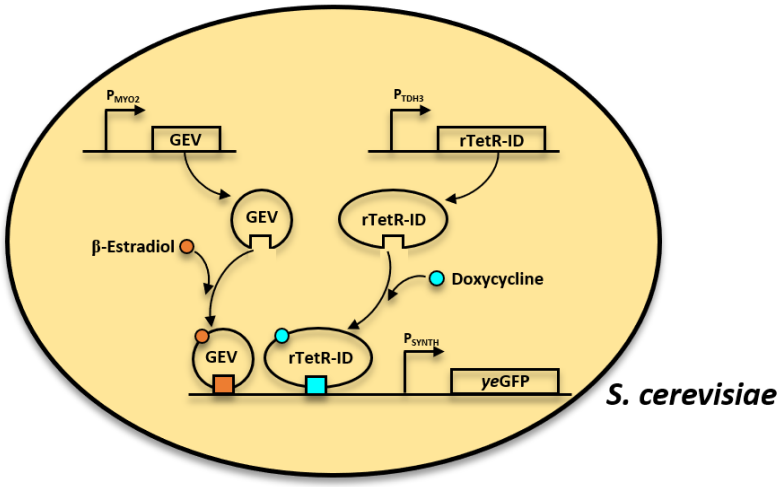

Figure 4: **Schematic of synthetic gene system.**
Includes two constitutively expressed TRs: the GEV transactivator is controlled by pMYO2 and induced by β-Estradiol. The rTetR-ID repressor is controlled by pTDH3 and induced by doxycycline. The GFP reporter is controlled by the synthetic GEV/rTetR-ID promoter, designated by P_{SYNTH} in the diagram. (Adapted from (Roney, 2016))

Compared to the Y2H system, using a synthetic system like ours is more ideal for several reasons. The Y2H system is consists of two plasmids and a reporter gene regulated by the *GAL* upstream activating sequence (UASG): a bait protein that contains the *GAL4* DNA-binding domain fused to a known protein and a prey protein that contains the *GAL4* activation domain fused to a protein from a library of yeast genomic DNA fragments (Fields and Song, 1989; Chien et al., 1991). If the proteins interact, they will bind at a *GAL4* binding site of the UASG and induce transcriptional activation of a reporter gene (Chien et al., 1991). This method is

limited to only telling us if the bait and prey proteins interact or not. Additionally, it requires that the proteins of interest can enter the nucleus and the bait itself is not a potent activator (Chien et al., 1991). This limits the number of proteins that can be tested with this system. Our system is structured differently since *yeGFP* transcription is activated by induced GEV and we select which protein that rTetR-ID interacts with by modifying its interaction domains. Therefore we detect two things: whether rTetR-ID and a protein interact and the strength of their interaction based on how strongly *yeGFP* transcription is repressed. Because GEV and rTetR-ID require an inducer to bind at our synthetic promoter, we can control when and by how much these fusion proteins activate or repress *yEGFP* transcription by altering the inducer concentrations. This gives us more control over our experimental system compared to the Y2H system.

Another reason our system is ideal to use is that the specificity of activator and repressor binding ensures that changes in *yEGFP* transcription levels are only due to our manipulation and not random endogenous proteins binding to the promoter. Finally, it is ideal and more practical to use an rTetR fusion protein rather than TetR because rTetR does not require the constant presence of doxycycline to maintain inactivity like TetR does (M Gossen and Bujard, 1995).

Chapter 2

Materials and Methods

2.1 Strains Used in Study

The following strains were used in this study. They will be referred to by their shortened nicknames to simplify descriptions.

Table 1: **Yeast strains used in this study.**

Four different synthetic promoter variants integrated at the native *ADE2* locus were used in this study to express the *yEGFP* reporter. Promoters are named $P_{GT\#}$ to show they receive input from the **G**EV activator and the r**T**etR-Sin3 repressor. The number (#) indicates the location of the *tet* binding sites in the promoter in relation to the transcription start site (TSS) (0= scrambled *tet* sites, 1=distal, 2=medial, 12=distal+medial). The GEV activator was integrated at the native *GAL4* locus and expressed by the Myo2 promoter. The rTetR-*SIN3* repressor was integrated at the native *ADE4* locus and expressed by the Tdh3 promoter.

STRAIN NAME	GENOTYPE	BACKGROUND	TYPE OF REPRESSION	NICKNAME
TRS00	*ade2::KanMX-rtAct1-pGT0-yEGFP* *gal4::His3-pMyo2-GEV-tPgk1* *ade4::NatMX-pTdh3-(rTetR-M2-SE-G72P)-SIN3-tPgk1*	BY4742	None	Negative Strain
TRS01	*ade2::KanMX-rtAct1-pGT1-yEGFP* *gal4::His3-pMyo2-GEV-tPgk1* *ade4::NatMX-pTdh3-(rTetR-M2-SE-G72P)-SIN3-tPgk1*	BY4742	Distal	Distal Strain
TRS02	*ade2::KanMX-rtAct1-pGT2-yEGFP* *gal4::His3-pMyo2-GEV-tPgk1* *ade4::NatMX-pTdh3-(rTetR-M2-SE-G72P)-SIN3-tPgk1*	BY4742	Medial	Medial Strain
TRS12	*ade2::KanMX-rtAct1-pGT12-yEGFP* *gal4::His3-pMyo2-GEV-tPgk1* *ade4::NatMX-pTdh3-(rTetR-M2-SE-G72P)-SIN3-tPgk1*	BY4742	Medial + Distal	Combination Strain
IR154	*ade2::KanMX-rtAct1-pGT12-yEGFP* *gal4::His3-pMyo2-GEV-tPgk1*	BY4742	None	Parental Strain
353	*MATa ura3Δ leu2Δ his3-1 met15Δ rpd3::NATMX*	BY4741	None	*RPD3* Delete Strain

2.2 Preparing rTetR-*SIN3* DNA

2.2.1 Genomic Extraction of Yeast Genomic DNA

DNA from a strain containing the rTetR-*SIN3* gene was genomically extracted. This strain contained a new rTetR variant with mutations that improve DNA binding and enhance doxycycline sensitivity (Roney et al., 2016). All steps involving phenol and chloroform:isoamyl alcohol were performed under a fume hood.

Yeast cells were struck out from freezer stock on a YPD agar plate containing G418 antibiotic selection. Plates were incubated at 30°C for 3 days. One isolated colony of yeast containing rTetR-*SIN3* DNA was selected from an agar plate and transferred to the bottom of a 1.5mL microcentrifuge tube with a sterile pipette tip. 20μL of breaking buffer and a pinch of acid-washed glass beads were added to the tube of cells. A master mix of phenol:chloroform:isoamyl alcohol(ratio of 25:24:1) was made and 20μL were added to the tube of cells. The tube was vortexed at maximum speed for 3min, then 20μL of TE buffer was added. The sample was vortexed briefly then spun down at 13000RPM for 5min. Then the supernatant was transferred to a fresh sterile microcentrifuge tube. 100μL of 100% ethanol (P016EAAN, Commercial Alcohols) was added to the tube and the tube was mixed by inverting 5-6 times. The tube was spun down at 13000RPM for 5min and the supernatant was discarded. The pellet was air-dried for 30min by placing the inverted tube over a sheet of Kimwipes. Then the pellet was resuspended in 20μL of TE buffer and incubated for 10min at room temperature. Samples were stored at 4°C.

2.2.2 PCR of rTetR-*SIN3* DNA

The desired rTetR-*SIN3* DNA piece was amplified into two pieces using Polymerase Chain Reaction (PCR) due to the large size of the gene. The standard PCR protocol was: 30 denaturing reaction cycles at 98°C for 15 seconds, 1 annealing reaction cycle using a gradient of 45°C-60°C for 20 seconds, and 1 extension reaction cycle at 72°C for 1min 18sec (15sec per 1000bp of product). The reaction was set up using the following products: 10pmol/μL of each forward and reverse primers (Integrated DNA Technologies), 1 unit of Phusion High Fidelity Polymerase (F530S, Fisher Scientific), 200ng of template DNA, 10mM of Deoxynuclotide mix (N0447S, New England Biolabs), and 1x High Fidelity Buffer (B0518S, New England Biolabs).

The PCR products were run on a 1% agarose gel for 45min at 85V to visualize the product sizes and confirm the correct DNA pieces were obtained.

2.3 Yeast Cell Transformation

DNA containing the rTetR-*SIN3* cassette was transformed at the *ADE4* locus in the genome of five different yeast BY4742 strains(MATα, his3Δ1, leu2Δ0, lysΔ0, ura3Δ0) (EUROSCARF). These parental strains were previously transformed in our lab with the promoter variants + *yEGFP* and GEV DNA cassettes at the *ADE2* and *GAL4* loci respectively. See Table 1 for strain information.

An adaption of a standard lithium acetate (LiAc) high-efficiency yeast transformation protocol was used for integrating the synthetic DNA into the parental strain genomes (Burke et al., 2000). For each parental strain, 5mL of 2xYPD was inoculated with a single yeast colony and grown overnight at 30°C in a tissue culture rotator at 200RPM (1640, Barnstead Lab-Line). The next morning, the overnight cultures were diluted to an OD_{600} of approximately 0.2 in 50mL of 2xYPD. Diluted cultures were grown for 3 to 5 hours at 30°C and

200RPM until they reached an OD_{600} of approximately 0.7. Cultures were spun down at 4000RPM (Precision) for 5min in a sterile centrifuge tube. The supernatant was then discarded, the pellet was resuspended in 10mL autoclaved ddH_2O, and subsequently spun down at 4000RPM (Precision) for 5min. The supernatant was discarded again, the pellet was resuspended in 1mL of 100mM LiAc (Sigma-Aldrich, L4158), and the cell suspension was transferred to a sterile 1.5mL microcentrifuge tube. The cell suspension was spun down at 130000RPM (VWR) for 10s and the supernatant was discarded. The pellet was resuspended to a final volume of 500μL with 100mM LiAc. An aliquot of single-stranded carrier DNA (ssDNA) (Deoxyribonucleic acid sodium salt type III from salmon testes, Sigma-Aldrich, D1626) was boiled at 100°C for 5min using a heat block (VWR), then immediately stored on ice at 4°C. The cell suspension was vortexed (VWR) and 50μL aliquots were transferred to labeled 1.5mL microcentrifuge tubes. These aliquots were spun down at 13000RPM for 10s and the supernatant was discarded. The following transformation mix components were added together in the order listed in one tube to make a master mix (number of transformation tubes x volume of each component = volume of each component to add to master mix): 240μL 50% polyethlyene glycol (PEG) (Sigma-Aldrich, P3640), 36μL 1M LiAc (Sigma-Aldrich, L4158), 50μL ssDNA (Sigma-Aldrich, D1626). 326μL of master mix were distributed to each of the cell suspension aliquots. 100ng of the each of the two rTetR-*SIN3* DNA pieces were mixed with sterile ddH_2O to a total of 36μL and added to each cell suspension aliquot as well. Samples were resuspended by vortexing at maximal speed for 1min. Then samples were heat shocked for 45min in a 42°C hot water bath (VWR). After heat shocking the samples, they were spun down at 6000-8000RPM for 2min and the supernatant was discarded. The pellet was resuspended in 1mL of YPD and incubated at 30°C with shaking for 3-4hrs to allow the antibiotic resistant genes to start being expressed. After the incubation period, samples were spun down again at 6000-8000RPM for 2min and the supernatant was discarded. The pellet

was resuspended in 150μL of sterile ddH_2O. All transformed cells were plated on YPD media supplemented with 2% glucose (Wisent Inc., 600-035-DG), 5% adenine (Sigma-Aldrich, A9126), and Nourseothricin (NAT) and Geneticin (G418) antibiotic selection.

Positive transformants were validated by PCR using primers designed to anneal at the upstream *ADE4* locus and in the middle of the rTetR-*SIN3* gene. PCR products were ran on a 1% agarose gel for 45min at 85V to visualize the product sizes and confirm correct integration into the *ADE4* locus.

2.4 Flow Cytometry

All flow cytometry experiments were completed using the BD FACSCelesta 12-colour Analyzer flow cytometer. Yeast strains were freshly grown from freezer stock on appropriate YPD selection plates 3 days in advance of flow cytometry experiments. Prior to analysis, yeast cultures were diluted 1:10 in 50nM sodium citrate buffer. The *yEGFP* fluorescence was detected using a 488nM laser and a 530/30 band-pass filter. A small forward and side scatter gate (about 40% of yeast population) was used to gate the yeast cells and reduce signal detection from extrinsic sources. MATLAB R2013a (Mathworks Inc, Natick, Massachusettes) was used to analyze FCS files and obtain mean fluorescent intensity values.

2.4.1 Strain Validation

5mL of YPD media supplemented with 5% Adenine was inoculated with one single colony from positive transformants of each strain plus a non-*yEGFP*-expressing BY4742 negative control. Samples were incubated overnight at 30°C and 200RPM.

The next day, overnight cultures were diluted to an OD_{600} of 0.1 in 3mL YPD and incubated for 3hrs at 30°C and 200RPM to bring cells into log phase. A deep-well incubation plate with beads, stored at -20°C, containing prepared drug concentrations in Synthetic Complete (SC) Media was incubated at 30°C with agitation 30min prior to cultures being added to the plate. After incubation, cultures were diluted to an OD_{600} of 0.1 in the incubation plate with varying combinations of β-Estradiol and doxycyline concentrations. The incubation plate was incubated at 30°C with agitation for 5hrs. Prior to flow cytometry analysis, induced cultures were diluted 1 in 10 in 50nM sodium citrate buffer.

2.4.2 Comparing the Effects of Different Inducer Preparation Techniques

Three replicates were made for fresh and prepared drug conditions using the same batch of 2x SC media supplemented with 5% Adenine. Desired drug concentrations were diluted in 2xSC media in deep-well incubation plates with beads the night before and the day of the experiment and stored at -20°C and 4°C, respectively.

3mL of 2x SC media was inoculated with a single yeast colony, then incubated overnight at 30°C and 200RPM. The next day, overnight cultures were diluted to an OD_{600} of 0.1 in 10mL of 2xSC media and incubated at 30°C and 200RPM for 3hrs. The freshly-prepared and pre-prepared deep-well incubation plates with beads were incubated at 30°C with agitation 30min prior to cultures being added to the plate. After incubation, cultures were diluted to an OD_{600} of 0.1 in the incubation plate with varying combinations of freshly-prepared and pre-prepared β-Estradiol and doxycyline concentrations. The incubation plates were incubated at 30°C with agitation for 4hrs. Prior to flow cytometry analysis, induced cultures were diluted 1 in 10 in 50nM sodium citrate buffer.

2.4.3 Optimizing Induction Times

3mL of SC media was inoculated with one single colony from selected yeast strains, then incubated overnight at 30°C and 200RPM. The next day, overnight cultures were diluted to an OD_{600} of 0.1 in 10mL of SC media and incubated at 30°C and 200RPM for 3hrs. A deep-well incubation plate with beads stored at -20°C containing prepared drug concentrations in SC Media was incubated at 30°C with agitation 30min prior to cultures being added to the plate. After incubation, cultures were diluted to an OD_{600} of 0.1 in the incubation plate with varying combinations of β-Estradiol and doxycyline concentrations. The incubation plate was incubated at 30°C with agitation for the specified number of hours for each experiment (4hrs, 5hrs, or 6hrs). Approximately 20min prior to flow cytometry analysis, induced cultures were diluted 1 in 10 in 50nM sodium citrate buffer so analysis could take place exactly at the 4hr, 5hr, or 6hr time point.

2.4.4 Characterizing Strains Using Dose-Response Experiments

3mL of SC media supplemented with 5% Adenine was inoculated with a single colony from selected yeast strains, then incubated overnight at 30°C and 200RPM. The next day, overnight cultures were diluted to an OD_{600} of 0.1 in 15mL of SC media and incubated at 30°C and 200RPM for 3hrs. A deep-well incubation plate with beads (stored at -20$^{/circ}$) containing prepared drug concentrations in SC media was incubated at 30°C with agitation 30min prior to cultures being added to the plate. After incubation, cultures were diluted to an OD_{600} of 0.1 in the incubation plate with varying combinations of β-Estradiol and Doxycyline concentrations. Dilutions of each strain were scheduled 10min apart from each other so that each plate would be analysed after exactly 5hrs of

incubation. The incubation plate was incubated at 30°C with agitation for 5hrs. 20min prior to analysis, induced cultures were diluted 1 in 10 in 50nM sodium citrate buffer.

2.5 TCA Protein Preps

Yeast cell cultures were grown overnight in YPD supplemented with 5% Adenine at 30°C and 200RPM. The next day, overnight cultures were diluted to an OD_{600} of 0.8 in a larger volume of media and incubated for 1hr at 30°C and 200RPM. After incubation, 6 OD equivalents of cells were spun down at 3000RPM (4°C) for 3min in falcon tubes, the supernatant was removed, and the pellet was washed with cold ddH_20. Cells were transferred to screw-cap microcentrifuge tubes and spun down at 13000RPM (4°C) for 3min. The supernatant was removed and cells were flash frozen in liquid nitrogen/dry ice. Samples were thawed on ice, then resuspended in 300μL of 20% TCA (Sigma-Aldrich T6399) and 100μL of glass microbeads. Tubes were bead beat twice for 3min and placed on ice for 1min in between bursts to lyse the cells. The supernatant (S1) was collected in a fresh sterile microcentrifuge tube and placed on ice. Samples in screw-cap microcentrifuge tubes were vortexed for 10s with 300μL of 5% TCA (Sigma-Aldrich) and the supernatant (S2) was pooled with S1. The pooled samples were spun down at 13000RPM (4°C) and the supernatant was removed. The pellet was then resuspended in 100μL of 2X-SDS-PAGE sample buffer with 100mM DTT by vortexing. Samples were boiled at 100°C and pelleted at 13000RPM. The supernatant was retrieved and placed in a fresh sterile microcentrifuge tube. Samples were stored at -20°C.

2.6 Whole Cell Extract Protein Preps

Yeast cell cultures were grown overnight in YPD supplemented with 5% Adenine at 30°C and 200RPM. The next day, overnight cultures were diluted to an OD_{600} of 0.7 and incubated for 4-5hrs at 30°C and 200RPM until they reached an OD_{600} of 1. After incubation, 100 OD equivalents of cells were spun down at 3000RPM (4°C) for 3min in falcon tubes, the supernatant was removed, and the pellet was washed with cold ddH_20. Cells were transferred to 2mL screw-cap microcentrifuge tubes and spun down at 13000RPM (4°C) for 3min. The supernatant was removed and cells were flash frozen in liquid nitrogen/dry ice. Pellets were thawed on ice and resuspended by vortexing in 750μL of cold lysis buffer (50 mM Tris-HCl pH8.0, 5 mM EDTA, 150 mM NaCl, 0.5% NP40) with inhibitors (10 mM Glycerol-2-phosphate, 5 mM NaF, 10mM sodium butyrate, 1.75mM PMSF, 1mM DTT, complete protease inhibitor tablet (without EDTA; Roche)). Approximately 1mL of acid-washed glass beads were added to the screw-cap tubes. Samples were lysed in the Bead Beater for eight cycles of 1.5min, with 2min rest on ice in between bursts. The bottom of the tubes were then pierced once with an 18 gauge needle and placed into fresh FACS tubes (falcon tube product number 352054) on ice. Samples were centrifuged at 2000RPM for 1min (4°C). All material from the FACS tubes was transferred to sterile microcentrifuge tubes on ice and then spun down at 17000g for 10min (4°C). The supernatant was removed and samples were spun down again at 17000g for 10min (4°C). The Whole Cell Extract (WCE) supernatant was collected and transferred to a sterile microcentrifuge tube. Samples were stored at -20°C if used for antibody validation purposes or samples were used immediately for immunoprecipitation experiments.

2.7 Immunoprecipitation (IP) of rTetR-Sin3 and Rpd3

100 OD equivalents of each yeast strain were prepared following the protocol in section 2.6. 500μL of WCE lysate for each sample were transferred to sterile microcentrifuge tubes to serve as Immunoprecipitation (IP) samples. 20μL of WCE for each sample were combined 1:1 with 1X SDS sample buffer and transferred to sterile microcentrifuge tubes to serve as inputs. Samples were boiled at 100°C for 5min. After boiling, inputs were placed on ice. The TetR monoclonal antibody (Clontech, 631131) was diluted 1:1000 in TBST. 20μL of antibody were added to each IP sample tube and mixed by gently pipetting up and down. All IP sample tubes were placed on an end-over-end rotator for 2hrs at 4°C.

20min before the end of the 2hr incubation period, the Protein G magnetic dynabeads (BioRad, 161-4023) were prepared. 20μL of bead mix per IP sample were placed into a sterile microcentrifuge tube using a clipped pipette tip. The following wash protocol was followed to wash the Protein G magnetic dynabeads: The bead mix tube was placed in a magnetic Dynabead rack for 30sec to allow beads to stick to the tube wall. Liquid was removed and discarded. The bead mix tube was removed from the magnetic Dynabead rack, resuspended with 750μL IP lysis buffer using a clipped pipette tip, and placed back in the magnetic dynabead rack for 20sec to allow beads stick to the tube wall. Liquid was removed and discarded and this washing protocol was repeated two more times. After the last wash, the liquid was removed and beads were resuspended in 50μL of IP lysis sample buffer per IP sample. Samples were incubated at room temperature for 2hrs.

After the 2hr incubation of IP samples, 50μL of prepared bead suspension were added to each IP sample using a clipped pipette tip. IP sample tubes were incubated for 50min in the end-over-end rotator at 4°C. After the second incubation period, the IP sample + bead mixes were spun down at 3000RPM for

5sec at 4°C. The beads were washed three times following the bead wash protocol. Tubes were placed on the magnetic Dynabead rack for 30sec and the liquid was removed. Tubes were spun down at 3000RPM for 5sec at 4°C, tubes were placed on magnetic Dynabead rack again for 30sec, and remaining liquid was removed. Each IP sample was resuspended in 50μL of IP sample buffer with a clipped pipette tip to elute my protein from the beads. IP samples were incubated for 10min at 65°C, with gentle mixing of tubes every 3min. After incubation, the IP samples were spun down at 3000RPM for 7 sec at room temperature and placed back in the magnetic Dynabead rack to allow beads to stick to the tube wall. The eluted protein supernatant was collected and transferred to a sterile microcentrifuge tube. Eluted IP protein samples were boiled for 5min at 100°C then spun down at 10000RPM for 4min at room temperature.

15μL of each input and IP sample were loaded into an 8% SDS-PAGE gel and run at 140V for approximately 1.5hrs. The gel was transferred to a polyvinylidene fluoride (PVDF) membrane (Immunobilon Transfer Membranes:IPVH00010) for 1.5hrs at 85V at 4°C overnight.

The transferred membrane was blocked for 30min with 5% milk in 1X TBST with agitation, then sliced horizontally with a razor blade right above the 75kDa marker. Membrane 1 (top half containing the rTetR-Sin3 protein) and membrane 2 (bottom half containing the Rpd3 protein) were incubated at 4°C with agitation overnight with TetR monoclonal (Clontech, 631131) and Rpd3 monoclonal (Santa Cruz, SC-514160) primary antibodies, respectively. Primary antibodies were diluted 1:1000 in 5% milk with TBST and 10% NaAzide. Afterward, the membranes were washed 3 times for 9min with 1X TBST and both membranes were incubated for 1hr with goat anti-mouse Light Chain Specific (LCS) secondary antibody (115-035-174, Cedarlane Labs, 1:10000 in 5% milk with TBST). Then membranes were washed 3 times for 12min with 1X TBST, incubated with Luminata Forte ECL Chemiluminescence HRP Substrate

(WBLUF0500, Fisher Scientific), and exposed to autoradiography film (Ultident, 39-20810) for 15sec, 30sec, 1min, 3min, and 7min. Membranes were stained with Ponceau S and digitally scanned.

Chapter 3

Results

3.1 Network Construction and Validation

The purpose of these experiments was to validate that the rTetR-ID cassette was integrated at the correct locus and our synthetic promoter variants showed expected transcriptional activation and repression behaviour upon induction with β-Estradiol or doxycycline. It is important that these strains have consistent and expected transcriptional activation and repression upon induction to ensure accurate results are obtained every time.

3.1.1 Network Construction

rTetR-*SIN3* Network Repressor

I chose to use yeast Sin3 to create the rTetR-Sin3 fusion protein to repress my synthetic transcriptional regulatory network. Sin3 is an enzymatic HDAC that serves as a scaffold on which other DNA-binding transcription factors and cofactors can assemble (Grzenda et al., 2009). Although it does not have repression abilities on its own, Sin3 can repress transcription as a part of the Sin3-Rpd3 HDAC complex. Once Sin3 recruits the Rpd3 complex, it can be targeted to gene promoters where it will interfere with RNA Pol II's ability to

bind, therefore repressing transcription (Silverstein and Ekwall, 2005; Kadosh and Struhl, 1998).

Using Sin3 to create our rTetR fusion protein is ideal because it is well-studied, with its interaction domains mapped out by previous researchers. For example, several researchers have created length variants of *SIN3* and completed glutathione S-transferase (GST) pulldown, yeast two-hybrid, and chromatin immunoprecipitation assays with the Sin3 variants and several Sin3 interactors (Grigat et al., 2012; Kliewe et al., 2017). This allowed them to assess where specific proteins bind within the Sin3 interaction domains and to create an interaction map of Sin3. Because these interaction domains are already mapped out, we can mutate specific parts of these domains and confidently predict the outcome. Additionally, using an rTetR-Sin3 fusion protein rather than using only the endogenous Sin3 to bind at our promoters allows us to make changes to the repressor without affecting endogenous Sin3 function.

An example of this map shows seven conserved regions that allow Sin3 to bind to the Sin3-Rpd3 complex transcription factors (TFs) and other proteins (Figure 5). It contains four paired amphipathic alpha-helix (PAH) domains, two histone deacetylase interaction domains (HID), and a highly conserved region (HCR) (Wang et al., 1990; Silverstein and Ekwall, 2005; Kliewe et al., 2017).

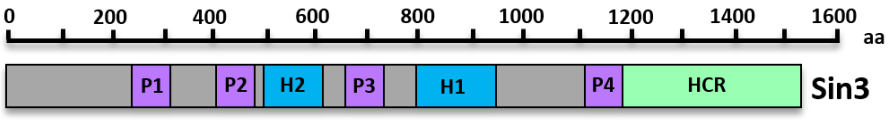

Figure 5: **Representation of the Sin3 interaction domains.**
Includes PAH1-PAH4 (P1-P4), two HDAC interaction domains (H1 and H2), and a highly conserved region (HCR). Figure adapted from (Kliewe et al., 2017).

Finally, using our synthetic regulatory network to study the relative importance of Sin3 can yield useful applications to human health research since *SIN3* is evolutionarily conserved from *S. cerevisiae* to *H. sapiens* (Kadosh and Struhl, 1997; Grigat et al., 2012). For example, studies have shown that

mutations in the mammalian endogenous *SIN3* gene have caused intellectual disabilities, brain development delays, and may cause de-regulation of cell growth-related genes, which leads to cancer (Witteveen et al., 2016; El-Hattab et al., 2010; Suzuki et al., 2008). If we can validate the relative importance of protein interactions between Sin3 and other proteins in yeast, we can potentially apply these findings to human health applications.

Synthetic Promoter Variants in Network

Four different promoter variants, schematized in Figure 6, were used to express the *yEGFP* reporter gene for our synthetic transcriptional regulatory network. All promoter variants contain three adjacent GAL4 binding sites at the native upstream activating sequences (UAS) region, which is within 1500 nucleotides of the TSS (Zaman et al., 2002). There are two tandem *tet* operators that are located at either a medial location(at the native upstream repressing sequences (URS) region, 100bp from the TSS), distal location (adjacent to GAL4 binding sites, 260bp from TSS), or a combination of medial and distal locations. A negative control promoter was used that contains the three adjacent *GAL4* binding sites but the *tet* operators are scrambled at both medial and distal locations. This allows us to positively control for activation and negatively control for repression.

Promoters containing distal repressor binding sites (P_{GT1} and P_{GT12}) sterically inhibit activator binding since they compete for the same DNA binding region. Alternatively, promoters containing medial repressor binding sites (P_{GT2} and P_{GT12}) enzymatically repress transcription through PIC formation interference. Therefore, it is necessary to test for any differences in repression strength between each variant. Then the variant with the strongest repression strength will be used for experiments.

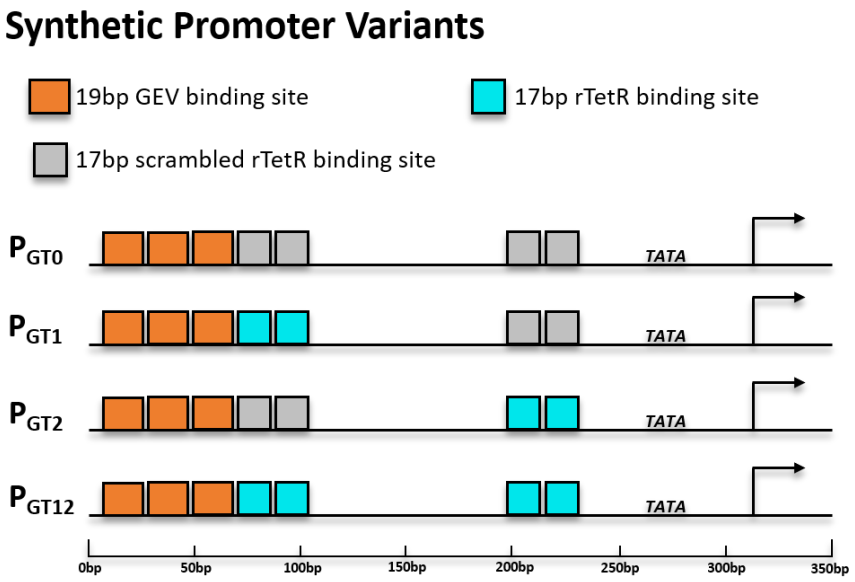

Figure 6: **Schematic of synthetic promoter variants.**

Shows location of DNA binding sites for the activator and repressor. P_{GT0}= negative strain (TRS00), P_{GT1}= distal strain (TRS01), P_{GT2}= medial strain (TRS02), P_{GT12}= combination strain (TRS12). Figure adapted from (Roney, 2016).

3.1.2 Network Validation

The rTetR-*SIN3* cassette, which includes NAT antibiotic selection, was integrated into the *ADE4* locus of parental strains and cells were grown on antibiotic selection plates. Successful integration into the *ADE4* locus causes a color change in yeast, from red to white (Ugolini and Bruschi, 1996). White colony growth indicated the DNA was integrated at the correct loci since the *ADE4* locus was disrupted. Colony growth on antibiotic selection plates offers additional confirmation of positive DNA integration.

Genomic DNA of white yeast colonies was amplified by PCR with primers upstream in the *ADE4* locus and in the middle of the rTetR-*SIN3* gene. Successful amplification of this region ensured that both DNA pieces were certainly transformed.

Finally, flow cytometry analysis was used as a method to validate proper network functionality. If the transformation was successful, activation or repression of *yEGFP* transcription in each strain should occur when induced

with β-Estradiol or doxycycline, respectively, depending on the strain's promoter type. All strains showed activation of *yEGFP* expression when induced with β-Estradiol, and the distal, medial, and combination strains showed repression of *yEGFP* expression when induced with doxycycline. The negative control strain did not show repression of *yEGFP* expression because its scrambled *tet* binding sites do not allow repressor binding. But the negative control strain did show activation of *yEGFP* expression since it does contain GEV binding sites.

3.2 Objective 1 Results

To ensure our synthetic gene regulatory network would yield accurate measurements of transcriptional regulation, the network needed to be examined and characterized. Typically, researchers use dose-response curves to characterize a system. These dose-response curves yield information about several aspects of network behaviour. For instance, we can obtain information about the inducer concentrations required to achieve minimal and maximal activation and/or repression, how changing inducer concentrations affects *yEGFP* transcription in the network, and how gene network mutations affect baseline *yEGFP* expression. By characterizing the gene regulatory network, we can know what behaviour to expect and eventually fit the dose-response curves to mathematical models to predict the outcome of making changes to the network.

3.2.1 Comparing the Effects of Different Inducer Preparation Techniques

The purpose of this experiment was to investigate whether variability in flow cytometry results occurred when using freshly-prepared drugs vs. drugs prepared the night before flow cytometry experiments. Ideally, preparing for experiments in advance can reduce the amount of preparation required the day of

experiments. As a result, this decreases the chance for user error due to time constraints on the day of flow cytometry experiments. The combination strain with distal + medial *tet* binding sites was used for this analysis.

In Figure 7, the dose-response curves show no major difference in activation of of *yEGFP* expression for all three replicates when using freshly-made or pre-made β-Estradiol.

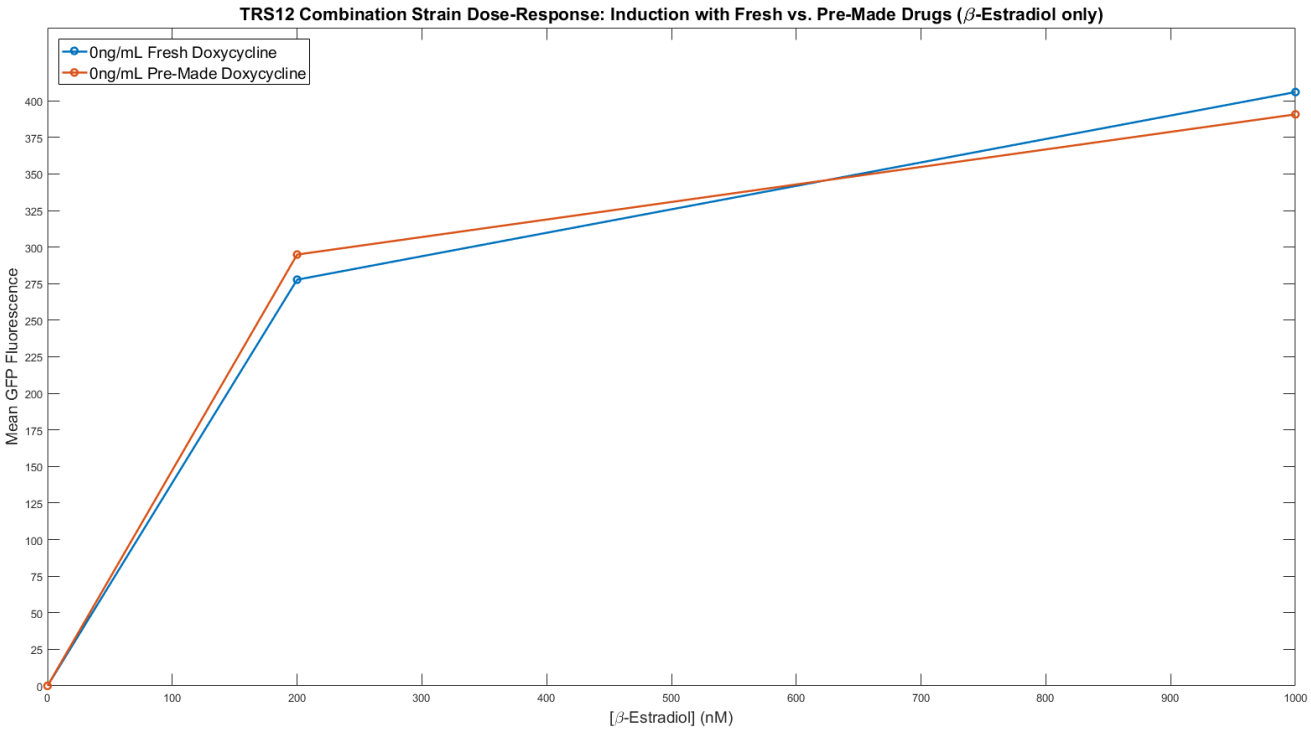

Figure 7: **Dose-response curve showing *yEGFP* activation in the combination strain using β-Estradiol only**

Fresh (blue) and pre-made (red) β-Estradiol was used at 0nM, 200nM, and 1000nM with no addition of doxycycline. Mean fluorescence is similar when using fresh or pre-made drug.

Figure 8 shows no major difference in repression of of *yEGFP* expression for all three replicates when using freshly-made or pre-made doxycycline. Therefore, I decided to prepare the drugs the day before future flow cytometry experiments to increase time efficiency and reduce the chance of user error.

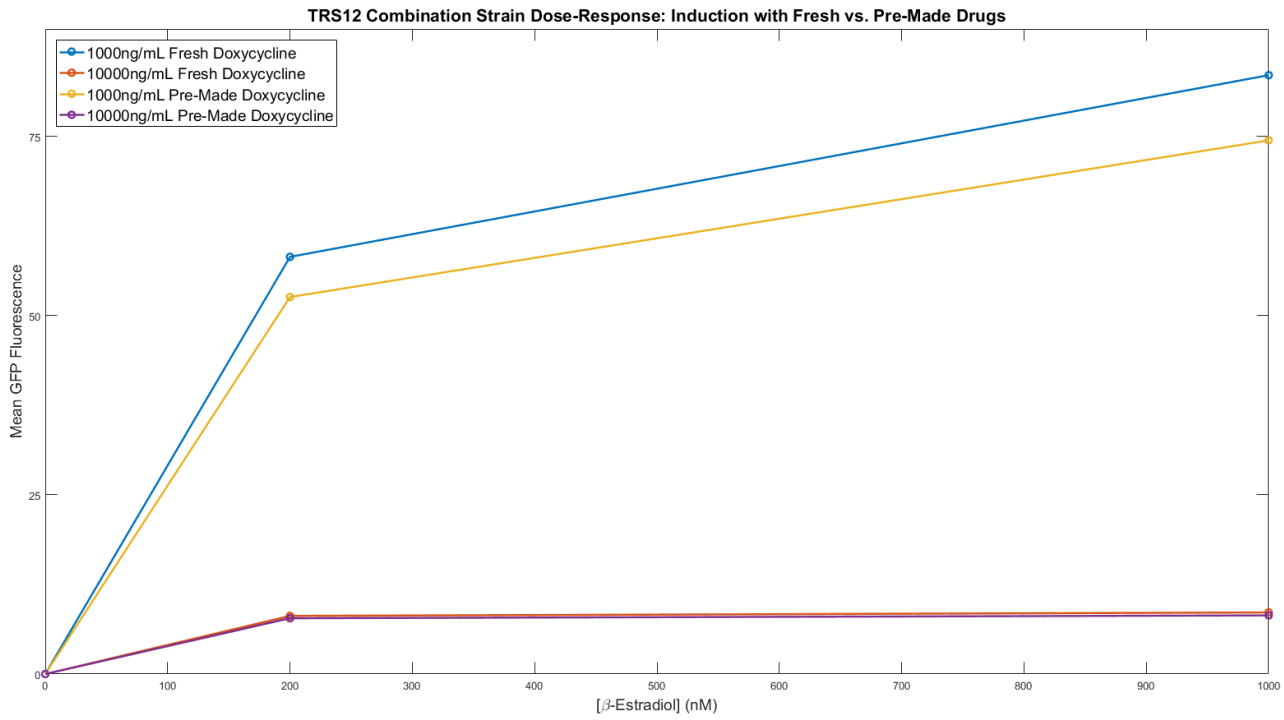

Figure 8: **Dose-response curve showing *yEGFP* activation in the combination strain using β-Estradiol and doxycycline**

Fresh doxycycline was used at 1000ng/mL (blue) and 10000ng/mL (red) with 0nM, 200nM, and 1000nM of fresh β-Estradiol. Pre-made doxycycline was used at 1000ng/mL (yellow) and 10000ng/mL (purple) with 0nM, 200nM, and 1000nM of pre-made β-Estradiol. Mean fluorescence is similar when using fresh or pre-made drug.

3.2.2 Testing Drug Induction Times

The purpose of this experiment was to test whether mean *yEGFP* fluorescence in our strains is similar after being induced with varying β-Estradiol and doxycycline concentrations for 4, 5, or 6 hours. I wanted to choose an induction time in which maximal activation and repression of *yEGFP* expression could occur. Using an induction time that doesn't allow for maximal activation and repression can reduce accurate reproducibility of experiments since *yEGFP* expression levels are still changing.

I measured fluorescence intensity in two replicates for the distal, medial, and combination yeast strains. The negative control yeast strain with scrambled *tet* binding sites and a BY4742 yeast strain lacking our entire synthetic system were

used as negative controls. I then compared the mean fluorescence of each strain at 4, 5, and 6 hour time points. Analysis was completed over three consecutive days, with one induction time experiment being completed per day.

There was a difference in *yEGFP* transcriptional activation levels between 4hrs and 5hrs of induction in both replicates for each strain (Figure 9A). Between 5hrs and 6hrs of induction, both replicates showed that there was no major difference in *yEGFP* transcriptional activation activation levels for each strain. In (Figure 9B), there is a difference in *yEGFP* transcriptional repression levels between 4hrs and 5hrs of induction for both replicates for each strain. But there is no major difference in *yEGFP* transcriptional repression levels between 5hrs and 6hrs of induction in both replicates for each strain. Therefore, I selected the 5hr induction time as the optimal induction time to use in future flow cytometry experiments.

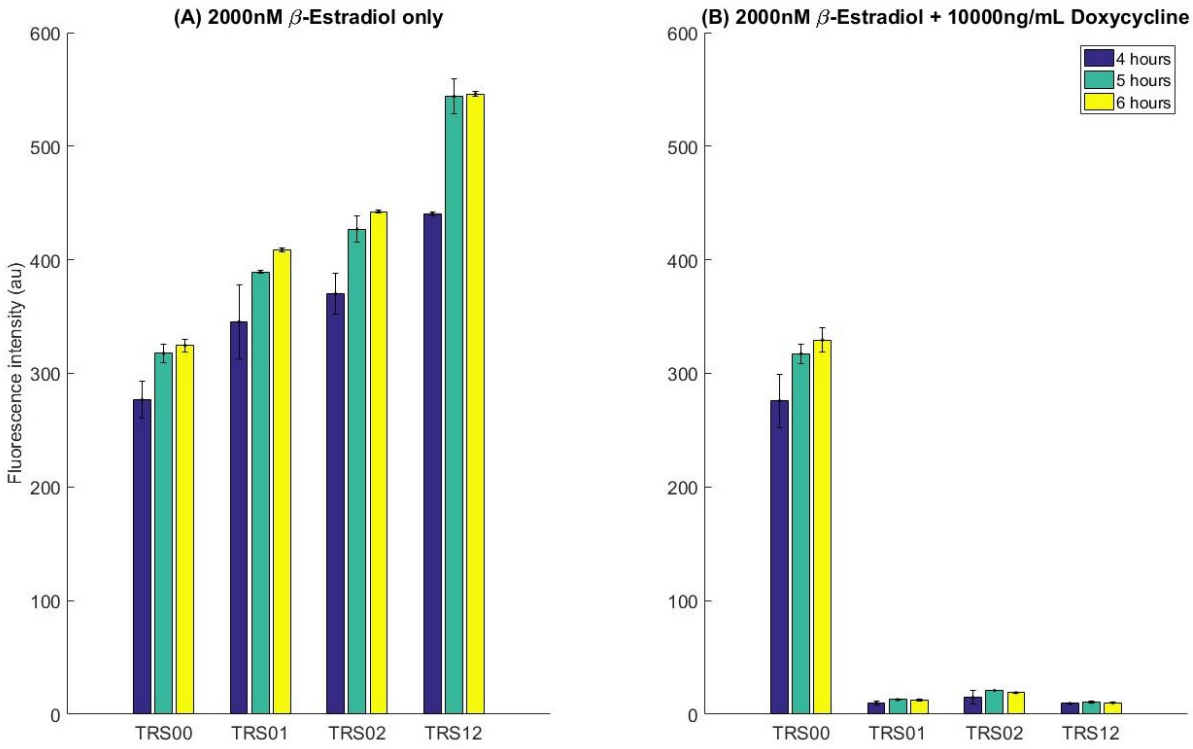

Figure 9: ***yEGFP* fluorescence intensity of rTetR-Sin3 variants** after: a) induction with β-Estradiol only or b) induction with β-Estradiol and doxycycline, for 4 hours (blue bar), 5hrs (teal bar), or 6 hours (yellow bar). There is a difference in *yEGFP* fluorescence intensity between 4 and 5 hours of induction, but no major difference between 5 and 6 hours of induction for each strain in both induction conditions.

3.2.3 Characterizing Strains Using Dose-Response Experiments

The purpose of this experiment was to compare levels of transcriptional activation and repression with varying concentrations of β-Estradiol and doxycycline to find the maximum and minimum levels of activation and repression. The medial, distal, and combination strains were tested and BY4742 lacking our synthetic regulatory network was used as a negative control.

Overall, the combination strain (TRS12) showed a wider range of sensitivity to drug induction for *yEGFP* transcriptional activation and repression and yielded steeper dose-response curves compared to the other promoter variants (Figure 10). At maximal doxycycline concentrations, the combination strain's repression nearly reached full saturation. The combination strain also showed

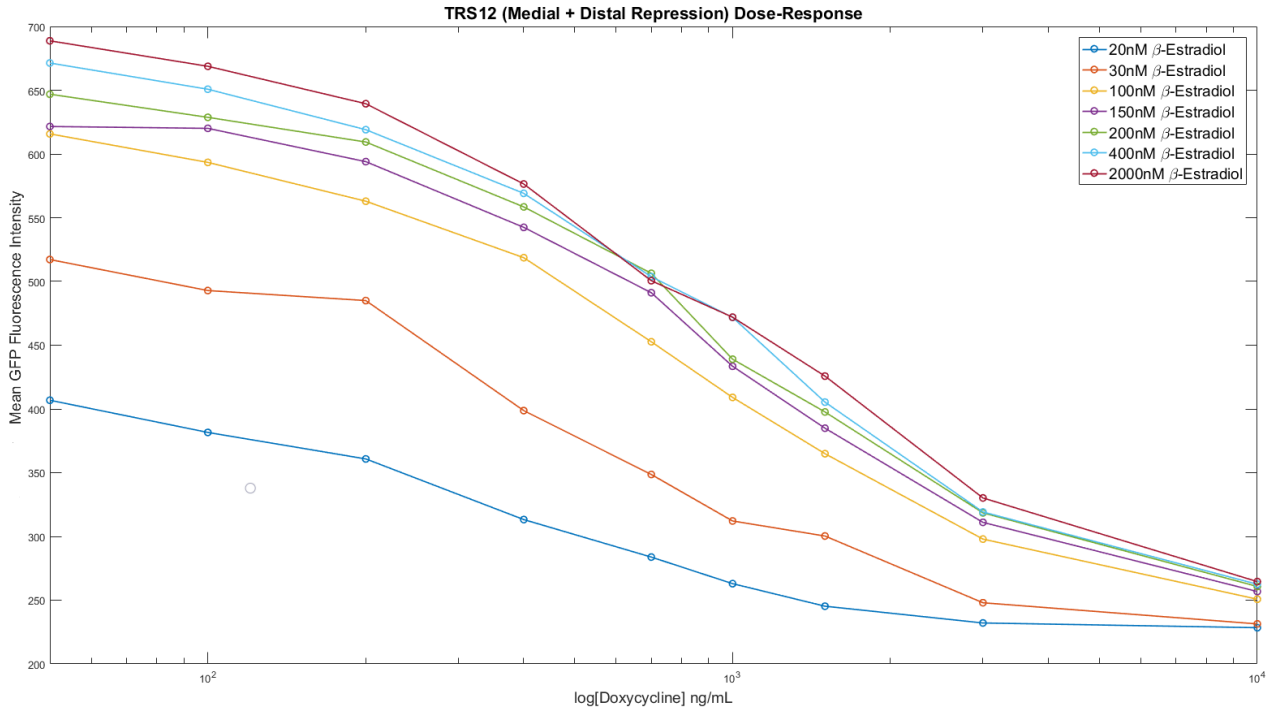

Figure 10: **Dose-response curves for the combination strain (TRS12) at varying β-Estradiol and doxycycline concentrations.**

Strains were induced for 5 hours under the following conditions: 20nM, 30nM, 100nM, 150nM, 200nM, 400nM, and 2000nM β-Estradiol and 0ng/mL, 50ng/mL, 100ng/mL, 200ng/mL, 400ng/mL, 700ng/mL, 1000ng/mL, 1500ng/mL, 3000ng/mL, and 10000ng/mL doxycycline. This strain shows a higher sensitivity to induction of *yEGFP* transcriptional activation and repression with steeper dose-response curves compared to the distal and medial strains, TRS01 and TRS02.

higher repression strength at maximal doxycycline concentrations. Furthermore, the onset of repression in the combination strain occurred at lower doxycycline concentrations than the onset of repression in the other strains.

The distal strain (TRS01) exhibited slightly sigmoidal dose-response curves that appear to achieve *yEGFP* transcriptional repression saturation almost as strong as the combination strain at maximal doxycyline concentrations (Figure 11). There is some variation in the 30nM, 150nM, 400nM, and 2000nM β-Estradiol curves that could potentially be from user error. But overall, its curves are less steep compared to the combination strain's curves, which means it requires a higher concentration of doxycycline to achieve the onset of *yEGFP*

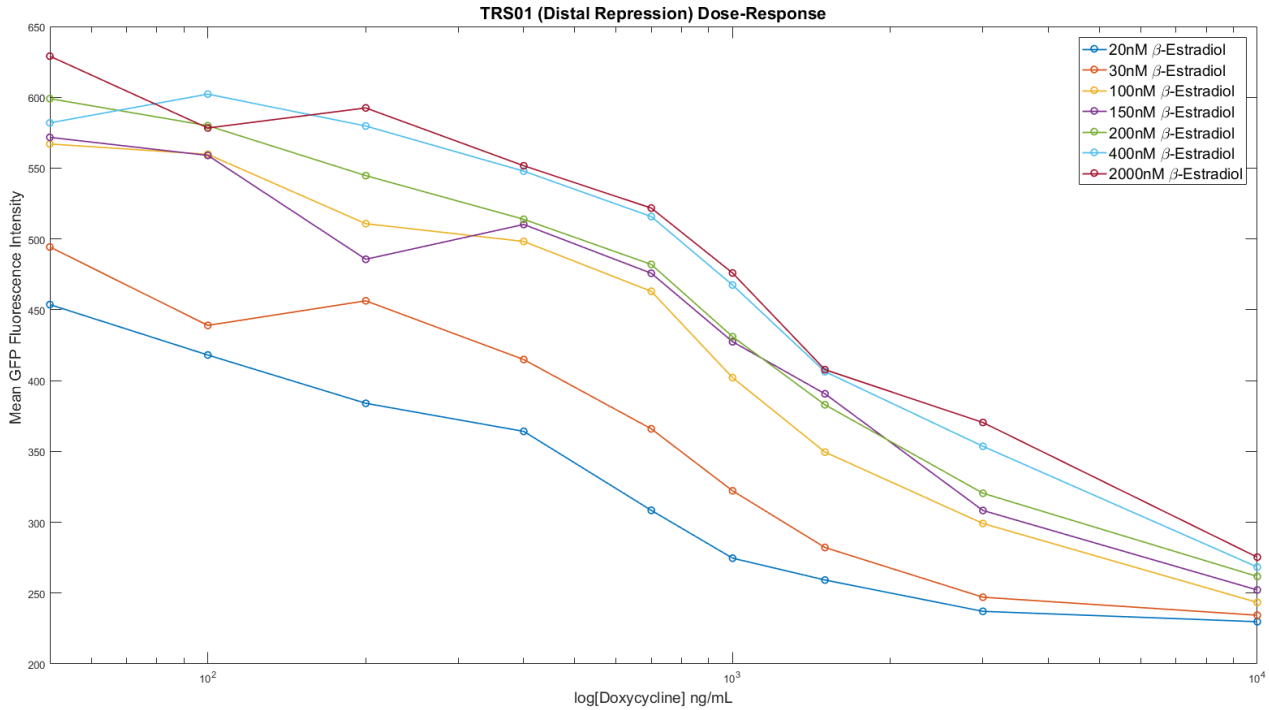

Figure 11: **Dose-response curves for the distal strain (TRS01) at varying β-Estradiol and doxycycline concentrations.**

Strains were induced for 5 hours under the following conditions: 20nM, 30nM, 100nM, 150nM, 200nM, 400nM, and 2000nM β-Estradiol and 0ng/mL, 50ng/mL, 100ng/mL, 200ng/mL, 400ng/mL, 700ng/mL, 1000ng/mL, 1500ng/mL, 3000ng/mL, and 10000ng/mL doxycycline. This strain shows a moderate sensitivity to induction of *yEGFP* transcriptional activation and repression. It requires a higher level of doxycycline to induce *yEGFP* transcriptional repression compared to the combination strain.

transcriptional repression. It has a moderate sensitivity to drug induction for *yEGFP* transcriptional activation and repression.

The medial strain (TRS02) exhibited the weakest sensitivity to drug induction for *yEGFP* transcriptional activation and repression (Figure 12). It yielded more linear dose-response curves compared to the other two strains, which means it requires an even higher concentration of doxycycline to achieve the onset of *yEGFP* transcriptional repression. At maximal doxycycline concentrations, the medial strain's *yEGFP* transcriptional repression did not come close to reaching maximal repression saturation compared to the distal and combination strains.

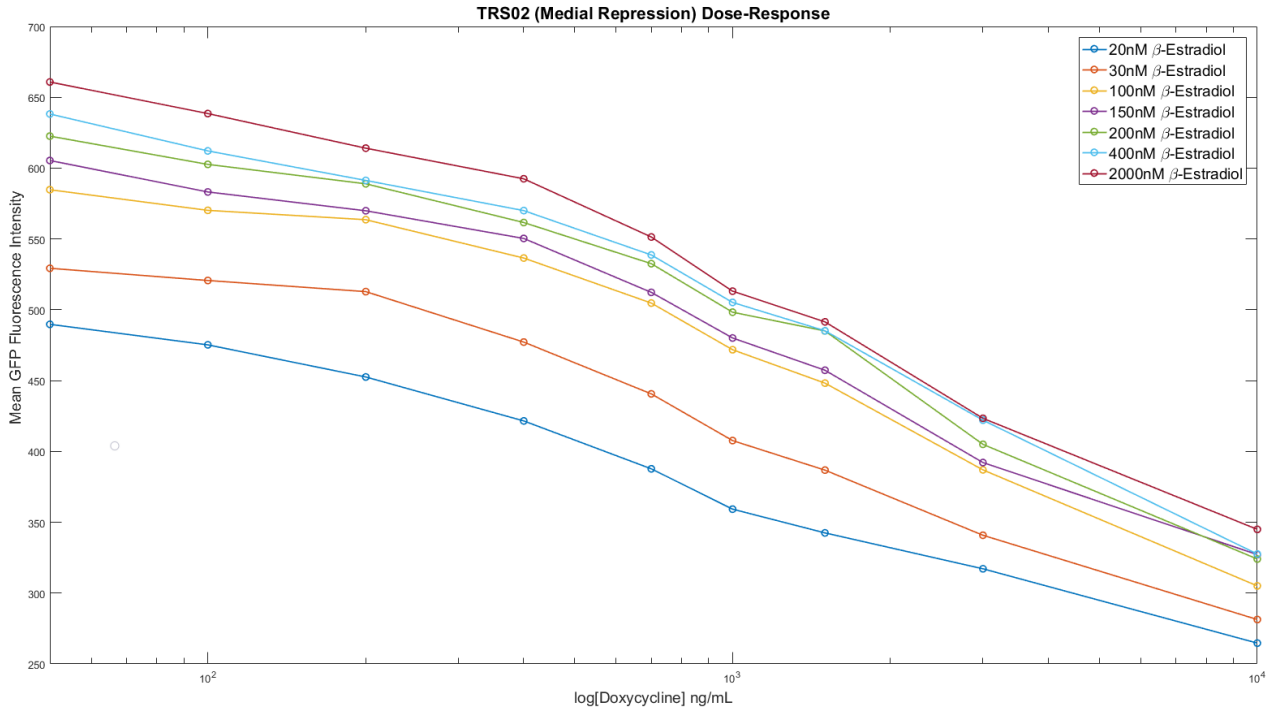

Figure 12: **Dose-response curves for the medial strain (TRS02) at varying β-Estradiol and doxycycline concentrations.**

Strains were induced for 5 hours under the following conditions: 20nM, 30nM, 100nM, 150nM, 200nM, 400nM, and 2000nM β-Estradiol and 0ng/mL, 50ng/mL, 100ng/mL, 200ng/mL, 400ng/mL, 700ng/mL, 1000ng/mL, 1500ng/mL, 3000ng/mL, and 10000ng/mL doxycycline. This strain shows a weaker sensitivity to induction of *yEGFP* transcriptional activation and repression compared to the distal and combination strains. It requires an even higher level of doxycycline to induce *yEGFP* transcriptional repression and yields a more linear dose-response curve.

3.3 Objective 2 Results

To perform future rTetR-*SIN3* mutation experiments to disrupt the rTetR-Sin3-Rpd3 interaction, we must first establish that the interaction does occur. It is important to perform this step before disrupting protein interactions with the platform to prove two things: 1) the protein interaction is real and not due to non-specific binding with another protein of similar size, and 2) to ensure that the mutations are actually disrupting the interaction and a loss of interaction result is not due to the proteins not interacting in the first place.

3.3.1 Immunoprecipitation (IP) of rTetR-Sin3 and Rpd3

The purpose of this experiment was to confirm the positive protein interaction between my rTetR-Sin3 protein and the endogenous yeast Rpd3 protein. If rTetR-Sin3 and Rpd3 positively interact, the complex will bind to the TetR antibody and be immunoprecipitated.

The TRS12 combination strain contains the rTetR-Sin3 and Rpd3 proteins. Strain IR154 is the parental strain for the combination strain, which lacks the rTetR-Sin3 protein and served as a negative control. The 353 *RPD3* deletion mutant (obtained from Dr. Michael Downey's Lab) lacks both of the proteins of interest and served as a second negative control. In Figure 13, the rTetR-Sin3 protein is visualized for TRS12 input and IP samples when blotting with αTetR. This demonstrates that TRS12 is the only strain that contains the rTetR-Sin3 protein. When blotting with αRpd3, the Rpd3 protein is visualized for TRS12 input, TRS12 IP, and IR154 input samples. The ponceau S demonstrates that an equal amount of protein was loaded in each lane and that the positive IP results are due to rTetR-Sin3 and Rpd3 positively interacting, not because any input samples are present in the IP samples. These results confirms that Rpd3 interacts with and immunoprecipitates with rTetR-Sin3.

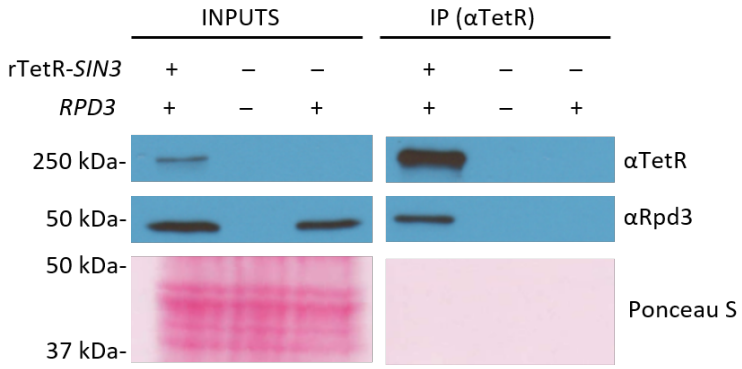

Figure 13: **rTetR-Sin3 and Rpd3 positively interact**

An IP using αTetR was performed on strains TRS12 (rTetR-*SIN3 RPD3* combination strain), 353 (*rpd3*Δ), and IR154 (*RPD3*). The rTetR-Sin3 protein is present in TRS12 input and IP samples when blotted with αTetR, but is not present in the other strains lacking rTetR-Sin3. The Rpd3 protein is present in TRS12 input and IP samples and in the IR154 input sample when blotted with αRpd3. Rpd3 is not present in samples lacking the rTetR-Sin3 protein when blotted with αRpd3, proving that Rpd3 interacts with rTetR-Sin3. Images are all from the same blot and exposure.

Chapter 4

Discussion

The goal of my thesis research was to contribute to the creation of a platform that can be used to validate and quantify the relative importance of protein interactions in transcriptional regulatory complexes. To work towards this goal, I successfully completed two objectives. I performed experiments to enable accurate measurements of transcriptional regulation and I demonstrated my synthetic system worked as expected by testing the interaction between my fusion protein and a transcriptional regulator. Upon completion of my research, I contributed to the advancement of knowledge in our field by laying the foundation for a tool that will expand our knowledge of protein interactions. Additionally, my tool will allow us to explore how important certain interactions are for transcriptional regulation and how we can quantify these interactions.

When creating a new system, I felt it was important to investigate five parameters to ensure accurate and reproducible results:

- The effect of induction time on achieving maximal reporter gene transcriptional activation and repression.
- How inducer preparation affects its ability to induce the system.
- The effect of changing the number and location of repressor binding sites in our system.

- How changing inducer concentration affects system behaviour.
- Confirming the system's fusion protein positively interacts with the transcriptional regulator of interest.

To enable accurate measurements of transcriptional regulation, it is important to optimize parameters that could affect accuracy and reproducibility of results. The first parameter we chose to optimize was the system induction time. In flow cytometry analysis, samples in a 96 well plate are not immediately analysed at the same time. There could be a 5 to 10 minute waiting time between the first and last sample analysis, which means some samples will be induced longer than the desired induction time. When this happens, we want to be sure reporter protein expression levels remain steady so that results are comparable at maximal induction levels. Otherwise, the data would be inaccurate and unusable if expression levels are still changing between the time the first and last samples are analysed. Using an optimal induction time also enables us to know that small changes in reporter protein expression are attributed to changes in protein interactions within our system, not because the system was not fully induced. In our system, I observed a difference between 4 and 5 hours of induction, but no difference between 5 and 6 hours of induction. Therefore, I chose to use a 5 hour induction time for future flow cytometry experiments.

The second parameter I investigated was whether inducer preparation affected its ability to induce our system. To decrease the chance of user error due to time constraints when preparing inducers, it might be helpful to prepare them prior to rather than on the experiment day. Since inducers would be stored at -20°C, there could be a higher chance of inducer molecule degradation compared to freshly prepared inducers that are not frozen. Results would not be accurate or reproducible if experiments were performed using inducers that behaved differently based on their preparation. In our system, I observed there was no major difference between using freshly prepared inducers and inducers prepared the day before. Therefore, I decided to prepare my inducers the day before experiments to decrease

the chance of user error.

Another important aspect of creating a new system is characterizing its behaviour in different experimental scenarios. By doing this, we will be able to predict our system's behaviour when changing certain experimental variables. I explored the third and fourth parameters through dose-response experiments to characterize my system. These experiments allowed me to investigate the effect of changing the number and location of repressor binding sites in my synthetic promoters and how changing inducer concentration affects transcription in my system. The goal of my third parameter was to discover which promoter type exhibited the strongest sensitivity to transcriptional activation and repression. Using a promoter with this quality would produce steeper, sigmoidal dose-response curves that enable me to see smaller shifts in transcription levels when introducing changes to our system. If the synthetic promoter contains distally located repressor binding sites (Distal Strain P_{GT1}), the GEV activator and rTetR-Sin3 repressor will be competing for the same DNA region at distal binding sites (Figure 14A). So this promoter should exhibit stronger sensitivity to repression. If the synthetic promoter contains medially located repressor binding sites (Medial Strain P_{GT2}), the repressor is interfering more with PIC assembly rather than activator binding (Figure 14B). Therefore, this promoter should exhibit weaker sensitivity to repression compared to the promoter with distal repression. A synthetic promoter containing both distally and medially located repressor binding sites (Combination Strain P_{GT12}) would interfere with both activator binding and PIC assembly (Figure 14C). Considering all of these hypotheses, the combination strain should exhibit the strongest sensitivity to repression compared to the other two promoter types. I observed that the combination strain did in fact exhibit stronger sensitivity to repression compared to the other promoters. If this promoter also exhibited greater sensitivity to changing inducer concentrations, I planned to use it in my system for future experiments.

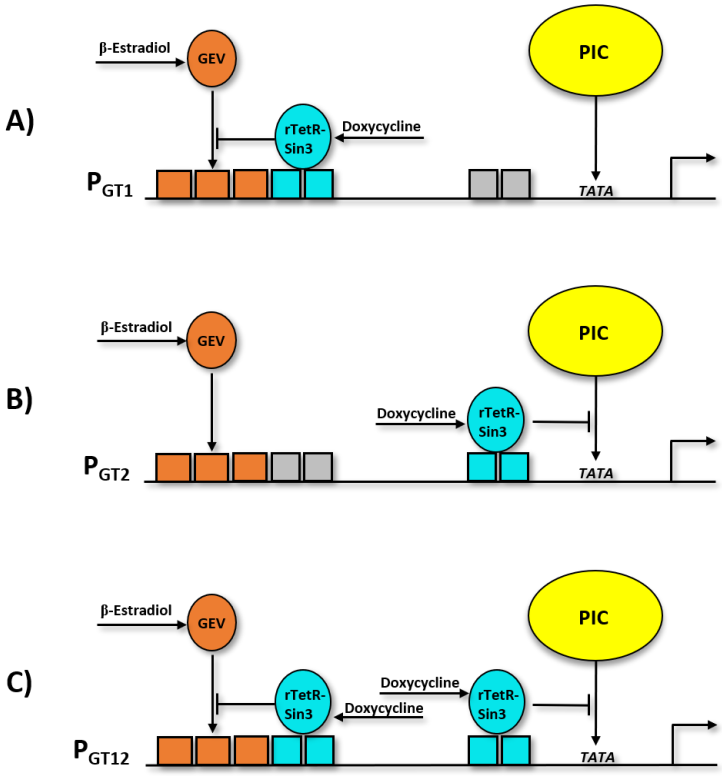

Figure 14: **Differences in repression mechanisms between distal (P_{GT1}), medial (P_{GT2}), and combination (P_{GT12}) strains.**

A) The rTetR-Sin3 repressor competes with the GEV activator for the same DNA binding region. B) The repressor interferes with PIC assembly and does not compete with the activator for the same DNA binding region. C) The repressor interferes with both activator and PIC binding.

The goal of the fourth parameter was to discover how changing inducer concentrations affects system behaviour. It was important to research this because a specific range of inducer concentrations is required to obtain the ideal steep sigmoidal dose-response curve mentioned in the third parameter discussion. Therefore, one must discover this concentration range that yields the most sensitive dose-response curve when characterizing a new system. In my research, I discovered that the combination strain exhibited a dose-response curve with a steeper more sigmoidal shape compared to the other strains that yielded more linear dose-response curves at the same inducer concentrations. The steeper sigmoidal shape enables us to see small increases and decreases in transcription levels along the curve, whereas linear curves would require larger changes in

transcription levels to detect any differences. Without this kind of sensitivity, the effects of changing experimental conditions or slight differences in protein interactions may not be as accurately detectable. Therefore, I chose to use the combination strain to optimize the accuracy of my system.

Before proceeding to future mutation experiments, we must first confirm that the interaction does occur between a synthetic fusion protein and a native protein of interest. Because I am not using the endogenous Sin3 protein to complete my research, my fifth parameter I investigated was confirming whether or not my rTetR-Sin3 protein interacts with the endogenous Rpd3 protein. Confirming this interaction sets the foundation for future mutation experiments in which we mutate Rpd3 DNA binding domains within the rTetR-*SIN3* gene. This experiment also allowed us to create a positive control for these mutation experiments. In my project, I performed IP experiments on protein extracts from three strains: one with Rpd3 and rTetR-Sin3 (combination strain), one with Rpd3 but lacking rTetR-Sin3 (IR154), and one lacking both Rpd3 and rTetR-Sin3 (Rpd3 delete strain). I observed that the combination strain containing both Rpd3 and rTetR-Sin3 yielded positive results when blotting with TetR and Rpd3 antibodies, therefore confirming that our system does interact with the endogenous Rpd3. The IR154 strain with Rpd3 but lacking rTetR-Sin3 only showed positive results when blotted with the Rpd3 antibody. The Rpd3 delete strain lacking both Rpd3 and rTetR-Sin3 showed no positive results when blotted with both antibodies. These two results further confirm the interaction between my fusion protein and the endogenous Rpd3 protein does occur and is not a result of any non-specific binding of other proteins.

Chapter 5

Future Directions

To further develop our system, it would be useful to determine how mutating rTetR-Sin3 interaction domains affects transcriptional regulation within our system. These mutations would disrupt specific protein interactions between rTetR-Sin3 and other proteins in complexes such as the Rpd3 HDAC complex. Rather than only telling us if the proteins interact or not, this system would allow us to investigate which of these protein interactions are important for Sin3's repression function as well.

It would also be interesting to see how mutations change the promoter variants' repression strength if there is not a complete loss of function. The truncated rTetR-Sin3 fusion protein could potentially decrease steric hindrance at GEV binding sites or cause less PIC assembly interference when binding at distal or medial repressor binding sites, respectively.

Since we have confirmed that rTetR-Sin3 and Rpd3 do interact, we could begin these studies by mutating the Rpd3 interaction domain within rTetR-Sin3. Previous studies have shown that Rpd3 positively interacts with Sin3's HID1 domain and a loss of interaction occurs when deleting that portion of *SIN3* (Grigat et al., 2012). By disrupting this interaction, we can confirm whether Rpd3 is integral for Sin3-Rpd3 complex assembly and Sin3's repression function, or if the complex can still assemble through rTetR-Sin3's interaction

with the other components in the complex.

Another protein that would be interesting to disrupt interactions with would be Sds3. Sds3 is a protein important for the structural and catalytic integrity of the Sin3-Rpd3 complex (Lechner et al., 2000). Previous studies have shown that disrupting the interaction between Sin3 and Sds3 results in a loss of function in the Sin3-Rpd3 complex (Clark et al., 2015). Therefore in our system, we would expect to see an extreme decrease or complete loss of repression of *yEGFP* transcription if we disrupted the rTetR-Sin3-Sds3 interaction. By completing these mutation studies, we can further understand the importance of specific proteins within the Sin3-Rpd3 complex.

This tool would also enable researchers to create a completely different rTetR fusion protein. One interesting outcome of doing this would be any changes in repression strength compared to rTetR-Sin3. Similar to the idea of the mutated rTetR-Sin3 protein, creating a fusion protein that is larger or smaller than rTetR-Sin3 could potentially increase or decrease repression strength. These potential changes are one of the reasons why system characterization and optimization experiments are very important for fine-tuning the accuracy of the system.

Our system would open the door to discovering the relative importance of protein interactions between many other transcriptional regulators. This allows us to quantify these reactions and gain a deeper understanding of the complicated mechanisms controlling transcriptional regulation.

Bibliography

Bannister, A. J. and Kouzarides, T. (2011). Regulation of chromatin by histone modifcations. Cell Research *21*, 381–395. 2

Bansal, N., David, G., Farias, E. and Waxman, S. (2016). Emerging Roles of Epigenetic Regulator Sin3 in Cancer. Advances in cancer research *130*, 113–135.

Browning, D. F. and Busby, S. J. W. (2004). The regulation of bacterial transcription initiation. Nature Reviews Microbiology *2*, 1–9.

Burke, D., Dawson, D. and Stearns, T. (2000). Methods in Yeast Genetics. 1st edition, Cold Spring Harbor Laboratory Press. Pg. 103. 11

Carrozza, M. J., Li, B., Florens, L., Suganuma, T., Swanson, S. K., Lee, K. K., Shia, W.-J., Anderson, S., Yates, J., Washburn, M. P. and Workman, J. L. (2005). Histone H3 methylation by Set2 directs deacetylation of coding regions by Rpd3S to suppress spurious intragenic transcription. Cell *123*, 581–592. 1, 2

Chaubal, A., Todi, S. V. and Pile, L. A. (2016). Inter-isoform-dependent Regulation of the Drosophila Master Transcriptional Regulator SIN3. The Journal of biological chemistry *291*, 11566–11571.

Chien, C. T., Bartel, P. L., Sternglanz, R. and Fields, S. (1991). The two-hybrid system: a method to identify and clone genes for proteins that interact with a protein of interest. Proceedings of the National Academy of Sciences of the United States of America *88*, 9578–9582. 7, 8

Clark, M. D., Marcum, R., Graveline, R., Chan, C. W., Xie, T., Chen, Z., Ding, Y., Zhang, Y., Mondragón, A., David, G. and Radhakrishnan, I. (2015). Structural insights into the assembly of the histone deacetylase-associated Sin3L/Rpd3L corepressor complex. Proceedings of the National Academy of Sciences of the United States of America *112*, E3669–E3678. 41

Cusick, M. E., Klitgord, N., Vidal, M. and Hill, D. E. (2005). Interactome: gateway into systems biology. Human molecular genetics *14 Spec No. 2*, R171–R181. 5

Danielle A. Guarracino, B. N. B. and Arora, P. S. (2011). Protein-Protein Interactions in Transcription: A Fertile Ground for Helix Mimetics. Biopolymers Journal *95*, 1–7.

de Ruijter, A. J. M., van Gennip, A. H., Caron, H. N., Kemp, S. and van Kuilenburg, A. B. P. (2003). Histone deacetylases (HDACs): characterization of the classical HDAC family. The Biochemical journal *370*, 737–749.

El-Hattab, A. W., Zhang, F., Maxim, R., Christensen, K. M., Ward, J. C., Hines-Dowell, S., Scaglia, F., Lupski, J. R. and Cheung, S. W. (2010). Deletion and duplication of 15q24: molecular mechanisms and potential modification by additional copy number variants. Genetics in medicine : official journal of the American College of Medical Genetics *12*, 573–586. 23

Farias, E. F., Petrie, K., Leibovitch, B., Murtagh, J., Chornet, M. B., Schenk, T., Zelent, A. and Waxman, S. (2010). Interference with Sin3 function induces epigenetic reprogramming and differentiation in breast cancer cells. Proceedings of the National Academy of Sciences of the United States of America *107*, 11811–11816.

Fields, S. and Song, O.-k. (1989). A novel genetic system to detect protein–protein interactions. Nature *340*, 245. 7

Gallinari, P., Di Marco, S., Jones, P., Pallaoro, M. and Steinkühler, C. (2007). HDACs, histone deacetylation and gene transcription: from molecular biology to cancer therapeutics. Cell research *17*, 195–211.

Grigat, M., Jaschke, Y., Kliewe, F., Pfeifer, M., Walz, S. and Schuller, H.-J. (2012). Multiple histone deacetylases are recruited by corepressor Sin3 and contribute to gene repression mediated by Opi1 regulator of phospholipid biosynthesis in the yeast Saccharomyces cerevisiae. Molecular Genetics Genomics *287*, 461–472. 22, 40

Grzenda, A., Lomberk, G., Zhang, J.-S. and Urrutia, R. (2009). Sin3: master scaffold and transcriptional corepressor. Biochimica et biophysica acta *1789*, 443–450. 21

Hahn, S. (2004). Structure and mechanism of the RNA polymerase II transcription machinery. Nature Structural and Molecular Biology *11*, 394–403. 1

Hantsche, M. and Cramer, P. (2017). Conserved RNA polymerase II initiation complex structure. Current opinion in structural biology *47*, 17–22. 1

Kadosh, D. and Struhl, K. (1997). Repression by Ume6 involves recruitment of a complex containing Sin3 corepressor and Rpd3 histone deacetylase to target promoters. Cell Press *89*, 365–371. 22

Kadosh, D. and Struhl, K. (1998). Targeted recruitment of the Sin3-Rpd3 histone deacetylase complex generates a highly localized domain of repressed chromatin in vivo. Molecular and cellular biology *18*, 5121–5127. 22

Kasten, M. M., Dorland, S. and Stillman, D. J. (1987). A large protein complex containing the yeast Sin3p and Rpd3p transcriptional regulators. Molecular and cellular biology *17*, 4852–4858.

Kliewe, F., Engelhardt, M., Aref, R. and Schüller, H.-J. (2017). Promoter recruitment of corepressors Sin3 and Cyc8 by activator proteins of the yeast Saccharomyces cerevisiae. Current genetics *63*, 739–750. 22

Kornberg, R. D. (2001). The eukaryotic gene transcription machinery. Biological chemistry *382*, 1103–1107. 1

Kouzarides, T. (2007). Chromatin modifications and their function. Cell Press *128*, 693–705.

Lechner, T., Carrozza, M. J., Yu, Y., Grant, P. A., Eberharter, A., Vannier, D., Brosch, G., Stillman, D. J., Shore, D. and Workman, J. L. (2000). Sds3 (suppressor of defective silencing 3) is an integral component of the yeast Sin3[middle dot]Rpd3 histone deacetylase complex and is required for histone deacetylase activity. The Journal of Biological Chemistry *275*, 40961–40966. 41

Lee, D. Y., Teyssier, C., Strahl, B. D. and Stallcup, M. R. (2005). Role of protein methylation in regulation of transcription. Endocrine reviews *26*, 147–170.

Liu, X., Bushnell, D. A. and Kornberg, R. D. (2013). RNA polymerase II transcription: structure and mechanism. Biochimica et Biophysica Acta (BBA)-Gene Regulatory Mechanisms *1829*, 2–8.

Lodish, H., Berk, A. and Zipursky, S. L., e. a. (2000). Molecular Cell Biology. Fourth edition, New York: W. H. Freeman.

Louder, R. K., He, Y., López-Blanco, J. R., Fang, J., Chacón, P. and Nogales, E. (2016). Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. Nature *531*, 604–609. 1

Louvion, J. F., Havaux-Copf, B. and Picard, D. (1993). Fusion of GAL4-VP16 to a steroid-binding domain provides a tool for gratuitous induction of galactose-responsive genes in yeast. Gene *131*, 129–134. 6

M Gossen, S Freundlieb, G. B. G. M. W. H. and Bujard, H. (1995). Transcriptional activation by tetracyclines in mammalian cells. Science *268*, 1766. 7, 8

Mann, M., Hendrickson, R. C. and Pandey, A. (2001). Analysis of protiens and proteomes by mass spectrometry. Annual Review Biochemistry *70*, 437–473. 3

McDaniel, S. L. and Strahl, B. D. (2013). Stress-free with Rpd3: a unique chromatin complex mediates the response to oxidative stress. Molecular and cellular biology *33*, 3726–3727. 3

McIsaac, R. S., Silverman, S. J., McClean, M. N., Gibney, P. A., Macinskas, J., Hickman, M. J., Petti, A. A. and Botstein, D. (2011). Fast-acting and nearly gratuitous induction of gene expression and protein depletion in Saccharomyces cerevisiae. Molecular biology of the cell *22*, 4447–4459. 6

Nasmyth, K., Stillman, D. and Kipling, D. (1987). Both positive and negative regulators of HO transcription are required for mother-cell-specific mating-type switching in yeast. Cell Press *48*, 579–587.

Roney, I. (2016). Construction and optimization of tetracycline-inducible gene expression systems. Master's thesis University of Ottawa. 2, 7, 24

Roney, I. J., Rudner, A. D., Couture, J.-F. and Kærn, M. (2016). Improvement of the reverse tetracycline transactivator by single amino acid substitutions that reduce leaky target gene expression to undetectable levels. Scientific reports *6*, 27697. 7, 10

Sadowski, I., Ma, J., Triezenberg, S. and Ptashne, M. (1988). GAL4-VP16 is an unusually potent transcriptional activator. Nature *335*, 563–564. 6

Sardiu, M. E., Gilmore, J. M., Carrozza, M. J., Li, B., Workman, J. L., Florens, L. and Washburn, M. P. (2009). Determining protein complex connectivity using a probabilistic deletion network derived from quantitative proteomics. PloS one *4*, e7310. 3

Silverstein, R. A. and Ekwall, K. (2005). Sin3: a flexible regulator of global gene expression and genome stability. Current genetics *47*, 1–17. 22

Sternberg, P. W., Stern, M. J., Clark, I. and Herskowitz, I. (1987). Activation of the yeast HO gene by release from multiple negative controls. Cell Press *48*, 567–577.

Suzuki, H., Ouchida, M., Yamamoto, H., Yano, M., Toyooka, S., Aoe, M., Shimizu, N., Date, H. and Shimizu, K. (2008). Decreased expression of the SIN3A gene, a candidate tumor suppressor located at the prevalent allelic loss region 15q23 in non-small cell lung cancer. Lung cancer (Amsterdam, Netherlands) *59*, 24–31. 23

Thomas, M. C. and Chiang, C.-M. (2006). The general transcription machinery and general cofactors. Critical reviews in biochemistry and molecular biology *41*, 105–178. 1

Titz, B., Schlesner, M. and Uetz, P. (2004). What do we learn from high-throughput protein interaction data? Expert review of proteomics *1*, 111–121. 3

Ugolini, S. and Bruschi, C. V. (1996). The red/white colony color assay in the yeast Saccharomyces cerevisiae: epistatic growth advantage of white ade8-18, ade2 cells over red ade2 cells. Current genetics *30*, 485–492. 24

Wang, H., Clark, I., Nicholson, P. R., Herskowitz, I. and Stillman, D. J. (1990). The Saccharomyces cerevisiae SIN3 gene, a negative regulator of HO, contains four paired amphipathic helix motifs. Molecular and cellular biology *10*, 5927–5936. 22

Wang, H., Huang, Z. Q., Xia, L., Feng, Q., Erdjument-Bromage, H., Strahl, B. D., Briggs, S. D., Allis, C. D., Wong, J., Tempst, P. and Zhang, Y. (2001). Methylation of histone H4 at arginine 3 facilitating transcriptional activation by nuclear hormone receptor. Science (New York, N.Y.) *293*, 853–857.

Wishart, J. A., Hayes, A., Wardleworth, L., Zhang, N. and Oliver, S. G. (2005). Doxycycline, the drug used to control the tet-regulatable promoter system, has no effect on global gene expression in Saccharomyces cerevisiae. Yeast (Chichester, England) *22*, 565–569.

Witteveen, J. S., Willemsen, M. H., Dombroski, T. C. D., van Bakel, N. H. M., Nillesen, W. M., van Hulten, J. A., Jansen, E. J. R., Verkaik, D., Veenstra-Knol, H. E., van Ravenswaaij-Arts, C. M. A., Wassink-Ruiter, J. S. K., Vincent, M., David, A., Caignec, C. L., Schieving, J., Gilissen, C., Foulds, N., Rump, P., Strom, T., Cremer, K., Zink, A. M., Engels, H., de Munnik, S. A., Visser, J. E.,

Brunner, H. G., Martens, G. J. M., Pfundt, R. and Kolk, T. K. . S. M. (2016). Haploinsufficiency of MeCP2-interacting transcriptional co-repressor SIN3A causes mild intellectual disability by affecting the development of cortical integrity. Nature Genetics *48*, 877–887. 23

Zaman, Z., Heid, C. and Ptashne, M. (2002). Telomere looping permits repression "at a distance" in yeast. Current biology *12*, 930–933. 23