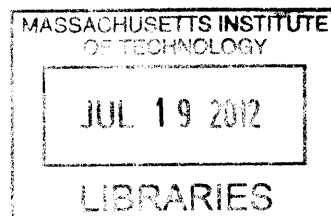


A Regulatory Role for Repeated Decoy Transcription Factor  
Binding Sites in Target Gene Expression

by

Tek Hyung Lee

B.S. Chemical Engineering  
Seoul National University, 2005



SUBMITTED TO THE DEPARTMENT OF CHEMICAL ENGINEERING IN PARTIAL  
FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY IN CHEMICAL ENGINEERING  
AT THE  
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

JUNE 2012

© 2012 Massachusetts Institute of Technology. All rights reserved.

Signature of Author .....

Department of Chemical Engineering  
May 22, 2012

Certified by .....

Narendra Maheshri  
Professor of Chemical Engineering  
Thesis Supervisor

Accepted by .....

Patrick S. Doyle  
Professor of Chemical Engineering  
Chairman, Committee for Graduate Students



# A Regulatory Role for Repeated Decoy Transcription Factor Binding Sites in Target Gene Expression

by

Tek Hyung Lee

Submitted to the Department of Chemical Engineering  
on May 22, 2012 in Partial Fulfillment of the  
Requirements for the Degree of Doctor of Philosophy in  
Chemical Engineering

## ABSTRACT

Repetitive DNA sequences are prevalent in both prokaryote and eukaryote genomes and the majority of repeats are concentrated in *intergenic* regions. These tandem repeats (TRs) are highly variable as the number of repeated units changes frequently due to recombination events and/or polymerase slippage during replication. While TRs have been traditionally regarded as non-functional ‘junk’ DNA, variability in the number of TRs present within or close to genes is known to lead to gross phenotypic changes and disease. However, whether *intergenic* TRs have a functional role is less understood. Recent studies reveal that many *intergenic* TRs contain transcription factor (TF) binding sites and that several TRs of TF binding sites indeed influence gene expression. A possible mechanism is that TRs serve as TF decoys, competing with a promoter for TF binding.

We utilized a synthetic system in budding yeast to examine if repeated binding sites serve as decoys, and alter the expression of genes regulated by the sequestered TF. Combining experiments with kinetic modeling suggests that repeated decoy binding sites sequester activators more strongly than a promoter binding site although both binding sites are identical in sequence. This strong binding converts a graded dose-response between activator and promoter to a sigmoidal-like response. We further find that the tight activator-decoy interaction becomes weaker with increasing activator levels, suggesting that the activator binding at the repeated decoy site array might be anti-cooperative. Finally, we show that the high affinity of repeated decoy sites qualitatively changes the behavior of a transcriptional positive feedback loop from a graded to bimodal, all-or-none response. Taken together, repeated TF binding sites play an unappreciated role as a gene regulator. Since repeated decoy sites are hypervariable in number, this variability can lead to qualitative changes in gene expression and potentially phenotypic variation over short evolutionary time scales.

Thesis Supervisor: Narendra Maheshri  
Title: Assistant Professor of Chemical Engineering





## **Acknowledgements**

First, I thank God for leading me to this good environment and directing each step. Without Him I would not have accomplished my degree successfully.

I wholeheartedly thank my advisor, Prof. Narendra Maheshri, for guiding my graduate study. He always motivated me to approach a problem in various ways and to proceed logically. He also encouraged me to be positive and confident in presenting my research to others. Every discussion with him inspired a new idea and brought valuable insight. He was a good mentor and a wonderful supervisor for me.

I also would like to thank my committee, Prof. Kevin J Verstrepen, Prof. K. Dane Wittrup and Prof. Christopher Voigt for their help and invaluable advice. Prof. Verstrepen's suggestions let me interpret my results more critically and his encouragement brought me great motivation. I also received priceless feedback during my study from Prof. Wittrup. His suggestions expanded my understanding of the project and enabled further progress. Prof. Voigt gave me critical suggestions for the final progress and great encouragement.

In addition, I am deeply grateful for former and current lab members T.L. To, C.J. Zopf, Bradley Niesner, Shawn Finney-Manchester, Nicholas Wren, Katie Quinn and Richard Joh. They gave me useful opinions for my project as well as critical discussion for my paper. I appreciate their help for improving my English speaking and writing skill as well.

I should not forget thanking my other friends including my classmates and Korean friends for their assistance. They helped me to adjust to the MIT life and enjoy it.

Finally, I would like to give thanks to my dad, mom, brother and girlfriend for their love and support. Their encouragement enabled me to continue my study and strengthened me when I was weary.



## Table of Contents

List of Figures and Tables.....	9
1 Introduction.....	11
1.1 Unstable tandem repeats (TRs) are functional.....	11
1.2 Intergenic TRs function as TF decoy sequestering TF and thus influencing target gene expression.....	13
1.3 References.....	16
2 Mathematical and computational modeling for TF sequestration by decoy TRs and target gene expression.....	20
2.1 Introduction.....	20
2.2 Model prediction for TF-decoy binding sites interaction and target gene expression.....	22
2.2.1 Strong TF-decoy sites interactions lead to a sigmoidal-like promoter dose-response.....	22
2.2.2 The stability of TF bound to repeated decoy sites affects the target promoter dose-response.....	27
2.3 Decoy TF binding sites do not alter the variation of promoter state at thermodynamic equilibrium.....	29
2.4 Conclusion.....	30
2.5 References.....	31
3 Investigating the effect of repeated decoy sites on target gene expression.....	32
3.1 Introduction.....	32
3.2 Mathematical model behavior for tTA sequestration by a tetO array and target gene expression.....	33
3.3 Materials and methods.....	38
3.3.1 Strains and plasmids.....	38
3.3.2 Varification of tetO array stability.....	38
3.3.3 Doxycycline (dox) titration.....	39
3.3.4 Methionine titration.....	39
3.3.5 Flow cytometry.....	40
3.3.6 2 $\mu$ plasmid copy number estimation.....	40
3.3.7 Fluorescence microscopy.....	41
3.3.8 Dual color reporter assay.....	41
3.3.9 YFP spot intensity measurement.....	41
3.3.10 Hysteresis assay.....	42
3.3.11 Quantitative chromatin immunoprecipitation (qChIP).....	42
3.4 Results and discussion.....	43
3.4.1 A tetO array converts the dose-response of the tetO promoter from graded to sigmoidal-like, suggesting the array affinity for tTA is higher than the	

promoter affinity.....	43
3.4.2 Split tetO arrays are more potent than a long contiguous tetO array.....	53
3.4.3 <i>in vivo</i> binding assay also shows that the array affinity for tTA is higher than the promoter's but it decreases with increasing tTA levels.....	55
3.4.4 A tetO array converts a graded transcriptional positive feedback response To a bimodal response.....	62
3.4.5 A tetO array does not alter gene expression noise.....	66
3.5 Conclusion.....	69
3.6 References.....	70
4 Overall conclusions and future directions.....	73
4.1 References.....	79
Appendices.....	82
Appendix A. A longer tetO array sequesters more tTA but exhibits a similar potency for lowering the target gene expression compared to a shorter array.....	82
References.....	84
Appendix B. Possible factors influencing tTA-promoter and tTA-array binding.....	85
Proteasome-mediated active disassembling of tTA in the promoter does not affect promoter binding affinity.....	85
The interaction between tTA activation domain and general transcription factors affects tTA-tetO binding affinity.....	87
The difference in chromatin architecture between promoter and decoy array does not affect their binding affinities for tTA.....	88
The interaction between tTA in the promoter and transcription- regulating proteins may not affect tTA-promoter binding.....	90
References.....	94
Appendix C. Yeast strains used in this study.....	97
Appendix D. Plasmids used in this study.....	99

## List of Figures and Tables

Figure 1	A simple model predicts an array of decoy binding sites qualitatively alters the dose-response of a TF and target promoter depending on the strength of the TF-binding site interaction.....	23
Figure 2	Target gene expression depends on the relative degradation rates of unbound and decoy-bound TF.....	28
Figure 3	Arrays of tetO decoy sites reduce target gene expression and convert the graded dose-response between tTA and its target promoter to a sigmoidal-like response.....	46
Figure 4	High copy plasmid-borne tetO arrays also create a sharp sigmoidal-like dose response of the tetO promoter.....	47
Figure 5	The tetO decoy array reduces target gene expression when the tTA synthesis rate is directly varied using a methionine-inducible promoter.....	48
Figure 6	Sensitivity of data fitting to model parameters.....	50
Figure 7	Noncontiguous tetO arrays sequester tTA more effectively than contiguous tetO arrays.....	54
Figure 8	tTA binds the tetO array and the tetO promoter with different strengths....	56
Figure 9	TF occupancy at the 7x and 1x tetO promoters and various tetO arrays....	58
Figure 10	6xtetO array and 7xtetO promoter create the sigmoidal dose response.....	61
Figure 11	Adding tetO decoy sites converts a positive feedback loop from a graded to a switch-like bimodal response.....	63
Figure 12	Addition of a tetO array alters the dose-response of a 1xtetO promoter in a manner similar to the 7xtetO promoter.....	64
Figure 13	The bimodal positive feedback response shifted from ‘off’ state is similar to that from ‘on’ state in the presence of tetO array.....	65
Figure 14	The chromosomally integrated tetO array does not affect noise in gene expression.....	68
Figure 15	tTA-tetO binding affinity may be altered under various environments.....	73
Figure 16	240x tetO array sequesters more tTA than 67x and 113x tetO arrays.....	84
Figure 17	Stabilizing tTA does not eliminate the sigmoidal-like response in the presence of the tetO array.....	86
Figure 18	tetR-YFP binding to tetO array is stronger than tTA-YFP binding at intermediate methionine levels.....	88
Figure 19	Positioning the tetO array near the tetO promoter does not eliminate the sigmoidal-like response.....	90
Figure 20	Sigmoidal-like response for <i>med9</i> and <i>spt8</i> gene disruption mutants.....	92
Figure 21	Sigmoidal-like response for <i>sin4</i> gene disruption mutant.....	93
Table 1	TRs containing TF binding sites.....	13
Table 2	Fit parameter estimates.....	52



## **1. Introduction**

### **1.1 Unstable tandem repeats (TRs) are functional**

The genome of numerous organisms contains repetitive nucleotide sequences. The existence of repeated sequences in the eukaryotic genome was found by DNA reassociation kinetics where repetitive sequences renature more rapidly than unique sequences (Britten et al, 1968). The fraction of repetitive sequences varies with species but is generally higher for eukaryotes than prokaryotes and archaea (Haubold et al, 2006). For example, over 45% of the human genome consists of repeated sequences compared to 7% in bacteria genome (Lander et al, 2001; Treangen et al, 2009). Based on the location of the repeat unit, these repeats are divided into two classes: interspersed repeats (IRs) and tandem repeats (TRs). In IRs the repeat units are scattered across the genome and are believed to originate by movement of transposons between different loci within the genome (Gemayel et al, 2010). On the other hand, the TR units are located adjacent to each other. TRs originate by local duplication (Gemayel et al, 2010). Many eukaryotes contain a significant amount of TRs within their genome. For example, 60% of the *D. nasutooides* (fruit fly) genome and almost half of *D. ordii* (kangaroo rat) genome are TRs (Lee et al, 1978). Though there is no clear definition, TRs are further divided into three subfamilies based on the size of a repeat unit, total length of repeats and number per genome: microsatellite, minisatellite and satellite (Richard et al, 2008).

TRs are highly variable. They expand or shrink rapidly at a rate ranging from  $10^{-2}$  to  $10^{-5}$  per generation, which is generally much faster than point mutation of  $10^{-4}$  to  $10^{-9}$  per generation (Rando et al, 2007). Two models explain the mechanism for the rapid TRs' variation: homologous recombination appearing in intra- or interchromosomal pairing, and strand slippage occurring during DNA replication or double strand break (DSB) repair (Richard et al, 2000; Paques et al, 1998; Gemayel et al, 2010). Since an early study showed TR variability did not bring any phenotypic change, they originally had been regarded as non-functional, "junk" DNA (Ohno, 1972).

However, recent studies implicate TRs in a number of various phenotypes. When TRs occur within the open reading frame of genes, their expansion/contraction directly affects protein structure or expression. For example, TRs within a yeast adhesion gene *FLO1* can influence their adhesive and flocculent properties (Verstrepen et al, 2005), TRs within the *Runx-1* transcription factor (TF) gene in dogs dictate skull morphology (Fondon & Garner, 2004), and changes in TRs number in contingency loci in many prokaryotes switch expression state by introducing frameshift (Rando & Verstrepen, 2007). TRs within intergenic regions adjacent to genes are also widely implicated in affecting gene expression. Expansion of trinucleotide repeats in untranslated regions or introns of genes play a causative role in triplet expansion diseases (Cummings & Zoghbi, 2000) often by silencing gene expression. Recent works in budding yeast demonstrate that TRs within promoters can influence gene expression by altering nucleosome structure or the number of TF binding sites (Vincens et al, 2009).



Compared to these TRs, whether intergenic TRs also affect phenotype is generally not clear. A possible role is that TRs serve as TF decoys and sequester TF, inhibiting its binding to the target promoter and thus affecting gene expression. Recent bioinformatic studies reveal that many intergenic TRs include specific TF binding sites (Horng et al, 2003), and several studies show that TRs of TF binding sites in heterochromatin affect gene expression by sequestering corresponding TFs (Table 1). For example, in mice, the major  $\alpha$ -satellite TRs including many binding sites for C/EBP $\alpha$  reduce the target gene expression recruiting available C/EBP $\alpha$  (Liu et al, 2007).

Table 1. TRs containing TF binding sites

TRs	TF	References
Gamma satellite	Ikaros	Brown et al, 1997; Cobb et al, 2000
Brown AG-rich satellite	GAGA	Platero et al, 1998
Satellite III	HSF1	Jolly et al, 2002
Gamma satellite	YY1	Shestakova et al, 2004
$\alpha$ -satellite	C/EBP $\alpha$	Liu X et al, 2007

## 1.2 Intergenic TRs function as TF decoy sequestering TF and thus influencing target gene expression

There are a number of biological examples where a competitive inhibitor, or decoy molecule, sequesters its target molecule and induce a functional change. The interaction of a ligand molecule with a target receptor can be modulated by the addition of decoy

molecule. For example, vascular epithelial growth factor receptor-1 (VEGFR-1), which is important for normal cellular development and angiogenesis, functions as a decoy receptor, inhibiting VEGF binding to VEGFR-2 (Meyer et al, 2006). Examples of decoys are not restricted to ligand/receptor interactions; other types of decoys are also used for changing gene expression. For example, small RNA (sRNA) or microRNA (miRNA) containing a complementary sequence for its target mRNA can function as a decoy, altering target mRNA stability or activity. (Mardin et al, 2009; Eiring et al, 2010; Mukherji et al, 2011). DNA binding sites can also serve as decoys, recruiting the cognate TF and affecting transcription of target genes. Decoy oligonucleotides containing one or two TF binding sites have been shown to effectively sequester TFs and prevent the gene expression (Morishita R et al 1995) and long repetitive DNA binding sites also showed that they can sequester TFs and affect target gene expression (Table 1).

However, since eukaryote DNA consists of repeated nucleosomes which possibly interfere TF binding for DNA binding sites, how strongly a long tract of repeated binding sites sequesters TF is not clear. In addition, whether entire binding sites within the repeats uniformly interact with TF and what is the relationship between TF-repeats interaction and target gene expression are still elusive. Moreover, whether the TF-decoy interactions change TF stability and/or affect variability in gene expression is also unknown.

To answer these questions, we employ both theoretical modeling and experimental analysis. First, we define the expected relationship between TF-decoy interaction and target gene

expression using a non-equilibrium kinetic model in Chapter 2. This model connects levels of nuclear TF, TF-decoy complex and TF-promoter complex with promoter activation, and yields a relationship (dose-response) between TF level and target gene expression in the presence and absence of decoy DNA sites. We use this model to predict how the size and affinity of decoy binding sites compared to promoter binding sites affects the dose-response behavior and to explore the possible effect of the repeated decoy binding sites on the stability of TF.

In Chapter 3, we experimentally investigate the effect of TF-repeated decoy interaction on target gene expression. Fitting the model behavior to experimental data, we examine how strongly decoy binding sites sequester TF, how the size of decoy sites correlates with TF sequestration and how effectively decoy repeats alter the dose-response of the promoter. Experimental results show that decoy binding sites sequester TF more strongly than promoter binding sites though the TF-decoy binding becomes weaker with greater occupancy of the TF. We also find that this large affinity difference induces a sigmoidal-like promoter dose-response. Next, we explore the possible underlying mechanism to explain the affinity difference between TF-promoter and TF-decoy sites. Finally, we show that repeated decoy sites do not alter variability in gene expression, but can qualitatively change gene network behavior.

Finally, in Chapter 4 we summarize our findings of the regulatory effects of decoy DNA binding sites on gene expression, and discuss future work to understand the basis for the

negative cooperative binding effect observed in binding of a TF to repeated decoy binding sites.

### 1.3 References

Britten, R.J., and Kohne, D.E. (1968). Repeated sequences in DNA. Hundreds of thousands of copies of DNA sequences have been incorporated into the genomes of higher organisms. *Science* 161, 529-540.

Brown, K.E., Guest, S.S., Smale, S.T., Hahm, K., Merkenschlager, M., and Fisher, A.G. (1997). Association of transcriptionally silent genes with Ikaros complexes at centromeric heterochromatin. *Cell* 91, 845-854.

Cobb, B.S., Morales-Alcelay, S., Kleiger, G., Brown, K.E., Fisher, A.G., and Smale, S.T. (2000). Targeting of Ikaros to pericentromeric heterochromatin by direct DNA binding. *Genes Dev.* 14, 2146-2160.

Cummings, C.J., and Zoghbi, H.Y. (2000). Trinucleotide repeats: mechanisms and pathophysiology. *Annu. Rev. Genomics Hum. Genet.* 1, 281-328.

Eiring, A.M., Harb, J.G., Neviani, P., Garton, C., Oaks, J.J., Spizzo, R., Liu, S., Schwind, S., Santhanam, R., Hickey, C.J., et al. (2010). miR-328 functions as an RNA decoy to modulate hnRNP E2 regulation of mRNA translation in leukemic blasts. *Cell* 140, 652-665.

Fondon, J.W., 3rd, and Garner, H.R. (2004). Molecular origins of rapid and continuous morphological evolution. *Proc. Natl. Acad. Sci. U. S. A.* 101, 18058-18063.

Gemayel, R., Vences, M.D., Legendre, M., and Verstrepen, K.J. (2010). Variable tandem repeats accelerate evolution of coding and regulatory sequences. *Annu. Rev. Genet.* 44, 445-477.

Haubold, B., and Wiehe, T. (2006). How repetitive are genomes? *BMC Bioinformatics* 7, 541.

Horng, J.T., Lin, F.M., Lin, J.H., Huang, H.D., and Liu, B.J. (2003). Database of repetitive elements in complete genomes and data mining using transcription factor binding sites. *IEEE Trans. Inf. Technol. Biomed.* 7, 93-100.

Jolly, C., Konecny, L., Grady, D.L., Kutsikova, Y.A., Cotto, J.J., Morimoto, R.I., and Vourc'h, C. (2002). In vivo binding of active heat shock transcription factor 1 to human chromosome 9 heterochromatin during stress. *J. Cell Biol.* 156, 775-781.

Lander, E.S., Linton, L.M., Birren, B., Nusbaum, C., Zody, M.C., Baldwin, J., Devon, K., Dewar, K., Doyle, M., FitzHugh, W., et al. (2001). Initial sequencing and analysis of the human genome. *Nature* 409, 860-921.

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

Mandin, P., and Gottesman, S. (2009). Regulating the regulator: an RNA decoy acts as an OFF switch for the regulation of an sRNA. *Genes Dev.* 23, 1981-1985.

Meyer, R.D., Mohammadi, M., and Rahimi, N. (2006). A single amino acid substitution in the activation loop defines the decoy characteristic of VEGFR-1/FLT-1. *J. Biol. Chem.* 281, 867-875.

Morishita, R., Gibbons, G.H., Horiuchi, M., Ellison, K.E., Nakama, M., Zhang, L., Kaneda, Y., Ogihara, T., and Dzaou, V.J. (1995). A gene therapy strategy using a transcription factor decoy of the E2F binding site inhibits smooth muscle proliferation in vivo. *Proc. Natl. Acad. Sci. U. S. A.* 92, 5855-5859.

Mukherji, S., Ebert, M.S., Zheng, G.X., Tsang, J.S., Sharp, P.A., and van Oudenaarden, A. (2011). MicroRNAs can generate thresholds in target gene expression. *Nat. Genet.* 43, 854-859.

Ohno, S. (1972). So much "junk" DNA in our genome. *Brookhaven Symp. Biol.* 23, 366-370.

Paques, F., Leung, W.Y., and Haber, J.E. (1998). Expansions and contractions in a tandem repeat induced by double-strand break repair. *Mol. Cell. Biol.* 18, 2045-2054.

Platero, J.S., Csink, A.K., Quintanilla, A., and Henikoff, S. (1998). Changes in chromosomal localization of heterochromatin-binding proteins during the cell cycle in *Drosophila*. *J. Cell Biol.* 140, 1297-1306.

Rando, O.J., and Verstrepen, K.J. (2007). Timescales of Genetic and Epigenetic Inheritance. *Cell* 128, 655-668.

Richard, G.F., and Paques, F. (2000). Mini- and microsatellite expansions: the recombination connection. *EMBO Rep.* 1, 122-126.

Shestakova, E.A., Mansuroglu, Z., Mokrani, H., Ghinea, N., and Bonnefoy, E. (2004). Transcription factor YY1 associates with pericentromeric gamma-satellite DNA in cycling but not in quiescent (G0) cells. *Nucleic Acids Res.* 32, 4390-4399.

Treangen, T.J., Abraham, A.L., Touchon, M., and Rocha, E.P. (2009). Genesis, effects and fates of repeats in prokaryotic genomes. *FEMS Microbiol. Rev.* 33, 539-571.

Verstrepen, K.J., Jansen, A., Lewitter, F., and Fink, G.R. (2005). Intragenic tandem repeats generate functional variability. *Nat Genet* 37, 986-990.

Vinces, M.D., Legendre, M., Caldara, M., Hagihara, M., and Verstrepen, K.J. (2009). Unstable tandem repeats in promoters confer transcriptional evolvability. *Science* 324, 1213-1216.

## **2. Mathematical and computational modeling for TF sequestration by decoy TRs and target gene expression**

### **2.1 Introduction**

The addition of decoys in the context of signaling and gene regulatory networks can lead to qualitative changes in the dose-response of modules within these networks, depending on the details of interactions. A previous analysis explains how simple binding of a repressor, acting as a decoy, to a signaling protein can qualitatively change the signaling response. Normally, there is a linear relationship between the total signaling protein's levels and the downstream response. However, with the decoy the slope of that relationship will decrease, provided the protein-decoy complex binding affinity is lower than the protein concentration. However, if the complex binds at high affinity, active, unbound protein levels remain low while total protein levels are increased, because the protein-decoy complex is strongly favored. When the decoy proteins are completely saturated, then additional increase in the amount of total signaling protein leads to a dramatic increasing in the active protein level, creating an ultrasensitive response. Another recent study extends this idea to interactions between TF and decoy protein (Buchler et al, 2009). It shows that when the binding between active TF and decoy protein is stronger than TF-promoter binding, the promoter activation becomes ultrasensitive with increasing TF levels because TF decoy protein sequester the TF up to its saturation, preventing TF binding to the promoter. Here, we further extend these theoretical studies into an interaction between TF and decoy DNA



binding sites and show that decoy repeats can also induce a sigmoidal-like promoter dose-response if repeated binding sites sequester TF more strongly than the promoter.

The previous study confines the decoy protein affinity for TF, which is higher than the promoter affinity, to be constant and to not alter with TF or decoy protein level (Buchler et al, 2009). However, in the context of TF-decoy DNA binding, whether clustered binding sites also possess a fixed and uniform binding affinity for TF and whether those binding strengths are identical to the promoter binding affinity are not clear. Therefore, our mathematical model introduces ensemble averaged parameters: “averaged” binding affinity and “effective” number of decoy binding sites, which does not reflect an actual number of binding sites.

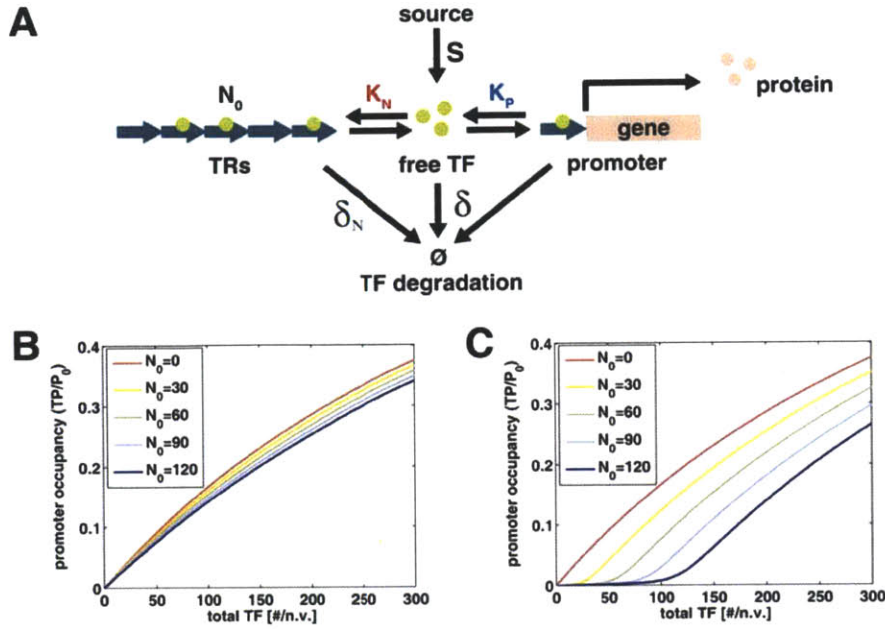
The decoy protein study also suggests that the stability of the decoy-active protein complex affects an induction of the ultrasensitive response of the free active protein level (Buchler et al, 2008). Whether repeated decoy binding sites stabilize TF from its degradation and, if so/not, how TF stabilization/destabilization alters the promoter response is unknown. To answer these questions, here, we theoretically explore possible effects of decoy DNA repeats on both TF stability and promoter dose-response and find the sigmoidal dose-response of the promoter appears once decoy repeats sequester TF more strongly than the promoter regardless of TF stability.

When the decoy DNA binding sites completely protect bound TF, it reduces the fluctuation of TF level in the transcriptional positive feedback system (Burger et al, 2010). However, whether these repeated decoy sites affect the variation of TF level and gene expression level in general TF regulatory system is not clear. To answer this question, we employ a stochastic simulation and find decoy sites of high TF binding affinity increase the fluctuation of TF level but not mRNA and protein level when the TF-decoy repeats interaction time scale is faster than mRNA lifetime.

## **2.2 Model prediction for TF-decoy binding sites interaction and target gene expression**

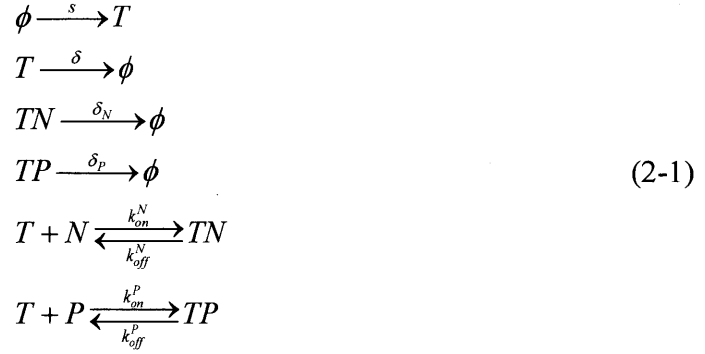
### **2.2.1 Strong TF-decoy sites interactions lead to a sigmoidal-like promoter dose-response**

To examine how repeated decoy sites affect target gene expression, we build a mathematical model to describe a dose-response between TF and promoter in the presence of decoy repeats based on non-equilibrium binding. Here, we consider the chemical transformations diagrammed in Figure 1A below.



**Figure 1. A simple model predicts an array of decoy binding sites qualitatively alters the dose-response of a TF and target promoter depending on the strength of the TF-binding site interaction.** (A) A simple model to describe the effects of TRs containing decoy binding sites on target gene expression. Important parameters include the number ( $N_0$ ) and binding affinity ( $1/K_N$ ) of decoy sites, the binding affinity ( $1/K_P$ ) of promoter binding sites, and production and degradation of each species. (B) Model predicted dose-response of expression versus total TF level,  $T_0$ , for various numbers,  $N_0$ , of decoy sites when the binding affinity of the TF for decoy and promoter sites are identical ( $K_N = K_P = 500$ ). Decoy sites reduce expression but do not change the graded nature of the response. (C) As in (B) but with promoter binding affinity set much lower than decoy binding affinity ( $K_N = 1$ ,  $K_P = 500$ ), which results in a more sigmoidal-like dose-response.

Each reaction is stated as:



The TF ( $T$ ) is a transcriptional activator which is produced constitutively and can potentially bind either decoy sites ( $N$ ) or the promoter ( $P$ ). Species balances for free and bound forms are:

$$\begin{aligned}
T_0 &= T + TP + TN \\
N_0 &= N + TN \\
P_0 &= P + TP
\end{aligned} \tag{2-2}$$

While not explicitly shown here, the DNA corresponding to decoys ( $N$ ) and the promoter ( $P$ ) are also synthesized and diluted as cells grow. These processes and the synthesis and degradation of the unbound TF ( $T$ ) are slow compared to fast binding and unbinding of the TF to the promoter or decoy sites (order 10's of minutes versus 10's of second). For now, we assume that decoy and promoter-bound TFs ( $TN$  and  $TP$ ) degrade at rates identical to the unbound TF ( $T$ ).

The differential equations describing the chemical transformations of equation (2-1) are:

$$\begin{aligned}
\frac{dT}{dt} &= S + k_{off}^P(TP) + k_{off}^N(TN) - k_{on}^P(T)(P) - k_{on}^N(T)(N) - \delta(T) \\
\frac{d(TP)}{dt} &= k_{on}^P(T)(P) - k_{off}^P(TP) - \delta_P(TP) \\
\frac{d(TN)}{dt} &= k_{on}^N(T)(N) - k_{off}^N(TN) - \delta_N(TN)
\end{aligned} \tag{2-3}$$

At steady state, the latter two equations in (2-3) results in

$$\begin{aligned}
K_P &= \frac{(T)(P)}{(TP)} = \frac{k_{off}^P + \delta_P}{k_{on}^P} \approx \frac{k_{off}^P}{k_{on}^P} \\
K_N &= \frac{(T)(N)}{(TN)} = \frac{k_{off}^N + \delta_N}{k_{on}^N} \approx \frac{k_{off}^N}{k_{on}^N}
\end{aligned} \tag{2-4}$$

Generally there is complete recovery of the FRAP curve on the minutes timescale (Karpova et al, 2008; Kumar et al, 2010). Therefore, while  $k_{off}^N > \delta_N$  it is possible they are of the same order of magnitude. We only estimate  $K_N$ , but  $\delta_N$  sets a upper bound on the residence time of strongly bound tTA of order 10 minutes. Promoter binding is probably reflected by the FRAP experiments, so  $k_{off}^P$  is much greater than the degradation rate of the promoter/TF complex,  $TP$ . Whether the large difference between  $K_N$  and  $K_P$  is solely due to differences in off rate (as might be expected since the on rate is thought to be diffusion controlled) is not clear. The residence times (off rates) in (2-4) in no way have to be governed by thermodynamics.

Combining (2-3) and (2-4) with the species balance (2-2) leads to:

$$\frac{S}{\delta} = T + \frac{\delta_N}{\delta}(TN) + \frac{\delta_P}{\delta}(TP) \approx T + \left( \frac{\delta_N}{\delta} \right) (N_0) \frac{T}{T + K_d^N} \tag{2-5}$$

Since the number of activator-promoter complexes ( $TP$ ) is much smaller than the number of free activators ( $T$ ) and activator-decoy complexes ( $TN$ ), the total number of activators is approximated as the sum of the numbers of free activator and activator-decoy complex.

Notice that only if  $\delta_N/\delta = 1$  does the total TF level  $T_0 = T + TN = S/\delta$ .

By using (2-5) we obtain

$$\frac{T}{T_0} = \frac{\left(1 - \frac{K_N}{T_0} - \left(\frac{\delta_N}{\delta}\right)\left(\frac{N_0}{T_0}\right)\right) + \sqrt{\left(1 - \frac{K_N}{T_0} - \left(\frac{\delta_N}{\delta}\right)\left(\frac{N_0}{T_0}\right)\right)^2 + 4\frac{K_N}{T_0}}}{2} \quad (2-6)$$

if  $\delta_N/\delta = 1$

$$\frac{T}{T_0} = \frac{\left(1 - \frac{K_N}{T_0} - \frac{N_0}{T_0}\right) + \sqrt{\left(1 - \frac{K_N}{T_0} - \frac{N_0}{T_0}\right)^2 + 4\frac{K_N}{T_0}}}{2} \quad (2-7)$$

We assume gene expression is proportional to  $TP/P_0$ , the TF occupancy at the promoter (Bintu et al, 2005):

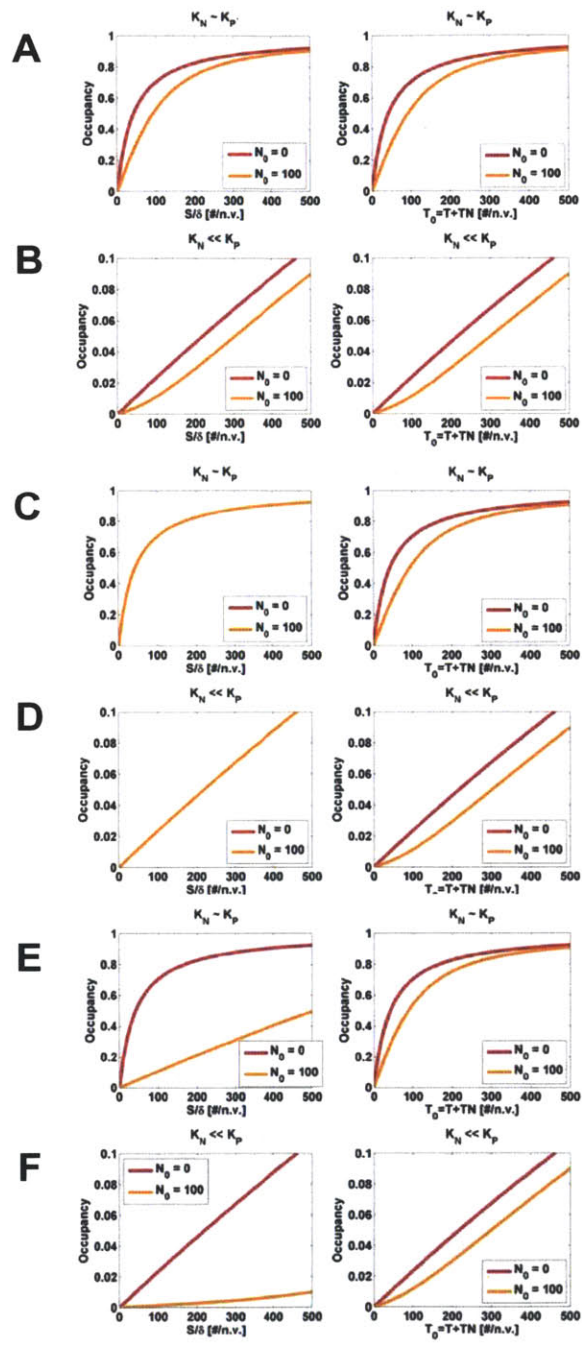
$$\frac{(TP)}{P_0} = \frac{\frac{T}{T_0}}{\frac{T}{T_0} + \frac{K_p}{T_0}} \quad (2-8)$$

We will employ concentration units of molecules per yeast nuclear volume (n.v.). Using equations (2-7) and (2-8), we plot the dose-response of TF occupancy to  $T_0 = T + TN$  when  $T_0$  is varied by changing the synthesis rate,  $S$ . Because  $T_0 = S/\delta$ , choosing to plot TF

occupancy versus  $T_0$  or  $S$  are equivalent within a constant. When  $K_N/K_P = 1$ , increasing decoy number  $N_0$  decreases target gene expression but does not change the shape of the dose-response curve (Figure 1B and Figure 2A). However, the shape changes when decoy sites have much higher affinity ( $K_N/K_P \ll 1$ ). As TF level increases, they bind and saturate decoy sites before leading to gene expression (Figure 1C and Figure 2B). The result is a sharper, threshold dose-response that has previously been discussed (Buchler & Louis, 2008).

### **2.2.2 The stability of TF bound to repeated decoy sites affects the target promoter dose-response**

The strength of TF-decoy repeats binding may be affected by the stability of TF bound decoy sites. We consider the non-equilibrium effects of varying  $\delta_N/\delta$  assuming  $\delta_N = \delta_P$ . In the first case where decoy-bound TF is protected from degradation and  $\delta_N/\delta \ll 1$ , when  $K_N/K_P = 1$ , faster turnover of unbound  $T$  decouples its level from the decoy-bound  $TN$  species, making it invariant to changes in decoy number  $N_0$  (Burger et al, 2010). But, increasing  $N_0$  does affect  $T_0$  as the decoy-bound species  $TN$  increases at any given synthesis rate  $S$ . Therefore, the dose-response curve of TF occupancy versus  $T_0$  is altered when decoys are added, whereas TF occupancy never changes with  $S$  (Figure 2C). TF occupancy also never changes with  $S$  when  $K_N/K_P \gg 1$ , although now there is the sharper, threshold dose-response versus  $T_0$  (Figure 2D). In the second case where  $\delta_N/\delta \gg 1$ , the dose-response is nearly identical to the case of  $\delta_N = \delta_P$ , but much larger changes in  $S$  are required to increase occupancy as the decoy-bound species degrades quickly (Figure 2E&F).





**Figure 2. Target gene expression depends on the relative degradation rates of unbound and decoy-bound TF.** (A,B) Model predictions of the TF occupancy at the target promoter ( $TP/P_0$ ) versus (*left*) the TF synthesis rate divided by the unbound TF degradation rate,  $S/\delta$ , and (*right*) the total unbound plus decoy-bound TF level  $T_0 = T + TN$ , when both the unbound and decoy-bound TF degrade at the same rate ( $\delta_N \sim \delta$ ). In this case,  $S/\delta$  is equivalent to  $T_0$  and so both plots on the left and right are identical. This is true when both (A)  $K_N/K_P \sim 1$  and the dose-response is graded or (B)  $K_N/K_P \ll 1$  and the dose-response is more sigmoidal, with positive concavity. Plots in (A) were generated with  $k_{on}^P = 0.01$  [#n.v./min],  $k_{off}^P = 0.42$  [/min],  $k_{on}^N = 0.01$  [/(#n.v.)/min],  $k_{off}^N = 0.42$  [/min],  $N_0 = 100$ ,  $\delta = 0.0069$  [/min] and  $\delta_N = 0.0069$  [/min]. Identical parameters for (B) were used except  $k_{off}^P = 42$  [/min]. (C, D) As in (A, B) but now when decoy-bound TF is protected and degrades more slowly than free TF ( $\delta_N \ll \delta$ ). In this case, the (*left*) both the free TF level at steady state (corresponding to  $S/\delta$ ) and promoter occupancy is unaffected by decoy number, regardless of the ratio of  $K_N/K_P$ . However, promoter occupancy does depend on total TF number in a manner similar to (A, B) for both (C)  $K_N/K_P \sim 1$  and (D)  $K_N/K_P \ll 1$ . Plots in (C, D) were generated with parameters as in (A, B), respectively, but with  $\delta_N = 0$  [/min] were used. (E, F) If decoy-bound TF degrades faster than unbound TF ( $\delta_N \gg \delta$ ), (*left*) the free TF level is no longer given by  $S/\delta$  but is lower in the presence of decoys because they increase the overall TF degradation rate. The response versus total TF remains identical. Plots (E, F) were generated with parameters in (A, B), respectively, except  $\delta_N = 0.069$  [/min].

### 2.3 Decoy TF binding sites do not alter the variation of promoter state at thermodynamic equilibrium

If gene expression is proportional to promoter occupancy by TF and if lifetime scales of mRNA and protein are similar to or less than the timescale of TF binding and unbinding, promoter state fluctuations will affect the variation in gene expression. If we add repeated decoy sites in this system, can the decoy sites alter the promoter state fluctuations? When

TF-promoter binding is at thermodynamic equilibrium where the total TF number is fixed, the promoter state fluctuations are expressed as the following.

$$\frac{\sigma_{TP}^2}{\langle TP \rangle^2} = \frac{\frac{T}{T + K_p} \frac{K_p}{T + K_p}}{\left( \frac{T}{T + K_p} \right)^2} = \frac{K_p}{T} \quad (2-9)$$

Where  $\sigma_{TP}$  denotes the standard deviation of promoter state,  $TP$ . Here, the free TF number determines the promoter state fluctuation. The decoy binding sites, which exhibit an identical affinity to the promoter, will reduce free TF level by sequestering TF and thus increase the noise in the promoter state. However, if we increase total TF level in the presence of decoy sites to keep the free TF level fixed, the promoter state noise will stay identical because the noise is determined by free TF number. This might not be true in the case that the decoy sites' affinity is much higher than the promoter affinity and the fluctuations in TF binding and unbinding to the decoy sites is much slower than the cell cycle. Because the system will not reach an equilibrium state, we will observe a significant increase in the promoter state noise.

## 2.4 Conclusion

To understand how intergenic repeated decoy TF binding sites influence target gene expression, we theoretically examined a dose-response of the promoter using a kinetic model. When the promoter binding affinity for TF is similar to the decoy sites' binding affinity, the dose-response is graded regardless of the number of the decoy sites. However, this graded response is switched to a sharp, threshold response when the decoy sites

sequester TF more strongly than the promoter. Our results also revealed that this sigmoidal-like dose-response behavior can be observed regardless of the stability of the TF-decoy binding complex. Finally, our analysis predicts that at thermodynamic equilibrium the decoy binding sites do not alter the promoter state noise at a fixed promoter state.

## 2.5 References

- Bintu, L., Buchler, N.E., Garcia, H.G., Gerland, U., Hwa, T., Kondev, J., and Phillips, R. (2005). Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.* 15, 116-124.
- Buchler, N.E., and Cross, F.R. (2009). Protein sequestration generates a flexible ultrasensitive response in a genetic network. *Mol. Syst. Biol.* 5, 272.
- Buchler, N.E., and Louis, M. (2008). Molecular titration and ultrasensitivity in regulatory networks. *J. Mol. Biol.* 384, 1106-1119.
- Burger, A., Walczak, A.M., and Wolynes, P.G. (2010). Abduction and asylum in the lives of transcription factors. *Proc. Natl. Acad. Sci. USA* 107, 4016-4021.
- Karpova, T.S., Kim, M.J., Spriet, C., Nalley, K., Stasevich, T.J., Kherrouche, Z., Heliot, L., and McNally, J.G. (2008). Concurrent fast and slow cycling of a transcriptional activator at an endogenous promoter. *Science* 319, 466-469.
- Kumar, M., Mommer, M.S., and Sourjik, V. (2010). Mobility of cytoplasmic, membrane, and DNA-binding proteins in *Escherichia coli*. *Biophys. J.* 98, 552-559.

### **3. Investigating the effect of repeated decoy sites on target gene expression**

#### **3.1 Introduction**

Intergenic TRs containing TF binding sites may serve as a competitive inhibitor, or decoy, sequestering TF. The number of Decoy binding sites have been implicated affecting fruit fly morphology (Janssen et al, 2000). However, a clear understanding of how number and relative position of these binding sites affects gene expression is lacking. Since TRs are highly variable, it is easy to imagine how the numbers of repeated decoy binding sites could vary over short evolutionary time scales, and perhaps lead to qualitative changes in gene expression and ultimately an organism's phenotype and fitness.

The repeated decoy binding sites for TF are often identical to a promoter binding site in sequence. Assuming repeated decoy binding sites have sequences nearly identical to those found in promoters, we hypothesize that they will affect gene expression by sequestering TF as discussed in Chapter 2. Specifically, 1) the decoy repeats possibly lower promoter activation without altering a graded nature of promoter response and 2) the reduction of gene expression increases with the number of decoy sites (Figure 2B). However, a clustered feature of the decoy sites, chromatin architecture and other unknown interactions between decoy sites-bound activator and another protein, which possess an affinity for the activator, may alter activator's binding to the decoy sites and thus affect the promoter response. To test these hypotheses, we utilize a synthetic tet-OFF system adapted for budding yeast (Gari et al, 1997). Briefly, tetracycline controlled transactivator (tTA) binds a tetracycline

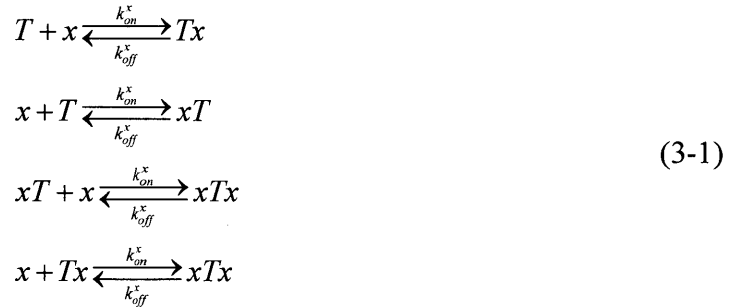
operator (tetO) site within the tetO promoter and activates downstream gene expression. We control the tTA level by using doxycycline (dox), which tightly binds to tTA and prevent its binding to the tetO site. We introduce various arrays of tetO binding sites of various number to serve as decoys, and measure promoter response by monitoring a downstream fluorescent reporter expression (Figure 3A). By using this synthetic system with a corresponding mathematical model, we examine how strongly and effectively the repeated decoy binding sites sequester activators and alter the promoter dose-response.

### **3.2 Mathematical model behavior for tTA sequestration by a tetO array and target gene expression**

For an accurate picture of the dose-response curve between tTA and fluorescent reporter expression in tet-OFF system, we use a kinetic model built in Chapter 2 with slight modification to connect to our particular experimental setup. As explained above, the constitutively produced tTA is titrated by dox. To obtain a relationship between total tTA level and reporter expression, we need to translate an experimentally set external dox concentration to the active tTA level. To do so, we extended a previously reported and experimentally verified model of the dox-tTA interaction (Murphy et al, 2007; To & Maheshri, 2010).

Key features of this model are 1) a constant flux of dox enters cells resulting in the intracellular dox concentration being linearly proportional to the external dox concentration, 2) two dox molecules bind each tTA dimer in a non-cooperative manner to abrogate its

DNA binding capability, and 3) free, promoter-, decoy-, and dox-bound tTA equilibrate on timescales faster than tTA degradation ( $\sim 15$  minute half-life – (To & Maheshri, 2010)). To incorporate interactions of the tetR dimer in tTA with dox following chemical transformations are added to (2-1), where  $x$  represents intracellular dox:



Differential equations describing both dox and decoy interactions are:

$$\begin{aligned}
 \frac{dT}{dt} &= S - \delta(T) + k_{off}^x(Tx) + k_{off}^x(xT) - 2k_{on}^x(T)(x) + k_{off}^N(TN) - k_{on}^N(T)(N) \\
 \frac{d(Tx)}{dt} &= k_{on}^x(T)(x) - k_{off}^x(Tx) + k_{off}^x(xTx) - k_{on}^x(Tx)(x) - \delta(Tx) \\
 \frac{d(xT)}{dt} &= k_{on}^x(T)(x) - k_{off}^x(xT) + k_{off}^x(xTx) - k_{on}^x(xT)(x) - \delta(xT) \\
 \frac{d(xTx)}{dt} &= k_{on}^x(Tx)(x) + k_{on}^x(xT)(x) - 2k_{off}^x(xTx) - \delta(xTx) \\
 \frac{d(TN)}{dt} &= k_{on}^N(T)(N) - k_{off}^N(TN) - \delta(TN) \\
 \frac{d(x)}{dt} &= Fx_{ext} - \delta_x(x) + k_{off}^x(Tx) + k_{off}^x(xT) - 2k_{on}^x(T)(x) + 2k_{off}^x(xTx) - \dots \\
 &\quad k_{on}^x(Tx)(x) - k_{on}^x(xT)(x) + \delta(Tx) + \delta(xT) + 2\delta(xTx)
 \end{aligned} \tag{3-2}$$

Here,  $Fx_{ext}$  is the net transport rate of dox into the cell and is proportional to the external dox concentration,  $x_{ext}$ . Dox is not degraded but lost by dilution due to growth. When dox-bound  $T$  species are degraded, the dox returns to the pool of free dox. We also assume dox-

bound tTA degrades at a rate identical to free tTA. The species balance yields the total amount of each dox species in its free and bound forms:

$$x_0 = x + xT + Tx + xTx \quad (3-3)$$

Only the free ( $T$ ) species is capable of binding the decoy or promoter sites. Combining differential equations in (3-2) with the species balance in (3-3) results in a relationship between the external and intracellular dox concentration:

$$\frac{dx_0}{dt} = Fx_{ext} - \delta_x x_0 \quad (3-4)$$

The key attribute of this model is the external dox concentration sets the total intracellular dox ( $x_0$ ) concentration at the steady-state:  $x_0 = Fx_{ext} / \delta_x$ . We define  $K_M = F / \delta_x$  as a lumped parameter akin to a membrane partition coefficient for dox. Then, combining this definition with (3-2) and (3-3) in the absence of decoy sites ( $N_0 = 0$ ), at the steady-state we arrive at following expressions

$$\frac{T}{T_0} = \frac{\frac{x_0}{T_0} - \frac{x}{T_0}}{2 \frac{x}{K_s} \left( 1 + \frac{x}{K_s} \right)} \approx \frac{\frac{x_0}{T_0}}{2 \frac{x}{K_s} \left( 1 + \frac{x}{K_s} \right)} \quad (3-5)$$

$$\frac{T}{T_0} = \left( 1 + 2 \frac{x}{K_s} + \left( \frac{x}{K_s} \right)^2 \right)^{-1} \quad (3-6)$$

Here,  $K_S = k_{off}^x / k_{on}^x$  is the thermodynamic affinity of dox to tTA. We do not fit data to the first expression in (3-5), because this introduces  $T_0$  as an additional free parameter. In fact, the second expression is an excellent approximation. Because dox binding is strong ( $K_S = 0.21$

dox molecules / n.v; Degenkolb et al, 1991), the actual value of  $T_0$  will not affect  $T/T_0$  provided  $T_0 \gg 1$  molecule / n.v. In other words,  $x/T_0 \ll 1$  over the range we fit, because free dox ( $x$ ) is always at low levels when there is expression. Therefore this term can be neglected, leading to the approximate expression.

It is worth noting that transport of doxycycline out of the cell at a rate proportional to the *intracellular* dox concentration is not described by our model and leads to different behavior. Modifying equation (3-4) for this case:

$$\frac{dx_0}{dt} = Fx_{ext} - \delta_x x_0 - F_{out} x \quad (3-7)$$

It is easy to see that now  $x_{ext}$  does not set  $x_0$ . In the limit that  $F_{out} \gg \delta_x$ ,  $x_{ext}$  sets  $x$ . In other words, depending on the relative value of  $F_{out} / \delta_x$ , the external dox concentration either sets the *total* intracellular dox level or the *unbound* intracellular dox level. We explored both extremes and found our data was better described by the fixed total intracellular dox model. Nevertheless, direct measurement of efflux rates could aid in improving the model.

We use the following expression, based on equation (2-8), to relate  $T/T_0$  to fluorescent reporter expression:

$$FP - FP_{\min} = (FP_{\max} - FP_{\min}) \times \frac{(TP)}{P_0} = k_{\max} \frac{\frac{T}{T_0}}{\frac{T}{T_0} + \frac{K_P}{T_0}} \quad (3-8)$$



$FP$  is the measured fluorescent reporter expression,  $FP_{min}$  is the basal expression in the absence of TF,  $FP_{max}$  is the maximum expression, and  $k_{max} = FP_{max} - FP_{min}$ . In previous work, we have established this model, with a “Hill coefficient” of 1, for the 7xtetO promoter (To & Maheshri, 2010). We can estimate the  $FP_{max}$  by measuring the expression of the promoter in positive feedback, and  $FP_{min}$  by measuring expression in strains without tTA or subject to very high dox levels (To & Maheshri, 2010). The CFP/YFP fluorescent signals reported are normalized with respect CFP/YFP signals measured in a yeast strain constitutively expressing the fluorescent protein from an *ADHI* promoter integrated at the *LEU2* locus. This allows direct comparison of fluorescent signals irrespective of fluorophore and method of measurement. Finally, when tTA expression is driven from the weak *MYO2* promoter, steady-state levels ( $T_0$ ) are low enough that  $T_0 < K_P$  and the dose-response is always in the linear range (To & Maheshri, 2010). This is confirmed when we use equations (3-5), (3-6), and (3-8) to fit the dose-response data in the absence of decoys, by varying two free parameters,  $K_P/T_0$  and  $K_M/T_0$ .

Equation (3-6) can be modified to fit the dose-response curves in the presence of decoy arrays:

$$\frac{T}{T_0} = \left( 1 - \left( \frac{N_0}{T_0} \right) \frac{T/T_0}{T/T_0 + K_N/T_0} \right) \left( 1 + 2 \frac{x}{K_s} + \left( \frac{x}{K_s} \right)^2 \right)^{-1} \quad (3-9)$$

Using (3-5), (3-8), and (3-9) with estimated values for  $K_M/T_0$  and  $K_P/T_0$ , we fit two new parameters:  $N_0/T_0$  and  $K_N/T_0$  (Table 2)

### **3.3 Materials and methods**

#### **3.3.1 Strains and plasmids**

*S. cerevisiae* synthetic tet-OFF system was built in W303 background using standard methods of yeast molecular biology (Thomas & Rothstein, 1989; Guthrie & Fink, 2004). Details of the tTA and tetO promoters are given in (To & Maheshri, 2010). Briefly, tTA gene driven by constitutive *MYO2* promoter was placed in *ADE2* locus and 1x or 7x tetO promoter driving yellow/cyan fluorescent protein reporter (YFP/CFP) was integrated into *URA3* locus. Various number of tetO arrays (15x, 37x, 67x, 113x, 127x and 240x tetO) were derived from a 9kb non-recombinogenic tetO array containing 240xtetO binding sites spanned by 10 or 30bp of random DNA sequence (a kind gift of D. Sherrat) (Lau et al, 2003) and located in *HIS3*, *TRP1*, *URA3* loci or on centromeric/high copy 2 $\mu$  plasmids.

#### **3.3.2 Verification of tetO array stability**

The tetO arrays built from a non-recombinogenic 240x tetO array (Lau et al, 2003) include 10 or 30 random nucleotides between each tetO unit, thereby significantly reducing the possibility of homologous recombination between tetO units and preventing uncontrolled expansion or contraction of the array. To confirm the stability of this synthetic tetO array, yeast centromeric plasmids containing various lengths of tetO arrays were extracted from cells grown for at least 10 generations during which time they had been subject to doxycycline titration and measurement. The size of the tetO array was verified by

restriction analysis using sites both within and flanking the array. Over 97% plasmids showed that the tetO size was not altered (data not shown).

### **3.3.3 Doxycycline (dox) titration**

Yeast cells were grown in synthetic medium with 2% glucose overnight at 30°C, then diluted ( $OD_{600} \sim 0.01-0.05$ ) and grown in the same medium with various concentrations of doxycycline (Sigma) in a 96 deep well plate for at least 8 hours, maintaining exponential phase. Then, cells were diluted again ( $OD_{600} \sim 0.01-0.05$ ) and grown for at least 8 hours to insure reporter expression reached the steady state. After incubation, cells were placed on ice or at 4°C and fluorescence intensities were measured by flow cytometry.

### **3.3.4 Methionine titration**

To titrate tTA expression driven from the methionine-inducible *MET3* promoter, yeast colonies were grown in synthetic medium with 20 mg/L methionine and 2% glucose overnight at 30°C, then diluted to low  $OD_{600}$  (0.01-0.05) and grown in the same medium with various concentrations of methionine (10-100 mg/L) in a 96 deep well plate for at least 3 days. Cells were maintained in exponential phase for steady-state expression of NLS-tTA-YFP by repeatedly diluting cultures to keep the  $OD_{600}$  below 1. The expression of fluorescence proteins (NLS-tTA-YFP, CFP, and RFP) was measured by flow cytometry.

### **.3.5 Flow cytometry**

Analytical flow cytometry on yeast cells were performed using a Beckton-Dickinson (BD) LSRII HTS equipped with a 405 nm laser and 450/50 nm filter (CFP), a 488 nm laser and 530/30 nm filter (YFP) and a 561 nm laser and 610/20 nm (RFP) filter. For each sample, at least 30,000 cells were measured. Yeast cells without fluorescent reporters or a strain constitutively expressing YFP or CFP from an *ADHI* promoter were always used as negative and positive controls, respectively. This enabled normalization and comparison of the YFP or CFP intensity from measurements performed on different days. Reported data includes the densest region of a forward versus side scatter plot of analyzed cells, representing 15% of population.

### **3.3.6 2 $\mu$ plasmid copy number estimation**

To identify cells with a particular copy number or 2 $\mu$  high copy plasmid, a tdTomato red fluorescent protein (RFP) driven from the constitutively active *ACT1* promoter was incorporated on the same 2 $\mu$  plasmid on which the tetO array was present. These cells were grown with varied amounts of doxycycline over 16 hours maintaining a exponential phase. The broad distribution of RFP, which is due to plasmid number fluctuation on timescales close to the cell growth rate, was compared to the RFP expression from a strain containing a chromosomally integrated *ACT1* promoter driving RFP. Specifically, cells were binned into subpopulations where the average RFP level in each bin is an integer multiple of the RFP expressed from the chromosomally integrated RFP gene and cells in each subpopulation were regarded to contain a given integer number of plasmids. The cells

expressing averaged RFP lower than RFP level driven by *ACT1* promoter were regarded to lose the plasmid.

### **3.3.7 Fluorescence microscopy**

Fluorescence microscopy was performed using a high NA (1.4) 100X objective with a fully motorized Zeiss Axioobserver. The excitation source was the Lambda LS xenon lamp (Sutter), and fluorescence imaging was done using a triple band pass dichroic and appropriate filters (500/535 YFP, 430/470 CFP) (Chroma Technologies). Images were captured using a Cascade II EMCCD (Photometrics) camera and Metamorph software (Molecular Devices).

### **3.3.8 Dual color reporter assay**

To measure gene expression noise (total noise and intrinsic noise), yeast diploid cells harboring 7xtetO YFP and 7xtetO CFP placed at *URA3* locus of each allele were grown in synthetic medium with 2% glucose overnight at 30°C at various dox concentrations, maintaining exponential phase, and fluorescent reporter expression was measured by fluorescence microscopy. Fluorescent protein expression data were processed in MATLAB (MathWork).

### **3.3.9 YFP spot intensity measurement**

Yeast haploid cells expressing NLS-tTA-YFP fusion under *MET3* promoter were grown in synthetic medium with 2% glucose at various methionine levels at 30°C for at least 16hr,

maintaining an exponential phase. To measure the number of cells with these spots and their intensity by microscopy, 12 z-stack images with 0.3 $\mu$ m separation were acquired and further processed by custom image processing routines written in MATLAB (MathWorks). Briefly, we first selected the image containing the brightest pixel for each cell and smoothed it with a Gaussian filter to minimize noise. Then for every individual cell we defined its brightest pixel and the 8 pixels around it as the spot region. We then calculated the spatial 2<sup>nd</sup> derivative across the spot, which we found to easily discriminate between a bona fide spot (which had a high spatial derivative) and a false positive.

### **3.3.10 Hysteresis assay**

To examine whether the bimodal response of the tTA positive feedback, which is caused by the presence of a tetO array, is bistable, we tested for hysteresis behavior. Yeast haploid cells harboring 1) tTA transcriptional positive feedback, 2) 1xtetO promoter driving YFP reporter and 3) various numbers of tetO array placed in either 2u high copy plasmid or chromosome were grown in synthetic medium with 2% glucose at 100 and 1000 ng/mL dox levels overnight. Then, cells were diluted to various dox levels and further grown for at least 16hr maintaining an exponential phase. After incubation, cells were placed on ice and fluorescence intensities were measured by flow cytometry.

### **3.3.11 Quantitative chromatin immunoprecipitation (qChIP)**

ChIP was performed as in (Aparicio et al, 2004) with slight modifications. Briefly, yeast strains grown overnight were diluted to OD<sub>600</sub> of ~0.001 in 200 mL synthetic medium with

2% glucose and then grown to mid-exponential phase (a final OD<sub>600</sub> between 0.7 and 1.0). Crosslinking was performed by resuspending cell pellets in 5.6mL of 37% formaldehyde and incubating for 20 min at room temperature, followed by addition of 10 mL of 2.5M glycine to quench the reaction. Fixed cells were vortexed with glass beads for 1hr at 4°C for lysis. Chromatin was sheared using a Microson Ultrasonic Cell Disruptor, with 6 x 10 second cycles at a power setting of 8. Chromatin was immunoprecipitated with Dynabead (Invitrogen) – Anti-HA High Affinity rat monoclonal antibody (Roche) complex as previously described (Lee et al, 2006). qPCR was performed on a Applied Biosystems 7300 real-time PCR machine. PCR efficiency of primers targeting the tetO promoter and array were confirmed to be > 1.85 using serial dilutions of either sheared chromosomal DNA or a highly concentrated IP DNA containing the tetO promoter and array. This also determined the threshold cycle (C<sub>T</sub>) range for linear amplification, and all C<sub>T</sub> values for INPUT and IP DNA were within this range.

### **3.4 Results and discussion**

#### **3.4.1 A tetO array converts the dose-response of the tetO promoter from graded to sigmoidal-like, suggesting the array affinity for tTA is higher than the promoter affinity**

To examine the effect of repeated decoy sites on target gene expression in this synthetic system, we first compared dose-response curves between dox and reporter expression in the absence and presence of the tetO array. We placed the array either on centromeric plasmid,

2 micron high copy plasmid or in *HIS3* locus of the genome. Adding decoy sites decreases expression at any given level of dox (Figure 3B&D, 4D). Since dox reduces active tTA level constitutively produced from the *MYO2* promoter, varying dox is equivalent to changing the tTA synthesis rate. Therefore, the decreased expression implies decoy-bound tTA is not protected from degradation and  $\delta_N/\delta$  cannot be much less than unity. We further verified the decoy array reduces expression at a given tTA synthesis rate by placing tTA expression under the control of the methionine-inducible *MET3* promoter (Figure 5). The simplest interpretation of these results is that decoy-bound and unbound tTA have the same degradation rate and we set  $\delta_N = \delta$ . We cannot exclude the possibility that  $\delta_N \gg \delta$ , but this does not change inferences about promoter and decoy binding strength (Figure 2).

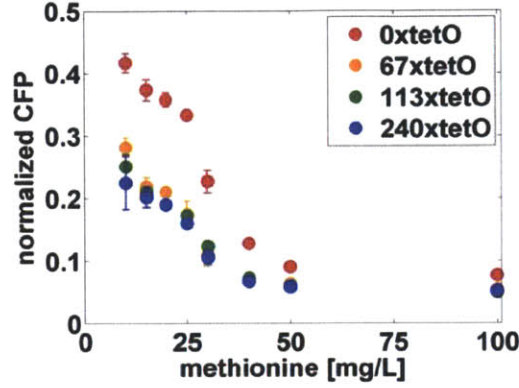
In order to examine a direct input-output relationship between tTA and gene expression affected by the decoy array, we need to find an active nuclear tTA level from external dox concentration. To do so, we apply a kinetic model derived in section 3.1.1 to estimate the total nuclear tTA level, the sum of free tTA level and array-bound tTA level. Surprisingly, the resulting dose-response of total tTA and reporter expression in the absence and presence of the decoy array shows that the array alters its shape from graded to sigmoidal-like response (Figure 3C&E, 4E). Since the dose-response of the reporter expression is graded when the promoter and array binding affinities are identical, this sigmoidal-like response suggests that array binding affinity for tTA is higher than the promoter affinity. Another finding that this altered response occurs for both plasmid-borne and



chromosomally integrated tetO array suggests that the array effect is not dependent upon its position.

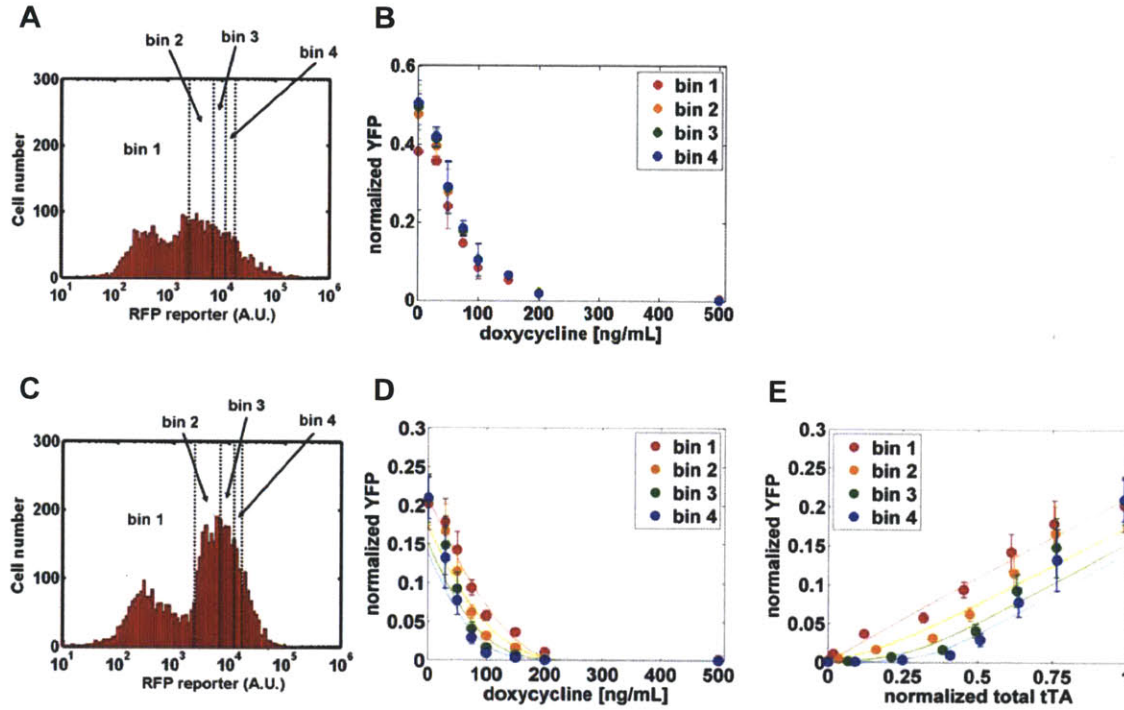
While the model captures the experimental data well at low tTA levels, it systematically overestimates the extent to which decoy sites decrease expression at high tTA levels where there is a decrease in the gap between target gene expression of strains with and without decoys. This feature cannot be explained for any choice of physical parameters by our model. The decreased gap implies either the decoys release bound tTA at higher tTA levels, increased array occupancy promotes gene expression at the promoter by an unknown mechanism, or total tTA levels change in the presence of decoy sites at low dox levels. The decreased gap is not dependent on our dox model because it remains even if the data are plotted as a function of dox rather than the total TF level. Therefore, we used only six data points at low tTA level for model-fitting. However, the fitting with all the data points does not change the conclusion (data not shown). Interestingly, we did not observe a significant decrease in the gap when we titrated tTA using the methionine-inducible promoter (Figure 5). Furthermore, when we used promoters stronger than *MYO2* to drive tTA expression, we did not observe significant increases in expression (data not shown).

more pronounced as plasmid copy number increases (corresponding to increases in the RFP bin number). Solid dots represent experimental data and error bars the s.d. of three replicates.

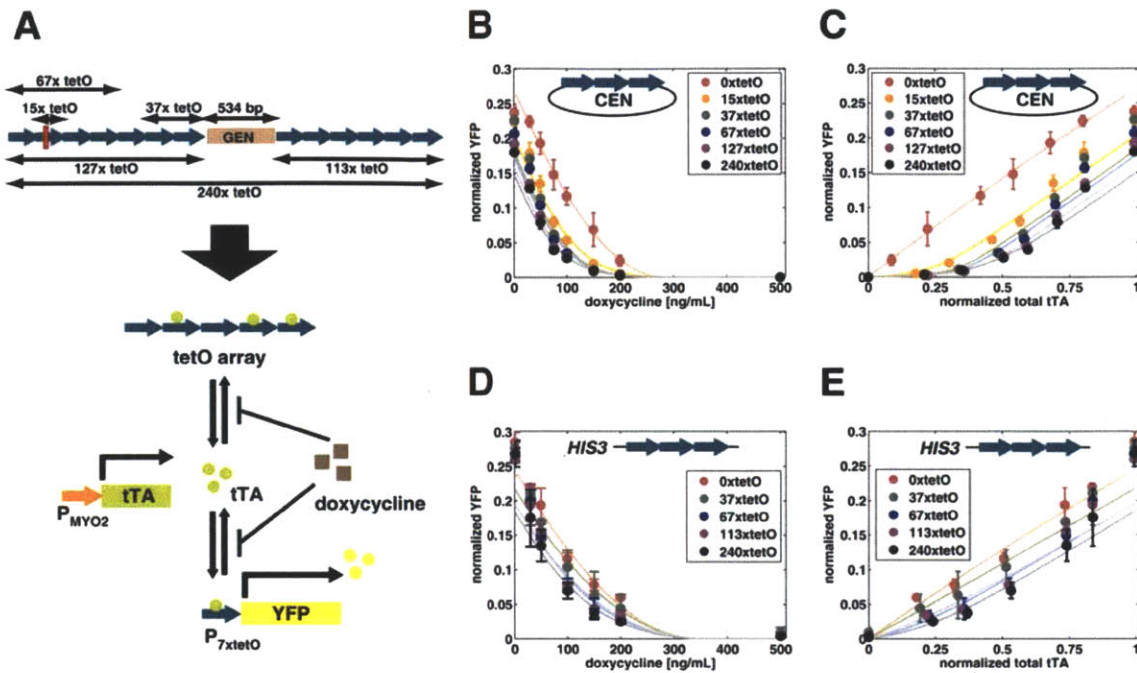


**Figure 5. The tetO decoy array reduces target gene expression when the tTA synthesis rate is directly varied using a methionine-inducible promoter.** A methionine-inducible *MET3* promoter driving expression of a NLS (nuclear localization signal) tagged tTA-YFP fusion was integrated at the *LEU2* locus in a strain background where a 7xtetO promoter drives CFP from the *URA3* locus. Increasing methionine levels result in reduction in CFP expression, suggesting that decoy-bound tTA is not protected from degradation and decoys have similar sequestration effects to doxycycline. Surprisingly, the effect of 67x tetO array is similar to those by a 113x and 240x tetO array, as in the case with titrating tTA using dox (Figure 3).

Two features of the fitting procedure deserve mention. First, we find estimates of  $K_M/T_0$  across different datasets are similar, as expected for a property that is independent of decoy number. Second, because both changing  $N_0/T_0$  and  $K_N/T_0$  can reduce target gene expression, we analyzed the covariance between these two parameters. This is best seen in Figure 6, where the sum of the squared residuals (SSR) of the fit is given for various values of these two parameters. For any given  $K_N/T_0$ , there is a narrow range of  $N_0/T_0$  that results in a good fit. In contrast, we found a range of  $K_N/T_0$  spanning several orders of magnitude results in a



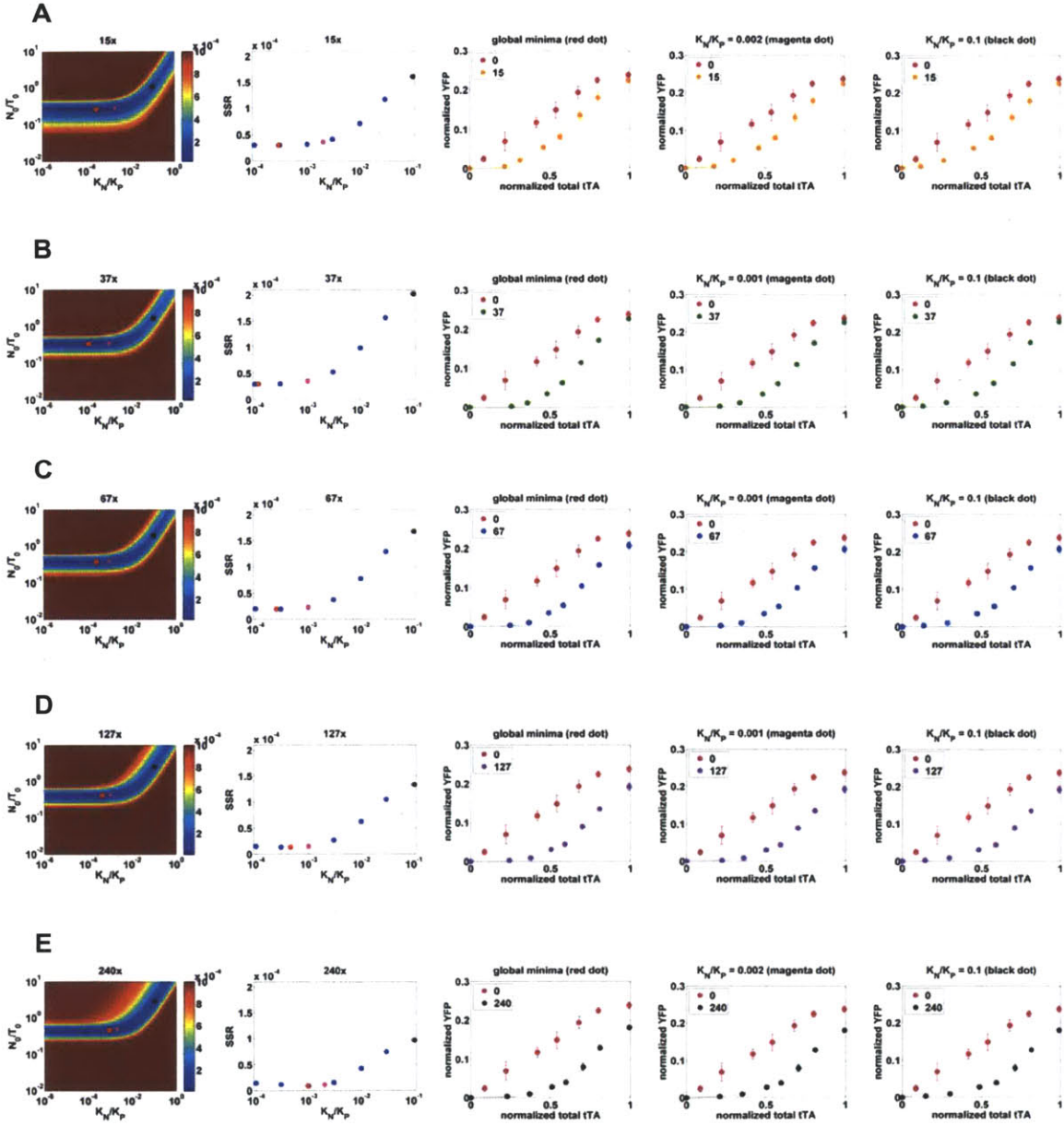
**Figure 4. High copy plasmid-borne tetO arrays also create a sharp sigmoidal-like dose response of the tetO promoter.** A high copy 2 micron plasmid containing a constitutively driven RFP (tdTomato) from the *ACT1* promoter and no (A&B) or a 67x tetO array (C&D), was introduced into cells containing constitutively expressed tTA from the *MYO2* promoter and a 7xtetO promoter driving YFP expression. The tTA activity was titrated by addition of doxycycline and both RFP and YFP expression was measured in single cells by flow cytometry. The combined RFP distribution of all the cells containing no array (A) or the 67xtetO array (C) was broad. Cellular autofluorescence has been subtracted. Cells were divided into four bins based on RFP expression, which should correlate with plasmid copy number. The mean expression and width of the second bin corresponds to RFP expression of a control strain which only contains a chromosomally integrated *ACT1* promoter driving RFP. The first peak and bin represents cells which have lost the plasmid. In panels B&D, the results of titrations disaggregated by bin are reported. Copy number of plasmids (B) without tetO sites has no significant effect on YFP expression, but (D) plasmids with the 67xtetO array, further inhibit expression. (E) The model-inferred, dose-response between total tTA and YFP expression corresponding to (D) also exhibits a sharp, sigmoidal-like response that is



**Figure 3. Arrays of tetO decoy sites reduce target gene expression and convert the graded dose-response between tTA and its target promoter to a sigmoidal-like response.** (A) tTA is expressed constitutively from a chromosomally integrated *MYO2* promoter at the *ADE2* locus, and its activity is titrated by addition of dox. Activation of a tTA-responsive 7xtetO promoter driving YFP integrated at the *URA3* locus was monitored by flow cytometry. Arrays of tetO binding sites of various sizes were created from a single 240x tetO array. (B) Target gene expression at the 7xtetO promoter is reduced at any dox concentration when tetO arrays are introduced on a centromeric plasmid. Dots represent experimental data and solid lines correspond to fits of the kinetic model to six data points at low tTA levels. (C) Using a model that accounts for dox-tTA interactions, the expression data in (B) can be plotted versus total tTA number (unbound + decoy-bound tTA). The arrays result in a qualitatively sharper change in the dose-response indicative of stronger binding of tTA to the array versus the promoter. (D, E) tetO arrays have a qualitatively similar, albeit weaker effect when integrated at the *HIS3* locus in the chromosome. Error bars represent s.d. of 2 biological replicates.

good fit. The SSR in this range is plotted in Figure 6 for various values of  $K_N/K_P$ . The minimum SSR value corresponds to a low  $K_N/K_P \sim 10^{-6}$ , but lies within a shallow plateau region. Because such a large change in affinity results in a physically nonsensical residence time for the TF, assuming a diffusion-controlled on-rate, we used the upper-bound of the plateau region as our estimate  $K_N/K_P$ . We defined this heuristically as when the SSR changes by less than 25%, leading to a reported  $K_N/K_P$  values that generally lie between  $10^{-2}$  and  $10^{-3}$  for all experiments (Table 2). Alternate ways of defining the end of the plateau region (based 1<sup>st</sup> and 2<sup>nd</sup> derivatives of the SSR versus  $K_N/K_P$  curve) gave similar answers. These values should be considered as order of magnitude estimates and clearly suggest a large difference in the strength of tTA binding to decoy sites versus productive binding events at the promoter exists. In Figure 6, we compare these fits to a case where  $K_N/K_P = 10^{-1}$ , which does not describe the data well.





**Figure 6. Sensitivity of data fitting to model parameters.** (A) (left) Contour plots of the sum of squared residuals (SSR) as function of parameters  $K_N/K_P$  and  $N_\theta/T_0$  from non-linear least squares fit of data from Figure 3B and C when a 15x tetO array was introduced (excluding two data points at high tTA levels). For  $K_N/K_P < 10^{-2}$ , there is a narrow range of  $N_\theta/T_0$  that yields a satisfactory fit but a much larger range of  $K_N/K_P$ . The red dot represents the global minima of SSR, but sits in a plateau

region that encompasses a large range of  $K_N/K_P$ . (*left-center*) Calculated SSR for various values of  $K_N/K_P$ , including the optima value (red dot),  $10^{-4}$ ,  $3 \times 10^{-3}$ ,  $10^{-3}$ ,  $3 \times 10^{-2}$ ,  $10^{-2}$ , and  $10^{-1}$  (black dot), using the best estimate for  $N_0/T_0$ . The magenta dot represents the  $K_N/K_P$  value probed where the SSR is 25% greater than the SSR for global minima. The criteria for the magenta dot were chosen so as to result in a point where the SSR begins to rapidly increase. These points are also indicated on the contour plot (*left*). The actual data and model fits for the (*center*) global minima (red dot), (*right-center*) our heuristic upper limit of  $K_N/K_P$  (magenta dot) and (*right*) when  $K_N/K_P = 0.1$  (black dot). Solid dots represent the experimental data and the solid line represents the model fit. Error bars are s.d. of biological duplicate samples. (B-E) are identical to (A) but corresponding to experimental data from Figure 3B for 37x, 67x, 127x, and 240x tetO arrays, respectively.

**Table 2. Fit parameter estimates.**

tetO array location	$\frac{K_M}{T_0} [\text{ng/mL}]^{-1}$	Number of tetO sites within array					
centromeric plasmid	0.0071	<sup>a</sup>	15	37	67	127	240
		$\frac{K_N}{K_P}$	0.002	0.001	0.001	0.001	0.002
		$\frac{N_0}{T_0}$	0.28 (1.1) <sup>b</sup>	0.35 (1.7)	0.38 (2.0)	0.43 (2.6)	0.49 (3.0)
		fold <sup>c</sup>	4.5	5.8	7.4	9	8.6
integrated ( <i>HIS3</i> )	0.0058		15	37	67	113	240
		$\frac{K_N}{K_P}$	-	0.08	0.01	0.02	0.007
		$\frac{N_0}{T_0}$	-	0.25 (0.29)	0.25 (0.71)	0.28 (0.65)	0.29 (0.97)
		fold	-	1	1.9	1.4	3.2
multiple centromeric plasmids	0.0066		67	67/67	67/67/67		
		$\frac{K_N}{K_P}$	0.004	0.0004	0.0002		
		$\frac{N_0}{T_0}$	0.23 (0.76)	0.47 (3.5)	0.57 (7.5)		
		fold	1.9	6.6	5.2		
2 $\mu$ Plasmid	0.0087		Bin 2	Bin 3	Bin 4		
		$\frac{K_N}{K_P}$	0.008	0.002	0.0008		
		$\frac{N_0}{T_0}$	0.21 (0.73)	0.30 (1.6)	0.36 (2.5)		
		fold	1.6	2.6	3.5		



<sup>a</sup> the reported  $K_N/K_P$  was chosen such that its sum squared of residuals (SSR) is only 1.25 times higher than the  $K_N/K_P$  found at the global minima, which is a good heuristic for the plateau region.

<sup>b</sup> The parenthetical value for  $N_0/T_0$  corresponds to the best estimate when  $K_N/K_P = 0.1$ .

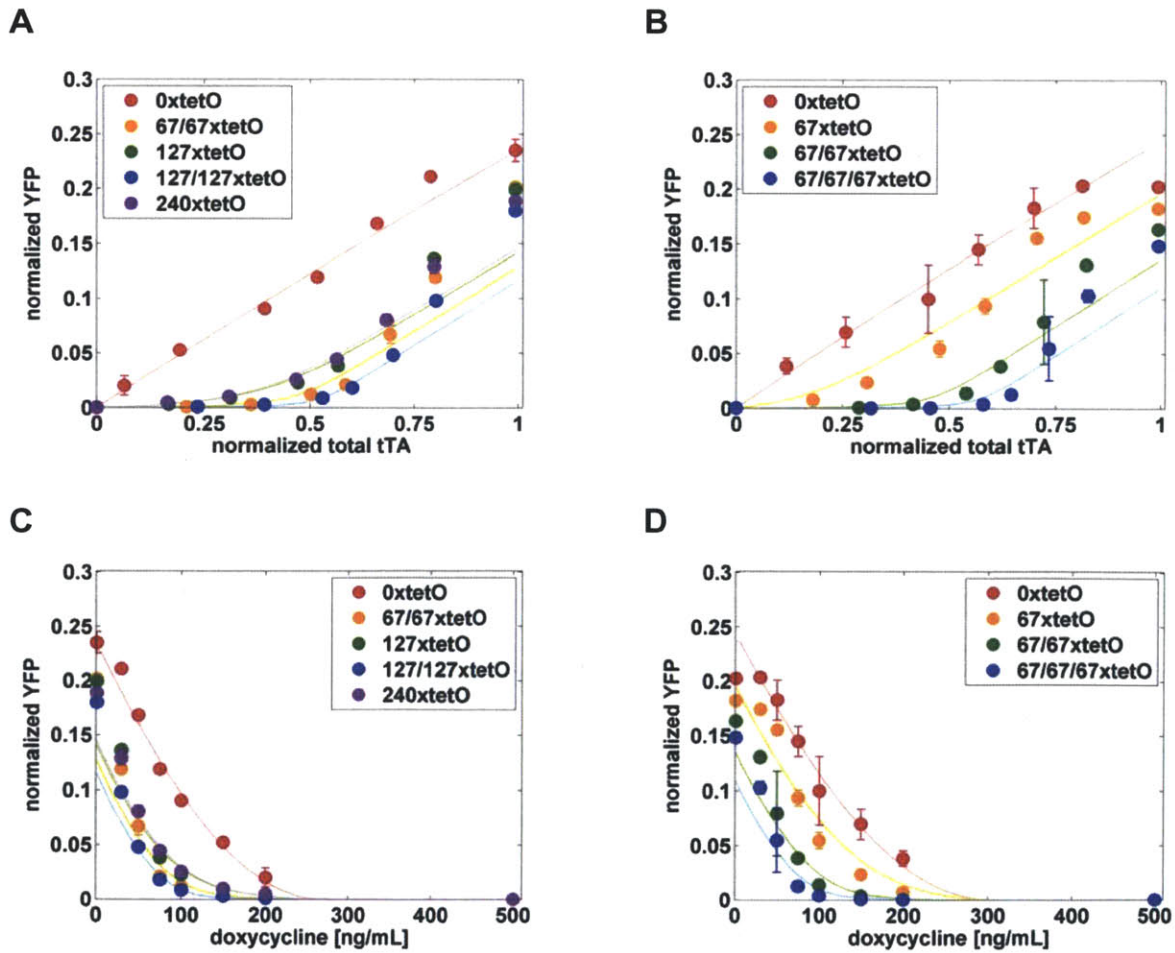
<sup>c</sup> “fold” represents the fold change of the SSR for  $K_N/K_P = 0.1$  versus the SSR for the reported  $K_N/K_P$ .

The sigmoidal-like shape of the dose-response curve suggest  $K_N/K_P \ll 1$ . However, describing the tTA-tetO interaction using a single thermodynamic affinity is likely a gross simplification. In reality, the residence time of tTA to either the promoter or decoy sites depends on a combination of (1) interactions with multiple tetO binding sites, (2) the conformation and chromatin state of the DNA, (3) interactions of tTA with other proteins that may also have affinity for DNA (general transcriptional machinery, chromatin remodeling factors, etc.), and (4) other unknown factors. Some of these interactions are not at thermodynamic equilibrium. Nevertheless, the model serves as a useful framework for pinpointing what interactions must be different.

### 3.4.2 Split tetO arrays are more potent than a long contiguous tetO array

One of the unexpected results is that at any given dox concentration, increasing array size beyond 67x tetO sites had little additional effect on decreasing target gene expression. In fact, the *actual* number of tetO sites in the array is not input into the model. The parameter  $N_0/T_0$  represents the effective number of tetO site and it is remarkably similar for 67x, 113x, 127x, and 240x array sizes (Table 2). However, when decoy site numbers are increased by using high copy plasmids  $N_0/T_0$  does increase (Figure 4, Table 2). To determine if the contiguous nature of the additional decoy sites was responsible for the decrease in their

effectiveness we constructed yeast strains that had multiple centromeric plasmid-borne decoy arrays. We found 2 copies of a plasmid-borne 67x array was more effective versus both 1 copy of a plasmid-borne 127x array and 1 copy of a plasmid-borne 240x array (Figure 7A&C). Further increases in the copy number of plasmid-borne arrays continue to lower gene expression (Figure 7B&D) confirming a split, non-contiguous array of tetO sites is more effective in sequestering tTA than a contiguous array. These results suggest that only a subset of the contiguous tetO array sequesters tTA and are responsible for the reduction of gene expression.



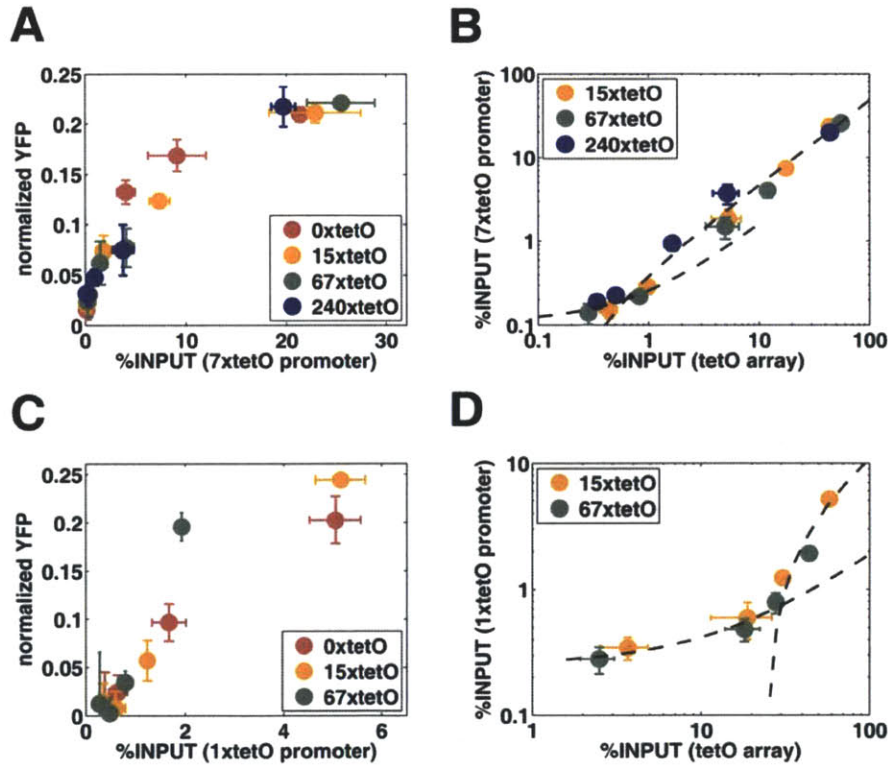
**Figure 7. Noncontiguous tetO arrays sequester tTA more effectively than contiguous tetO**

**arrays.** (A) The dose-response of cells with a 7xtetO promoter driving YFP in the presence of two separate centromeric plasmid-borne 67xtetO arrays (67/67) versus one 127xtetO array (127) or two centromeric plasmid-borne 127xtetO arrays (127/127) versus one 240xtetO array (240). In strains with a single contiguous array, a second empty plasmid is also present. Separating the array into two different locations increases its effectiveness in sequestering tTA. (B) A similar effect is seen when comparing expression in strains containing 1, 2, or 3 copies of 67x centromeric plasmid-borne arrays. Error bars represent s.d. of 3 biological replicates. (C, D) The dose-response of reporter expression versus doxycycline that corresponds to the data for split arrays shown in (A) and (B), respectively.

**3.4.3 *in vivo* binding assay also shows that the array affinity for tTA is higher than the promoter's but it decreases with increasing tTA levels**

If only a fraction of binding sites within an array allows an effective tTA sequestration for lowering a free tTA level, the remaining part of the array should either weakly interact with tTA or completely protect the array-bound tTA. A preliminary result using fluorescence recovery after photobleaching (FRAP) experiment using YFP tagged nuclear localized tTA showed that the fluorescence in the photobleached region of the array is recovered in a few minutes, suggesting the array does not protect tTA (data not shown). Therefore, there should be two binding regimes within the tetO array: effective tTA binding for the reduction of gene expression and non-effective binding. The affinity of the effective tTA binding regime should be higher than the tTA-promoter binding affinity to lead a sigmoidal-like dose-response of the tetO promoter. To directly measure the binding affinity of the promoter and the array, we introduced quantitative ChIP experiments to measure

occupancy of a 3X-HA-tagged tTA at both the promoter and the array. We first measured tTA occupancy at the promoter for cells with 0x, 15x, 67x, and 240x tetO sites at various dox levels. The “% INPUT” ChIP signal used here reflects the percentage of input chromatin that is immunoprecipitated by an HA-specific antibody. In Figure 8A we observe that promoter occupancy is roughly proportional to expression at the 7xtetO promoter.

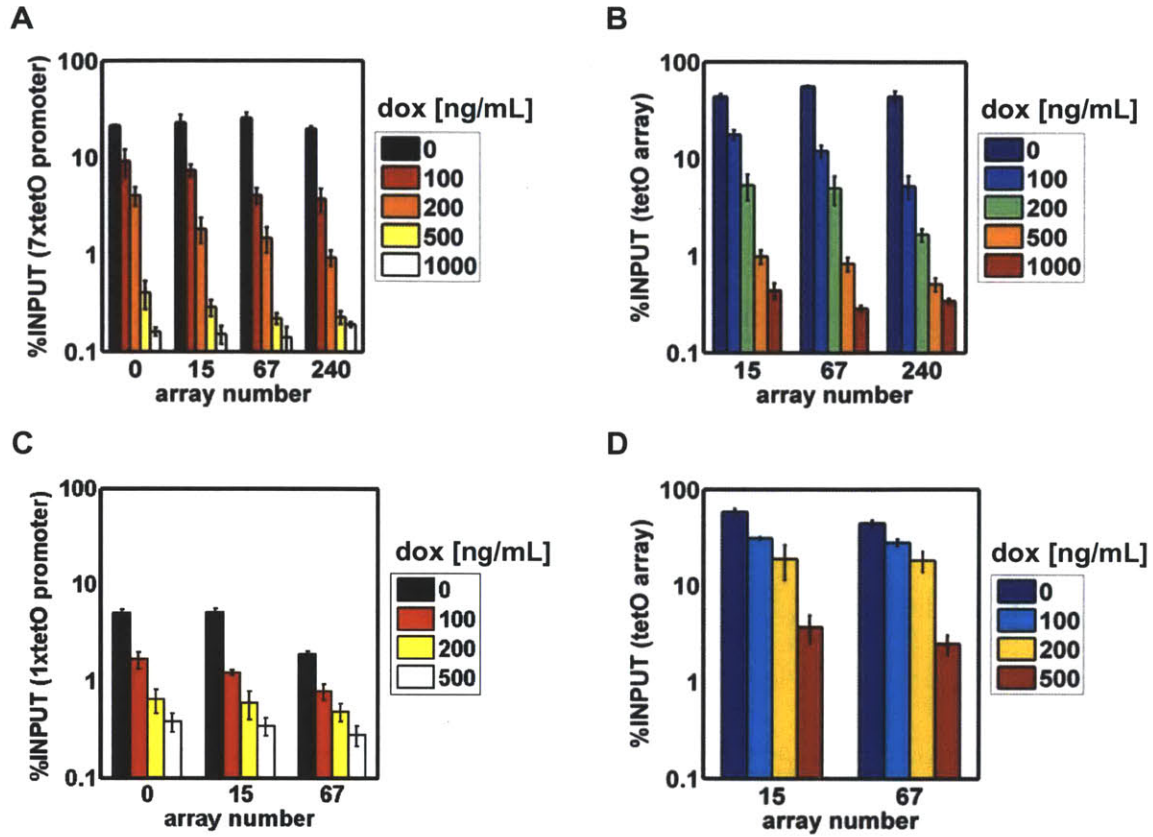


**Figure 8. tTA binds the tetO array and the tetO promoter with different strengths. (A)**

Occupancy of a C-terminal 3X-HA-tagged tTA is monitored at the 7xtetO promoter by ChIP in the presence and absence of centromeric-borne tetO arrays at 0, 100, 200, 500 and 1000 ng/mL dox. A region including only the 7xtetO binding sites is probed. ChIP signal (% of INPUT DNA measured in the immunoprecipitated (IP) sample) of tTA varies nearly linearly with target gene expression.

(B) tTA occupancy was also monitored at a particular region in the tetO arrays encompassing 6xtetO sites (red region denoted in Figure 3A). Promoter occupancy is plotted versus tetO array occupancy on a log-log scale. The relative change in promoter versus tetO occupancy is slightly different at low tTA levels versus higher tTA levels. Dotted lines are straight lines (that appear curved on a log-log plot) that could represent two binding regimes. (C) and (D) are identical to (A) and (B), respectively, but with a 1xtetO promoter. Expression varies nearly linearly with tTA promoter occupancy, but there is a dramatic difference in the relative binding of tTA to the promoter versus the array as the array occupancy increases. Error bars represent s.d. of triplicate chromatin samples prior to IP.

Using the same chromatin samples, we next measured tTA occupancy at a particular region shared among the different tetO arrays (red region in Figure 3A). Array occupancy clearly increases with increasing tTA level and not just at low tTA levels (Figure 8B and Figure 9B). At any given tTA level, the ChIP signal is weaker for the 240x array compared to the 67x array. This is expected since the 67x and 240x arrays are equally potent in reducing target gene expression. Therefore, these data provide direct support that the “per site” occupancy of tTA on the 67x array is several-fold higher than the 240x array. The extremely high ChIP signals we observe at the array when high levels of tTA are expressed may be because any sheared fragment encompassing the probed region also has additional tTA molecules present on tetO sites adjacent to the probed region. Therefore, not only is there a strong interaction between the 3X-HA tag and the anti-HA antibody, but also many antibodies will be bound to each probe, resulting in highly efficient precipitation of tTA.



**Figure 9. TF occupancy at the 7x and 1x tetO promoters and various tetO arrays.** Data presented here are identical to data in Figure 8, but in a different format. Occupancy as measured by % INPUT (IP/INPUT) is reported for the (A) 7xtetO promoter and (B) various tetO arrays in cells containing the 7xtetO promoter, at different dox concentrations. Similar experiments were done to measure the occupancy of the (C) 1xtetO promoter and (D) of the tetO arrays in cells containing the 1xtetO promoter. For the array, a 6xtetO region within the array was amplified (red region denoted in Figure 3A). Error bars represents the standard deviation of triplicate INPUT and IP samples.

To determine if tTA preferentially binds the tetO array as compared to the promoter at low tTA levels, we plotted the ChIP signal for the array versus the promoter (Figure 8B). At

lower tTA levels the slope of this plot is smaller, showing array occupancy changes more significantly than promoter occupancy. This supports the previous model prediction that the binding sites within the array have higher affinity for tTA versus promoter binding sites. As the tTA level increases, tTA occupancy at both the promoter and array increases, indicating that the array does not saturate before promoter activation. Surprisingly at higher tTA levels the slope is higher than at low tTA levels. If the ChIP data faithfully represent occupancy, then the slope corresponds to the ratio of tTA binding affinity to the promoter and the array. The change in slope indicates either the tTA-promoter binding becomes stronger or tTA-array binding becomes weaker as tTA levels increase. If the tTA-promoter affinity positively correlates with tTA level, the slope of the dose-response between tTA and reporter should also increase with tTA levels in a non-linear manner in the absence of the array. Since we observed a linear promoter dose-response (Figure 3C&E), we rule out this possibility; the data strongly suggest that the array affinity decreases with increasing tTA levels.

The difference in slopes that marks a transition from tight tTA-tetO interactions on the array to weaker interactions is more subtle than anticipated; especially given the sharp change in concavity we observe in the expression data. We hypothesized that perhaps not all the binding sites in the 7xtetO promoter are weaker binding as compared to the array. ChIP measures an “average occupancy” across multiple binding sites (as the chromatin was sheared to an average size of 300 bp that encompasses multiple sites on average). To address this, we repeated the experiment using a 1xtetO promoter. We see a similar linear

relationship between promoter occupancy and expression (Figure 8C). Strikingly, when comparing the ChIP signal for the decoy arrays versus the 1xtetO promoter, the transition between the binding regimes is much more dramatic (Figure 8D).

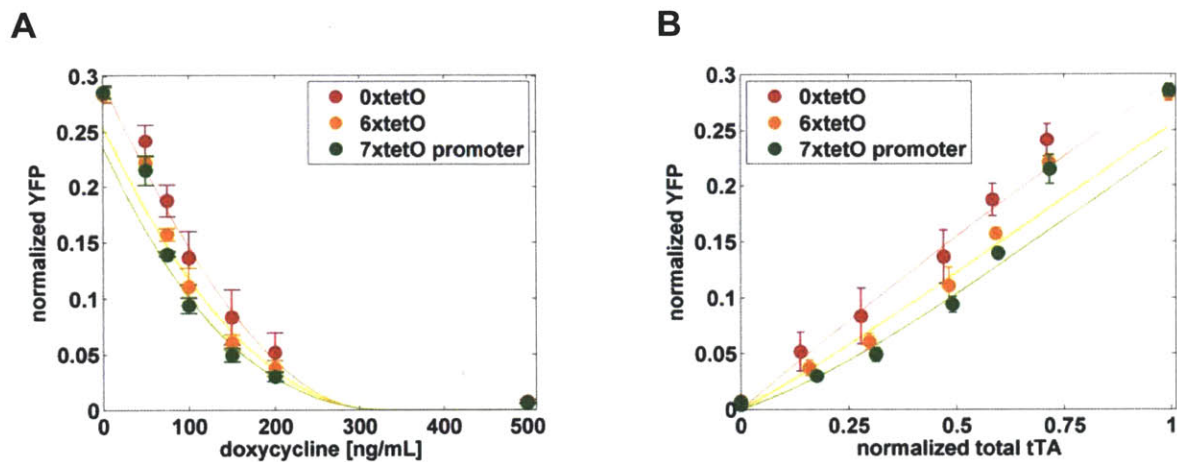
Together with our model, these data support the notion that the residence time of tTA on both the 1xtetO promoter and a subset of binding sites in the 7xtetO promoter is shorter than on decoy tetO sites, but only when the tetO array region is relatively free of bound tTA. At higher tTA levels, tTA continues to bind the remaining vacant tetO sites in the array, albeit weakly. This weak binding does not appear to alter gene expression levels. The origin of the array affinity change remains unclear. It is possible that as more tTA occupy the array structural changes induced by the binding, such as bending or twisting the DNA strand, could alter subsequent binding (Wang et al, 2005).

If some tetO sites within the 7xtetO promoter do bind tTA as tightly as those in the array, then an additional copy of the promoter should affect gene expression in a manner similar to the array. To test this, we constructed a centromeric plasmid containing the 7xtetO promoter driving CFP and introduced it into the usual yeast strain expressing tTA and containing an integrated 7xtetO promoter driving YFP. Addition of this plasmid reduces target YFP expression and changes the concavity of the dose-response (Figure 10).

Introducing a centromeric plasmid containing 6xtetO sites from the 7xtetO promoter, but without the minimal *CYCI* promoter or the CFP open reading frame, has similar effects. Therefore, a subset of the tetO sites in the 7xtetO promoter binds tTA as strongly, like the



tetO sites present in the array region. This would explain why 7xtetO promoter binding is relatively strong even at low tTA levels (Figure 8B). It also implies that the location of tetO sites or the regional chromatin environment does not contribute significantly to stronger binding of tTA versus the productive promoter binding.

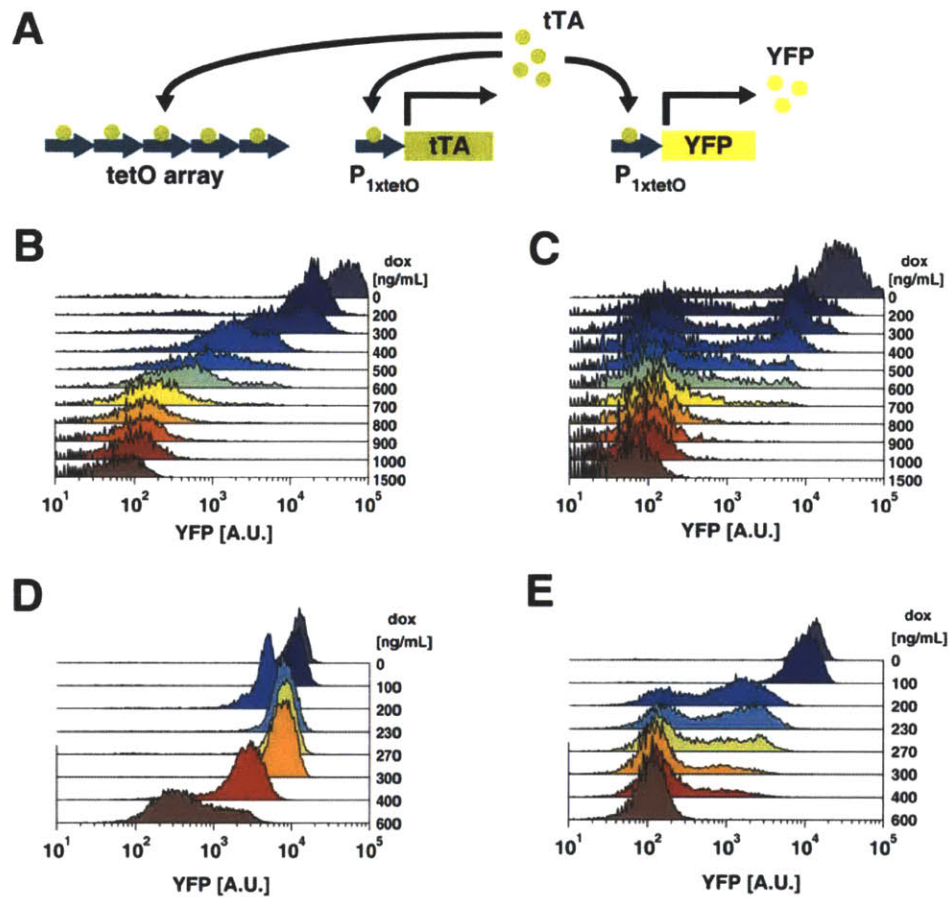


**Figure 10. 6xtetO array and 7xtetO promoter create the sigmoidal dose response. A**

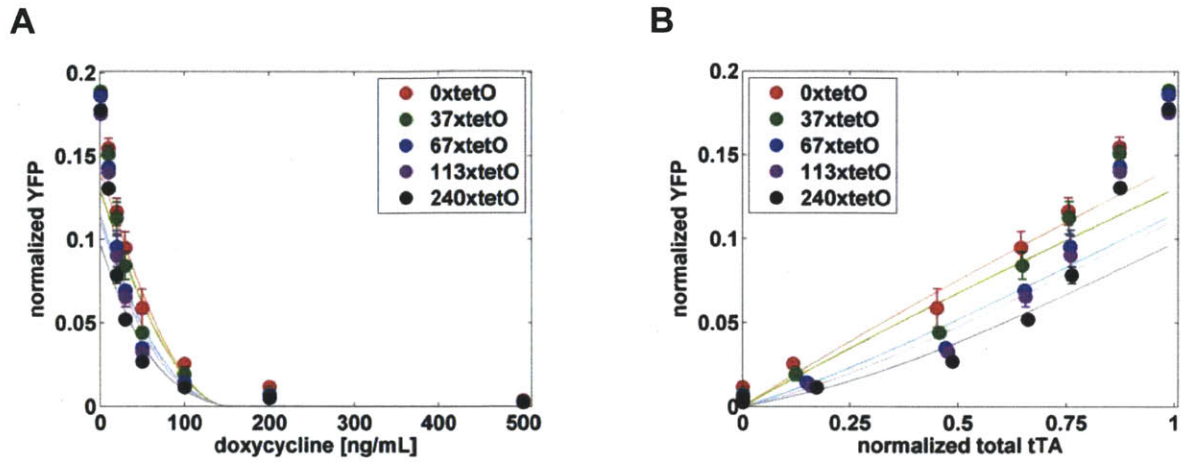
centromeric plasmid containing 6xtetO or 7xtetO promoter was introduced into the cells in which tTA was constitutively expressed from a *MYO2* promoter integrated at the *ADE2* locus and a 7xtetO promoter driving YFP expression integrated at the *URA3* locus. The tTA activity was titrated by addition of doxycycline and YFP expression was measured in single cells by flow cytometry. (A) The dose-response between dox and normalized YFP expression. (B) The dose-response between normalized total tTA and normalized YFP expression. Normalization was done as in Figure 3. Solid dots represent experimental data and error bars the s.d. of three replicates.

#### **3.4.4 A tetO array converts a graded transcriptional positive feedback response to a bimodal response**

The dox titration data demonstrate that repeated decoy arrays are effective at decreasing target gene expression and do so in a manner that converts the linear dose-response to one with an inflection point. As an additional, more stringent test of this conversion occurring, we added tetO arrays to a strain containing a 1xtetO promoter driving tTA expression in a transcriptional positive feedback (Figure 11A). The tTA levels were indirectly assayed by expression from a 1xtetO promoter driving YFP expression. When we titrate the feedback strength of a 1xtetO promoter in positive feedback using dox, we observe a graded response (Figure 11B&D), as has been previously shown as the 1xtetO promoter response in the absence of feedback is gradual and nearly linear (To & Maheshri, 2010) (Figure 12). If the decoy sites generate an inflection in the dose-response, then their addition could lead to bistable gene expression when in positive feedback. Indeed, we see bimodal expression when we introduce a 2 $\mu$  plasmid containing a 127x tetO array, integrate two copies of a 67x tetO array in the genome (Figure 11C&E), or add a centromeric plasmid containing a 240x tetO array (data not shown). This result also suggests that repeated decoy sites could create a qualitative change of a biological network behavior. Since all-or-none responses play a central role in cellular differentiation and apoptosis, the repeated decoy sites may be served as an alternative mechanism to regulate these cellular behaviors.



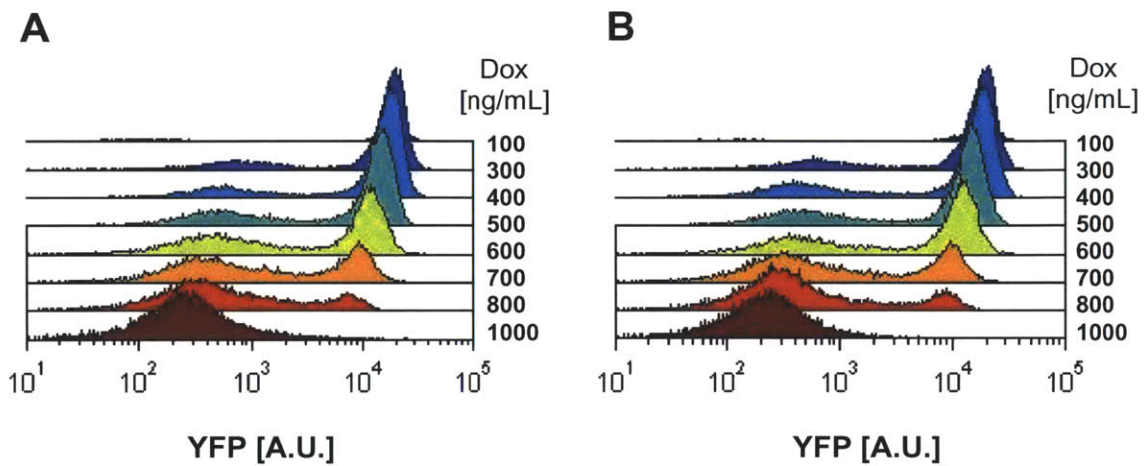
**Figure 11. Adding *tetO* decoy sites converts a positive feedback loop from a graded to a switch-like bimodal response.** (A) Two  $1\tetO$  promoters driving *tTA* and *YFP* expression are chromosomally integrated at the *HIS3* and *LEU2* locus, respectively. (B) Analytical flow cytometry of single cell fluorescence distributions at various dox concentrations (ng/mL). In the absence of the *tetO* array, the behavior response of the positive feedback where the promoter is embedded is graded. The feedback strength was changed by controlling dox concentration. (C) In the presence of a  $2\mu$  plasmid-borne  $127\tetO$  array, a graded response was changed to a bimodal response. Chromosomal integration of two  $67\tetO$  arrays at the *TRP1* and *URA3* locus is less potent but also converted the (D) graded positive feedback response into (E) the bimodal response over a smaller range of dox concentrations.



**Figure 12. Addition of a tetO array alters the dose-response of a 1xtetO promoter in a manner similar to the 7xtetO promoter.** Experiments were performed as in Figure 3, but with a 1xtetO YFP reporter replacing the 7xtetO YFP reporter at the *URA3* locus. The tetO array was chromosomally integrated at the *HIS3* locus. Solid dots represent experimental data and solid lines are a model fit. Error bars represent the s.d. of two replicates.

In a bistable response, input signals can shift the steady-state response between two fixed points. The signal strength required to make the switch depends on the initial state (hysteresis). To determine if the bimodal response observed with the positive feedback loop in the presence of tetO array (Figure 11) was bistable, we tested for hysteresis. Yeast cells containing the positive feedback loop were grown at two different dox levels (100 and 1000 ng/mL) until they reached a steady-state (on or off, depending on dox level). They were then transferred to fresh medium containing various dox levels, and grown for at least 16hrs. If the hysteresis occurs in the presence of the tetO array, two different initial states of positive feedback (on and off states) will result in different windows of tTA level in which on-off switching occurs. However, in the presence of 127x tetO array, the switching

window for the feedback response from on state was identical to that from off state (Figure 13). Though this result do not confirm a bistable positive feedback response, if a random fluctuation between on and off states occurs more slowly than the YFP reporter life timescale, the identical switching windows will be observed. (Octavio et al, 2009).



**Figure 13. The bimodal positive feedback response shifted from ‘off’ state is similar to that from ‘on’ state in the presence of tetO array.** (A) Cells containing 1xtetO promoter driving tTA at *HIS3* locus, 1xtetO promoter driving YFP reporter at *LEU2* locus and 2 $\mu$  plasmid harboring 127xtetO array and *ACT1* promoter driving tdTomato were grown at the low dox level (100 ng/mL), maintaining ‘on’ state of YFP expression. Then, cells were diluted to various dox levels and further grown until they reached a steady state. At intermediate dox levels,(300-800 ng/mL), the bimodal YFP expression is observed. (B) Cells initially grown at the high dox level (1000 ng/mL) were transferred into medium containing various dox levels as (A). At the steady state, they also showed a similar bimodal distribution to (A) at intermediate dox levles.



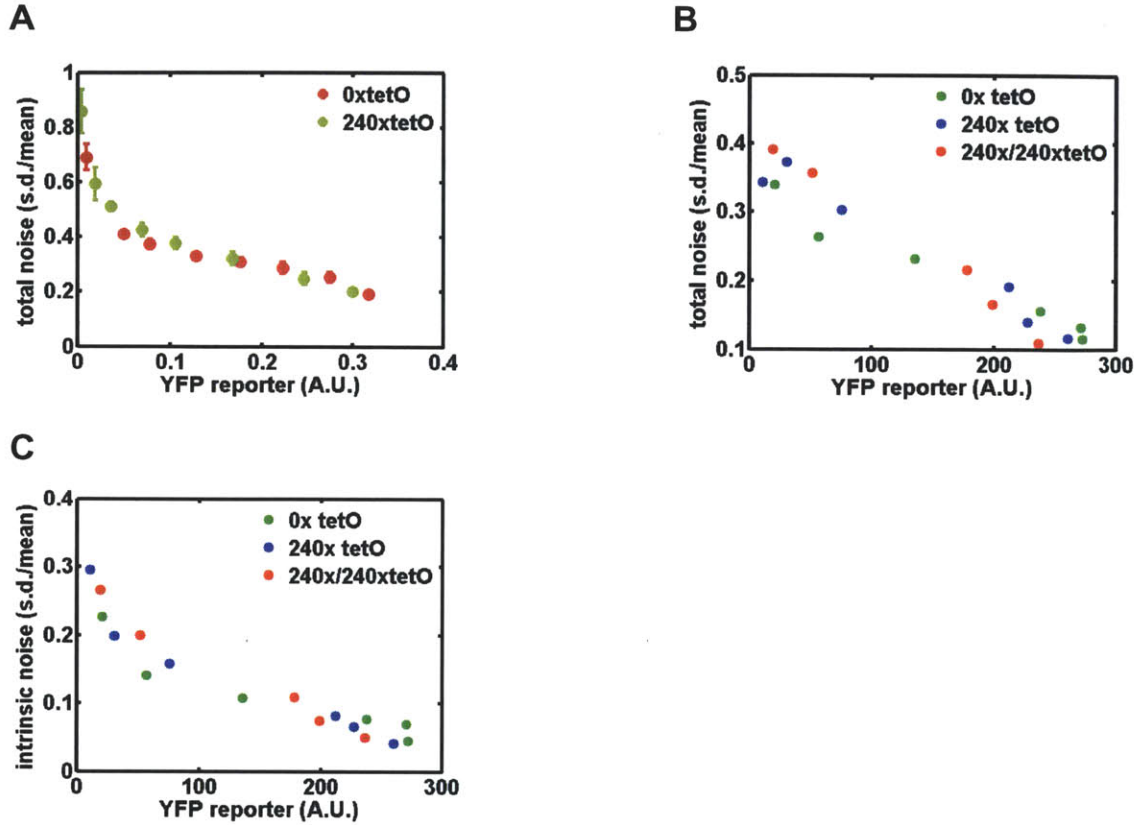
### 3.4.5 A tetO array does not alter gene expression noise

We tested whether tTA-tetO array binding affects the variation of target gene expression. If a time scale of tTA fluctuation affected by array sites is shorter than mRNA or protein life time scale, the tetO array will not contribute on gene expression noise. We found that addition of decoy sites has no effect on the variation in gene expression (Figure 14). This suggests not only that tTA binding and unbinding is faster than mRNA or protein lifetime but also that the array-driven bimodal positive feedback behavior is not due to the noise-induced bimodality (To & Maheshri, 2010).

If the residence time of tTA bound to the tetO array is on order of cell cycle time or larger, it is conceivable that the decoys will increase intrinsic noise in gene expression when tTA numbers are low. This is because some cells would have active (not bound to array) tTA while others would not. Alternatively, one might expect a reduction in the extrinsic noise with the addition of decoys, as decoy-bound tTA might buffer extrinsic fluctuations in free tTA (also suggested in (Burger et al, 2010)). However, in all our experiments the total noise measured at the same mean level of expression was identical, regardless of whether the tetO array was present or not (Figure 15A&B). To further confirm this result, we measured the intrinsic and extrinsic noise of gene expression from a diploid yeast strain containing an integrated 7xtetO promoter driving YFP and CFP from the same *URA3* locus, with and without repeats. Intrinsic noise was calculated as in (Elowitz et al, 2002). Again, at the same mean level of expression, both intrinsic and extrinsic fluctuations were identical (Figure 14C, extrinsic noise not shown).

The fact that decoy sites do not alter noise in expression might be explained by the dynamics of the various protein/DNA interactions. We have argued that an upper limit on the residence time of tTA bound to the tetO array is order 10 minutes. Given that the burst frequency of the 7xtetO promoter ranges from 0.1-1 burst per every 15 minutes, the promoter likely samples several tTA binding, unbinding, and degradation events per burst. This could explain why the experimentally measured intrinsic noise does not vary.

The extrinsic fluctuations may not be affected by decoy number if they are not due to pathway-specific components like tTA that are affected by the decoys. Even if they are affected by tTA, the tetO array still may not buffer extrinsic fluctuations in free tTA. For example, (1) tetO array copy number could be correlated to free tTA fluctuations; (2) extrinsic fluctuations in free tTA could occur on timescales equal to or faster than the burst frequency or (3) extrinsic fluctuations in tTA could be proportional to *total* tTA and not free tTA. To elaborate on (3), if two strains with and without decoy sites have the same mean expression level, then the strain with decoys have more total tTA. Then although total tTA fluctuations will be necessarily larger in magnitude, the decoy sites could buffer tTA fluctuations such that free tTA fluctuations are similar to the case where no decoy sites are present. It might be worth noting that the free tTA fluctuations we speak about here are really nuclear tTA fluctuations and so are governed by nucleocytoplasmic transport processes as well.



**Figure 14. The chromosomally integrated tetO array does not affect noise in gene expression.**

Noise of the reporter expression (standard deviation of the expression divided by mean of the expression) is plotted at various YFP mean levels. (A) Total noise in the expression of a 7xtetO promoter integrated at the *URA3* locus was measured at different levels of doxycycline in a haploid yeast strain, in the absence or presence of a 240xtetO array integrated at the *HIS3* locus. Error bars represent s.d. of three replicates. To verify that intrinsic noise in YFP expression is not affected by the tetO array, we repeated this experiment in diploid strains containing the 7xtetO promoter driving YFP and CFP at the *URA3* locus of each of homologous chromosome. Total (B) and intrinsic noise (C) measured was the identical, regardless of the presence or absence of one (240x) or two (240x/240x) integrated tetO arrays.



### 3.5 Conclusion

Intergenic repeated TF binding sites may play a role of competitive inhibitor for the promoter and affect target gene expression. We verified this hypothesis by using yeast synthetic system where an activator (tTA) induces gene expression. We found that the repeated activator binding sites convert a dose-response of a promoter from graded to sharp, sigmoidal-like one, suggesting 1) activator-decoy sites binding is much stronger than activator-promoter binding and 2) the effect of decoy repeats as competitive inhibitor does not further increase beyond a threshold size of array.

By using ChIP assay, we further confirmed a large difference of promoter/array affinities, measuring the effective activator-decoy sites binding regime which leads the reduction of target gene expression. ChIP results revealed that a strong activator-decoy sites binding becomes weaker with increasing activator levels, suggesting a possibility of the negative cooperative activator-decoy sites interaction. Moreover, ChIP assay showed only a subset of multiple activator binding sites within a promoter strongly sequesters the activator as the decoy binding sites do.

Finally, we showed that the repeated decoy sites can alter a qualitative behavior of a transcriptional positive feedback where a target promoter is embedded. This result suggests that hypervariable decoy DNA sites may lead a rapid qualitative change of gene network behaviors as their size alters in short time scale.

Additionally, we measured the number of sequestered activators in the decoy binding sites by microscopy to directly monitor an activator-decoy repeats binding (Appendix A). We found that a longer array sequesters more activator at high activator levels but not all array-bound tTAs are effective for reducing target gene expression.

To explore underlying mechanisms for the large affinity difference between promoter and repeated decoy sites we examined the effect of chromatin architecture for promoter and decoy sites by using a genetic approach considering dynamic interactions between activator and other transcription-involved proteins such as proteasome, chromatin remodelers and transcriptional cofactors (Appendix B). Our finding showed that a number of these interactions as well as chromatin structure are not responsible for affinity change though there are still remaining possibilities for unknown interactions.

### **3.6 References**

Aparicio, O., Geisberg, J.V., and Struhl, K. (2004). Chromatin immunoprecipitation for determining the association of proteins with specific genomic sequences in vivo. *Curr. Protoc. Cell. Biol.* Chapter 17, Unit 17.7.

Burger, A., Walczak, A.M., and Wolynes, P.G. (2010). Abduction and asylum in the lives of transcription factors. *Proc. Natl. Acad. Sci. U. S. A.* 107, 4016-4021.

Degenkolb, J., Takahashi, M., Ellestad, G.A., and Hillen, W. (1991). Structural requirements of tetracycline-Tet repressor interaction: determination of equilibrium binding constants for tetracycline analogs with the Tet repressor. *Antimicrob. Agents Chemother.* 35, 1591-1595.

- Elowitz, M.B., Levine, A.J., Siggia, E.D., and Swain, P.S. (2002). Stochastic gene expression in a single cell. *Science* 297, 1183-1186.
- Gardner, T.S., Cantor, C.R., and Collins, J.J. (2000). Construction of a genetic toggle switch in *Escherichia coli*. *Nature* 403, 339-342.
- Gari, E., Piedrafita, L., Aldea, M., and Herrero, E. (1997). A set of vectors with a tetracycline-regulatable promoter system for modulated gene expression in *Saccharomyces cerevisiae*. *Yeast* 13, 837-848.
- Janssen, S., Cuvier, O., Muller, M., and Laemmli, U.K. (2000). Specific gain- and loss-of-function phenotypes induced by satellite-specific DNA-binding drugs fed to *Drosophila melanogaster*. *Mol. Cell* 6, 1013-1024.
- Lau, I.F., Filipe, S.R., Soballe, B., Okstad, O.A., Barre, F.X., and Sherratt, D.J. (2003). Spatial and temporal organization of replicating *Escherichia coli* chromosomes. *Mol. Microbiol.* 49, 731-743.
- Lee, T.I., Johnstone, S.E., and Young, R.A. (2006). Chromatin immunoprecipitation and microarray-based analysis of protein location. *Nat. Protoc.* 1, 729-748.
- Murphy, K.F., Balazsi, G., and Collins, J.J. (2007). Combinatorial promoter design for engineering noisy gene expression. *Proc. Natl. Acad. Sci. U. S. A.* 104, 12726-12731.
- Octavio, L.M., Gedeon, K., and Maheshri, N. (2009). Epigenetic and conventional regulation is distributed among activators of FLO11 allowing tuning of population-level heterogeneity in its expression. *PLoS Genet.* 5, e1000673.

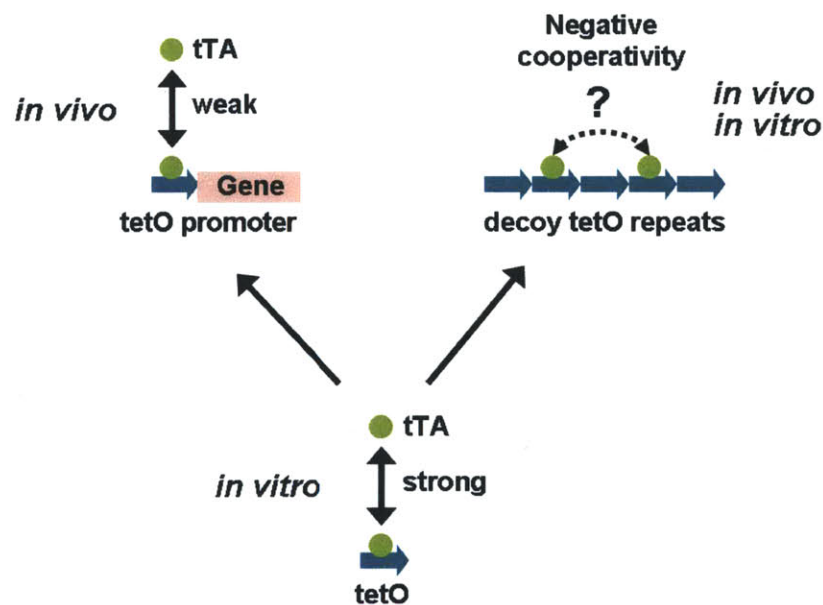
Thomas, B.J., and Rothstein, R. (1989). Elevated recombination rates in transcriptionally active DNA. *Cell* 56, 619-630.

To, T.L., and Maheshri, N. (2010). Noise can induce bimodality in positive transcriptional feedback loops without bistability. *Science* 327, 1142-1145.

Wang, Y.M., Tegenfeldt, J.O., Reisner, W., Riehn, R., Guan, X., Guo, L., Golding, I., Cox, E.C., Sturm, J., and Austin, R.H. (2005). Single-molecule studies of repressor–DNA interactions show long-range interactions. *Proc. Natl. Acad. Sci. U. S. A.* 102, 9796-9801.

#### 4. Overall conclusions and future directions

Our study demonstrates that repeated decoy TF binding sites in intergenic regions can regulate the TF's target gene expression by effectively sequestering the TF. When a TF is an activator, the TF-decoy interaction not only reduces gene expression but also alters the shape of the target promoter's dose-response from graded to sigmoidal-like. This is because TF-decoy binding, at low TF levels, is much stronger than TF-promoter binding. Since the sequence of decoy binding sites is identical to that of the promoter binding site, a large difference in binding affinities between decoy sites and promoter is unexpected. This raises the following two questions: 1) Why do repeated decoy sites bind TF more tightly than a target promoter and 2) why does the binding strength of decoy sites decrease with TF levels?



**Figure 15. tTA-tetO binding affinity may be altered under various environments.** *in vitro* tTA binding to the tetO site is very strong (~nM dissociation constant) (Kamionka et al, 2004). This binding affinity could be reduced during *in vivo* active transcription if the tetO site is embedded within a promoter and transcriptional regulatory proteins actively disassemble tTA from the promoter. The strong binding may also be weakened when tTA interacts with a tetO array if first strong tTA binding inhibits neighboring sites' binding (negative cooperativity). Whether this negative cooperative binding effect results from unknown interactions between tTA and other proteins *in vivo* or whether it comes purely from direct tTA-DNA binding, for example by DNA bending upon tTA binding, is generally unclear.

A possible scenario addressing the first question is that a dynamic interaction between TF bound to the promoter and other regulatory proteins reduces the stability of the TF-promoter complex during active transcription; the same interaction does not occur for TF bound to decoy regions where transcriptional regulatory proteins are not recruited (Figure 15). In this study, we did not find a specific regulatory protein responsible for affecting only TF-promoter interaction (Appendix B). However, further genetic approaches could help to confirm this possibility. If this scenario is correct, do all the TF binding sites within a promoter weakly associate with the TF due to the interaction with those potential regulatory proteins? We found that the binding affinity of a promoter containing seven tetO binding sites is strong and comparable to the tetO array affinity while the affinity of a single tetO site promoter is not (Figure 8). With the binding affinities of 7xtetO promoter and tetO array predicted in Figure 4, this implies that only a subset of the seven binding sites bind tTA weakly and those sites are responsible for downstream gene expression. It is possible that only binding sites close to a transcription start site weakly interact with tTA,

inducing gene expression. In this case, as the tTA level increases, the binding sites of strong affinity will sequester tTA prior to weak affinity binding sites, which may create a delay in gene expression.

If a promoter containing multiple TF binding sites is used in a single input module (SIM) where a single TF, the master regulator, targets multiple promoters and affects their expression (Alon, 2006), what will happen? A SIM can synchronize the expression of a group of genes within the module in a temporally ordered manner (Kalir et al, 2001; Zaslaver et al, 2004). The temporal ordering is thought to be encoded by the affinity of each gene's promoter to the master regulator. Upon activation of SIM, as the TF level slowly increases the promoter of strong affinity will activate earlier because the TF first bind the strong affinity promoter before its binding to the weak (Kalir et al, 2001). In light of our results, there may be additional, previously unappreciated ways that lead to distinct temporal separation. If a promoter contains multiple binding sites for the master regulator, since a subset of those binding sites strongly sequester the regulators, other promoters containing fewer sites or a single binding site will activate slowly creating a threshold response which could induce a potential temporal ordering. Moreover, because the number of multiple binding sites is likely to vary over time due to the repeats' variability, the expression delay of weak affinity promoters could vary with the number of binding sites within the promoter containing multiple TF binding sites. If the promoter affinity correlates with the number of binding sites within the promoter, we could tune the gene expression delay by changing the number of the binding sites.

Multiple TF binding sites within a promoter may not only create a temporal ordering but also lead to variability in gene expression. It has been shown that multiple binding sites within a promoter increase gene expression noise by up-regulating burst size (To & Maheshri, 2010). Whether a high affinity of the multiple binding sites correlates with the burst size is still not clear. However, it is possible that strong, multisite promoter binding may increase the local concentration of TF and thus facilitate a transcriptional re-initiation.

A possible mechanism to answer the second question could be that an initial TF binding to a decoy site, which is strong, prevents additional TF binding to neighboring sites (Figure 15). Several studies suggest that the interaction between DNA binding proteins associated with multiple clustered target sites may be negatively cooperative. For example, *in vitro* measurements for lacI repressor binding to a 256xlacO array, where each lacO unit is separated by 15bp of nonspecific sequence, reveal that most of the lacO sites are unoccupied even at high lacI levels, and each occupied site is distant from neighboring occupied sites (Wang et al, 2005). Another study also indirectly shows that tetR proteins bound to multiple tetO sites negatively impacts the binding affinity of tetR to neighboring tetO sites (Amit et al, 2011).

This negative cooperativity in protein binding to clustered binding sites could be due to direct interactions between the bound proteins or indirect effects whereby protein binding alters structural properties of nearby DNA. If the length of a spacer between clustered binding sites is smaller than ~5-6 bp, a steric effect will cause the negative cooperativity



upon protein binding (Pan et al, 2011). Although previous studies (Wang et al, 2005; Amit et al, 2011) used an array where each binding site is separated from one another by at least 15bp to eliminate the steric effect, they still found a negative cooperative binding effect. Therefore, it is unlikely that the observed negative cooperativity is driven by steric hindrance. A suggested indirect effect has been that initial protein binding bends or twists DNA hampering subsequent protein binding to nearby sites (Hogan et al, 1987; Wang et al, 2005; Chen et al, 2010). Interestingly, numerous TFs induce DNA bending upon their binding (Kim et al, 1989; van der Vliet et al, 1993). Since the DNA persistence length is around 150bp, it may be possible that the upper limit that the bending effect propagates is ~150bp. However, we found that negative cooperativity in binding could occur across 500bp (Figure 4) because two 127x tetO arrays separated by 500bp were equally potent to one 240x tetO array (Lee & Maheshri, 2012). If a TF binding to repeated decoy sites inhibits the neighbor site binding through negative cooperativity, the strong binding sites, which are distant from one another by the range of negative cooperative binding, will sequester TF first, and the remaining binding regions will sequester TF with weak affinity. This may induce the affinity decrease with increasing TF levels as we found in the ChIP measurements (Figure 8).

However, it is still possible that this negative cooperative binding effect stems from unknown interactions between the TF and other proteins *in vivo*. To find whether it comes purely from TF and DNA interactions, *in vitro* binding measurements will be useful. When the TF interacts with repeated decoy sites, it may create various complexes containing

various ratios of TF to decoy sites, as has been shown in non-specific TF binding for DNA (Graham et al, 2011). In the absence of negative cooperative binding, equilibrium binding determines the mean ratio. However, if there is a negative cooperative binding effect, the mean ratio will reduce with TF levels. By using a gel shift assay, we might differentiate each TF-decoy DNA complex and measure the amount of each complex as well as the mean ratio between TF and decoy DNA of the complex. By comparing this ratio with a thermodynamic equilibrium model behavior, we might test whether negative cooperative binding originates solely from the TF-DNA interaction.

Whether the range of this negative cooperative effect varies with the total number of binding sites is not clear. If the propagation range of the negative cooperative effect is fixed, the number of strong TF binding will increase with the array size and therefore a long contiguous array will be more effective than a small array for altering gene expression. However, our finding that a small array of binding sites is as potent as a long contiguous array suggests that the range of negative cooperative effect may increase with the number of binding sites. If so, the number of repeated decoy sites need not be large enough to effectively alter gene expression. Based on our results, the threshold number of binding sites beyond which further contiguous sites have little or no effect lies between 37 and 67 (Figure 3). Though the array effect for influencing gene expression is restrained by the array size, this can be overcome by multiplying the copy number of the array (Figure 7). Probably, whole gene duplication or segmental duplication of decoy binding sites during evolution may significantly increase the decoy effect if the total TF level stays constant.

Is utilizing intergenic decoy TF binding sites for gene regulation general in natural system? For yeast, due to its compact genome, intergenic TRs are not predominant, but even here decoy TF binding sites may have a functional role (Richard et al, 1997; Richard et al, 2008). For example, previous studies show that there are repeated decoy binding sites for Arr1p, an activator of *ARR2* and *ARR3* gene expression during arsenite detoxification (Maciaszczyk et al, 2004; Mukhopadhyay et al, 1998; Wysocki et al, 1997). If these decoy sites strongly sequester Arr1p, which is activated by intracellular arsenate, Arr1p-responsive arsenate-arsenite conversion and arsenite transport will exhibit a threshold-like dose-response with arsenate levels. This threshold response may be useful to prevent undesired *ARR2* and *ARR3* activation by Arr1p noise possibly appearing at low levels.

#### 4.1 References

- Alon, U. (2006). An Introduction to Systems Biology: Design Principles of Biological Circuits. Boca Raton: CRC
- Amit, R., Garcia, H.G., Phillips, R., and Fraser, S.E. (2011). Building enhancers from the ground up: a synthetic biology approach. *Cell* 146, 105-118.
- Chen, B., Xiao, Y., Liu, C., Li, C., and Leng, F. (2010). DNA linking number change induced by sequence-specific DNA-binding proteins. *Nucleic Acids Res.* 38, 3643-3654.
- Graham, J.S., Johnson, R.C., and Marko, J.F. (2011). Concentration-dependent exchange accelerates turnover of proteins bound to double-stranded DNA. *Nucleic Acids Res.* 39, 2249-2259.

Hogan, M.E., and Austin, R.H. (1987). Importance of DNA stiffness in protein-DNA binding specificity. *Nature* 329, 263-266.

Kalir, S., McClure, J., Pabbaraju, K., Southward, C., Ronen, M., Leibler, S., Surette, M.G., and Alon, U. (2001). Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. *Science* 292, 2080-2083.

Kamionka, A., Bogdanska-Urbaniak, J., Scholz, O., and Hillen, W. (2004). Two mutations in the tetracycline repressor change the inducer anhydrotetracycline to a corepressor. *Nucleic Acids Res.* 32, 842-847.

Kim, J., Zwieb, C., Wu, C., and Adhya, S. (1989). Bending of DNA by gene-regulatory proteins: construction and use of a DNA bending vector. *Gene* 85, 15-23.

Lee, T.H., and Maheshri, N. (2012). A regulatory role for repeated decoy transcription factor binding sites in target gene expression. *Mol. Syst. Biol.* 8, 576.

Maciaszczyk, E., Wysocki, R., Golik, P., Lazowska, J., and Ulaszewski, S. (2004). Arsenical resistance genes in *Saccharomyces douglasii* and other yeast species undergo rapid evolution involving genomic rearrangements and duplications. *FEMS Yeast Res.* 4, 821-832.

Mukhopadhyay, R., and Rosen, B.P. (1998). *Saccharomyces cerevisiae* ACR2 gene encodes an arsenate reductase. *FEMS Microbiol. Lett.* 168, 127-136.

Pan, Y., and Nussinov, R. (2011). The role of response elements organization in transcription factor selectivity: the IFN-beta enhanceosome example. *PLoS Comput. Biol.* 7, e1002077.

Richard, G.F., and Dujon, B. (1997). Trinucleotide repeats in yeast. *Res. Microbiol.* 148, 731-744.

Richard, G.F., Kerrest, A., and Dujon, B. (2008). Comparative genomics and molecular dynamics of DNA repeats in eukaryotes. *Microbiol. Mol. Biol. Rev.* 72, 686-727.

To, T.L., and Maheshri, N. (2010). Noise can induce bimodality in positive transcriptional feedback loops without bistability. *Science* 327, 1142-1145.

van der Vliet, P.C., and Verrijzer, C.P. (1993). Bending of DNA by transcription factors. *Bioessays* 15, 25-32.

Wang, Y.M., Tegenfeldt, J.O., Reisner, W., Riehn, R., Guan, X., Guo, L., Golding, I., Cox, E.C., Sturm, J., and Austin, R.H. (2005). Single-molecule studies of repressor–DNA interactions show long-range interactions. *Proc. Natl. Acad. Sci. U. S. A.* 102, 9796-9801.

Wysocki, R., Bobrowicz, P., and Ulaszewski, S. (1997). The *Saccharomyces cerevisiae* ACR3 gene encodes a putative membrane protein involved in arsenite transport. *J. Biol. Chem.* 272, 30061-30066.

Zaslaver, A., Mayo, A.E., Rosenberg, R., Bashkin, P., Sberro, H., Tsalyuk, M., Surette, M.G., and Alon, U. (2004). Just-in-time transcription program in metabolic pathways. *Nat. Genet.* 36, 486-491.

## Appendices

### **Appendix A. A longer tetO array sequesters more tTA but exhibits a similar potency for lowering the target gene expression compared to a shorter array**

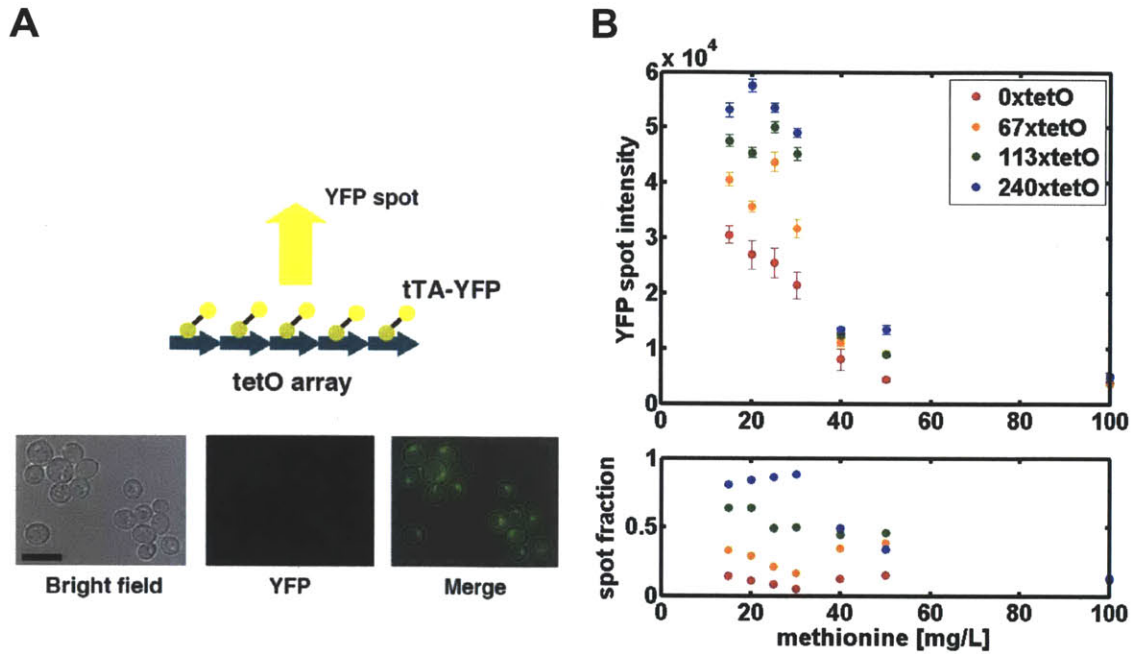
Since 67x, 113x and 240x tetO arrays showed similar reduction for the target gene expression, an emerging hypothesis is that those arrays sequester a similar amount of tTAs. To test this hypothesis in addition to the ChIP, we expressed a nuclear-localized tTA-YFP fusion from the *MET3* promoter in a cell containing each array. To monitor promoter activation, we placed a 7xtetO promoter driving CFP reporter.

When we titrated tTA-YFP expression by methionine, corresponding CFP dose-response were similar for the three arrays consistent with Figure 3, but always lower than the expression of a cell without an array (Figure 5). As the tTA-YFP level increased, a bright YFP spot appeared inside the cell (Figure 16A) and we assumed that the YFP spot intensity is proportional to the amount of tTA bound in the tetO array as in the previous study (Karpova et al, 2008). Due to the high autofluorescence level (mean YFP autofluorescence ~5000 A.U., data not shown), we were unable to measure the amount of sequestered tTA at high methionine levels where only a small fraction of cells expressed a visible YFP dot (Figure 16B). At high tTA levels, most of cells expressed clear YFP spots and we found that the spot intensity for 240x tetO array is higher than those of 67x and 113x tetO arrays, suggesting that a longer array sequesters more tTA. Since the reporter expression is similar for three arrays, we reason that not all the tTAs bound to the array are effective for altering

the promoter response as we also suggested from our ChIP measurement. Therefore, we propose that the number of effective tTA binding leading to a down-regulated gene expression is similar for three arrays.

ChIP results showed that a per base occupancy of 67x tetO array is several fold higher than that of 240xtetO array at low to intermediate tTA levels where an inflection of the sigmoidal-like dose-response occurs, suggesting that a similar number of tTAs are sequestered for three arrays. This may correspond to the effective tTA binding regime leading to the reduction of gene expression. However, at high tTA levels, the difference in per base occupancy decreased probably due to the occurrence of non-effective tTA binding. This means the total number of tTAs bound to the 240x array becomes higher than the 67x tetO at this tTA level, which is consistent with the YFP spot result.

If the size of the effective tTA binding regime of the 67xtetO array is identical to that of the 240xtetO array as shown here and in ChIP measurement, it can explain why two 67xtetO arrays are more potent than a 240x array as discussed in Chapter 3.



**Figure 16. 240x tetO array sequesters more tTA than 67x and 113x tetO arrays.** (A) When tTA-YFP is sequestered at repeats, it can be visualized as a bright spot using fluorescence microscopy. Therefore, the amount sequestered can be monitored by determining the fraction of cells with a bright spot and/or the intensity of the spot (representing sequestered tTA-YFP). Scale bar represents 5 $\mu$ m. (B) Spot fraction and intensity was from single cell data. A custom algorithm to determine spots was used, and the red dots represent a false positive rate corresponding to “spots” detected using images of cells containing no repeats. At low methionine levels (high tTA levels), the spot intensities for 240xtetO are higher than 67x and 113x arrays, indicating that a longer array sequesters more tTA-YFP. Error bars represent bootstrap s.d. of spot intensities measured at each methionine level.

## References

Karpova, T.S., Kim, M.J., Spriet, C., Nalley, K., Stasevich, T.J., Kherrouche, Z., Heliot, L., and McNally, J.G. (2008). Concurrent fast and slow cycling of a transcriptional activator at an endogenous promoter. *Science* 319, 466-469.



## **Appendix B. Possible factors influencing tTA-promoter and tTA-array binding**

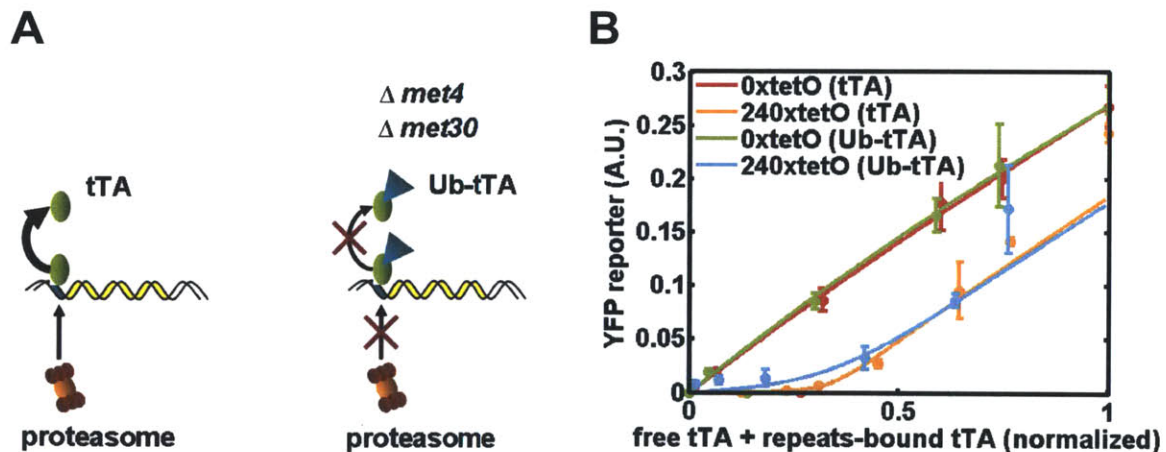
The mathematical model derived in the Chapter 2 and 3 explains solely the interactions between tTA, promoter and array. However, a cellular nucleus contains numerous other nuclear proteins that also may dynamically interact with tTA specifically and non-specifically. In the presence of these interactions, the binding affinities of promoter and array for tTA may be altered. In addition, array-driven multivalency effect or chromatin structure close to tTA binding region may also affect the tTA binding for its target sites. Here, we discuss potential factors that might change the binding properties of tTA for the promoter and the array.

### **Proteasome-mediated active disassembling of tTA in the promoter does not affect promoter binding affinity**

A number of eukaryotic TFs are degraded by proteasome-mediated pathway. They are covalently ubiquitylated and recognized by proteasome. Interestingly, the domain for ubiquitylation (degron) often overlaps with TF's transcriptional activation domain, suggesting a dual role of ubiquitylation (Salghetti et al, 2001): TF degradation and promoter activation. For example, a study for the fusion transcriptional activator of LexA DNA binding domain and VP16 activation domain shows that when the VP16 domain is mono-ubiquitylated it can induce downstream gene expression though it is resistant to the degradation (Salghetti et al, 2001). Other studies show that the promoter-bound TF are likely to be rapidly degraded by proteasome (Molinari et al, 1999). These results suggest a possibility that sequentially ubiquitylated TFs, which are bound to the promoter, may

actively disassembled by proteasome. If so, TF/promoter binding affinity will be reduced, creating the difference compared to the array affinity come weaker. Since TF recruited to the array dose not induce any gene expression, the active TF disassembling might not occur in the array.

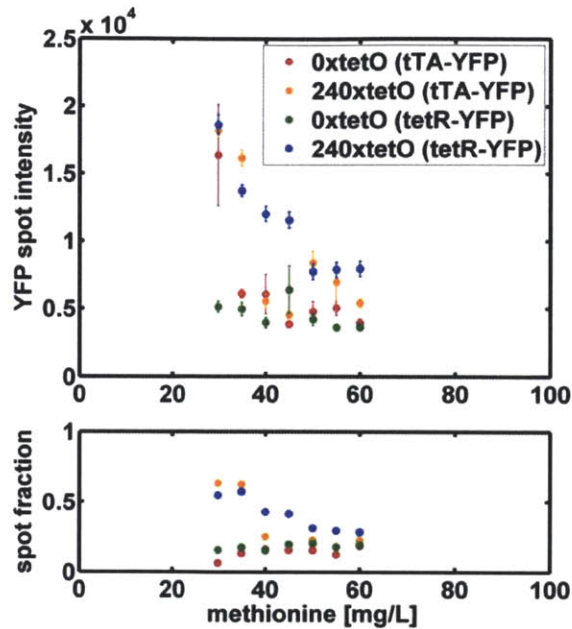
To test whether the stabilized tTA increases its binding affinity for promoter and thus decrease the extent of a sigmoidal dose-response, we expressed a monoubiquitinated tTA in a strain background that lacks the E3 ubiquitin ligase *met30* specific for tTA (To & Maheshri, 2010). However, we found that there was no difference in the sigmoidal dose-response, suggesting that interactions with ubiquitin ligases or proteasome did not influence binding at the promoter (Figure 17). This also provides evidence that the repeats were not having an inhibitory effect by somehow enhancing tTA degradation.



**Figure 17. Stabilizing tTA does not eliminate the sigmoidal-like response in the presence of the tetO array.** (A) To stabilize tTA, we expressed a monoubiquitylated form in a strain background where Met30p, the tTA-specific Ub-ligase, is deleted. *MET4* is also deleted to prevent lethality (Salghetti SE et al. 2001). (B) In the presence of 240xtetO array cells expressing Ub-tTA exhibit the equivalent sigmoidal-like response to cells expressing tTA. Error bars represent the s.d. of two replicates.

### **The interaction between tTA activation domain and general transcription factors affects tTA-tetO binding affinity**

The binding of tTA to repeated decoys may be altered by its ability to recruit transcriptional machinery, potentially forming a multiprotein complex that is more stable and hence possesses a higher affinity. If true, tetR (tTA without the VP16 activation domain) should have a reduced affinity for the decoy sites. When we compared spot fraction and spot intensity of methionine-titrated tetR-YFP to the tTA-YFP, we found no difference at high (>50 mg/L) and low (<35 mg/L) methionine levels (Figure 18). However, at the intermediate methionine level (40-45 mg/L) where a sharp slope change of the dose-response occurred, the spot intensity of tetR-YFP was much higher than that of tTA-YFP, suggesting tetR-YFP binding is stronger than tTA-YFP binding. This suggests that the interaction between the VP16 activation domain and transcriptional machinery may weaken the tTA-YFP binding to decoy sites or leads active tTA degradation for reducing the free tTA level.



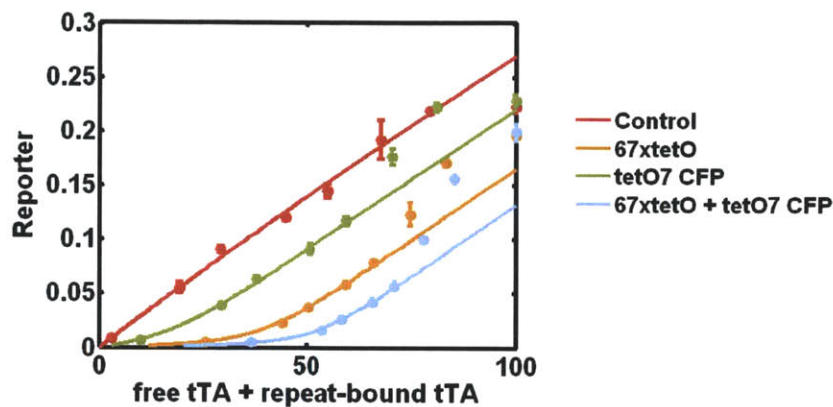
**Figure 18. tetR-YFP binding to tetO array is stronger than tTA-YFP binding at intermediate methionine levels.** The YFP spot intensities for sequestered tTA-YFP and tetR-YFP were measured by fluorescence microscopy as Figure 16. Spot fraction and intensity were from single cell data. A custom algorithm to determine spots was used, and the red and green dots represent a false positive rate corresponding to “spots” detected using images of cells containing no repeats. At low methionine levels (high tTA levels), the spot intensities for tTA-YFP and tetR-YFP are similar. On the other hand, at the intermediate methionine level (40-45mg/L), the intensity for tetR-YFP is higher than for tTA-YFP, indicating that the array binding affinity for tetR-YFP is stronger than for tTA-YFP.

### **The difference in chromatin architecture between promoter and decoy array does not affect their binding affinities for tTA**

Eukaryotic chromatin structure consists of repetitive nucleosomes. Here, 147bp DNA sequences wrap a histone octamer to form a nucleosome. Several studies reveal that the

nucleosome structure negatively affects the gene expression. For example, nucleosome loss in budding yeast activates downstream *PHO5* promoter, leading gene expression (Han et al, 1988). Recruiting proteins involved in chromatin remodeling and histone modification into the promoter facilitates the gene expression by removing or shifting nucleosome and thus enabling the access of TFs (Li et al, 2007). However, recent studies also show that upon the gene expression the nucleosome arrangement of most yeast promoters is unchanged raising a question for nucleosome repositioning (Zawadzki et al, 2009).

If the chromatin structure in the promoter, though whether the nucleosome moves or not is unclear, negatively affects TF recruitment to the promoter, it may affect TF-promoter binding during active transcription. To test it, we introduced a low copy centromeric plasmid containing 67x repeats directly upstream of a 7xtetO promoter. We measured a promoter dose-response from a chromosomally integrated 7xtetO reporter driving YFP expression. Positioning repeats near an active promoter had no effect on their inhibitory potential and the sigmoidal-like response remained, suggesting the regional chromatin environment is not a source of promoter binding affinity for tTA (Figure 19).



**Figure 19. Positioning the tetO array near the tetO promoter does not eliminate the sigmoidal-like response.** (A) A centromeric plasmid-borne 67xtetO array was placed either alone or next to another 7xtetO promoter driving CFP, and the dose response of a chromosomally integrated YFP reporter was monitored. Placing the 67xtetO array upstream of the 7xtetO promoter did not diminish its potency and the sigmoidal-like response.

### **The interaction between tTA in the promoter and transcription-regulating proteins may not affect tTA-promoter binding**

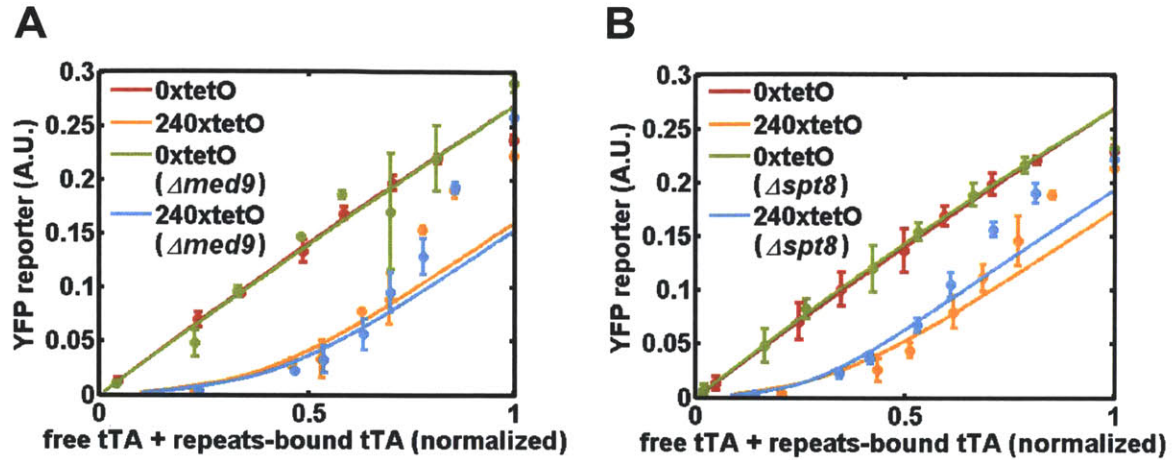
The promoter affinity for tTA may be lower than the tetO array affinity because of active, energy-dependent processes occurring at these sites, mediated by proteins associated with transcription such as chromatin remodeler and cofactor. For example, Gcn4p, an activator that responds to amino acid starvation interacts with RSC, SAGA, SWI/SNF and SRB mediator (Swanson et al, 2003). A deletion of the *RSC2* gene that encodes a chromatin remodeler, Rsc2p, leads to impaired recruitment of Ace1p into *CUP1* promoter, which may lower the Ace1p binding affinity for the promoter (Karpova et al, 2008). We tested possible

effects of unknown interactions between tTA and transcription-regulating proteins on tTA-promoter binding by employing a genetic approach with mutant strains for genes.

Recent studies have implicated Rsc2p in actively disassembling non-specifically bound activators from DNA (Karpova et al, 2004). However, at the *CUP1* promoter in budding yeast, a *rsc2* may serve to decrease promoter occupancy of the activator (whether the effect is direct or indirect is not known) (Karpova et al, 2008). Therefore we disrupted *RSC2* to determine if the binding affinity for activators at the promoter would decrease, further increasing the sigmoidal-like response. However, the *rsc2* mutant behaved identically to wild type (data not shown).

We observed similar results when we disrupted *MED9/CSE2*, a component of the Middle domain of Mediator known to have both positive and negative effects at specific genes; and *SPT8*, a subunit of SAGA which inhibits binding of SAGA to PIC's in the absence of activator but stimulates transcription in the presence of activator (Warfield et al, 2004) (Figure 20). Expression was lower in a *med9*, suggesting that the Med9p/Med10p module plays a positive role at the tetO promoter.



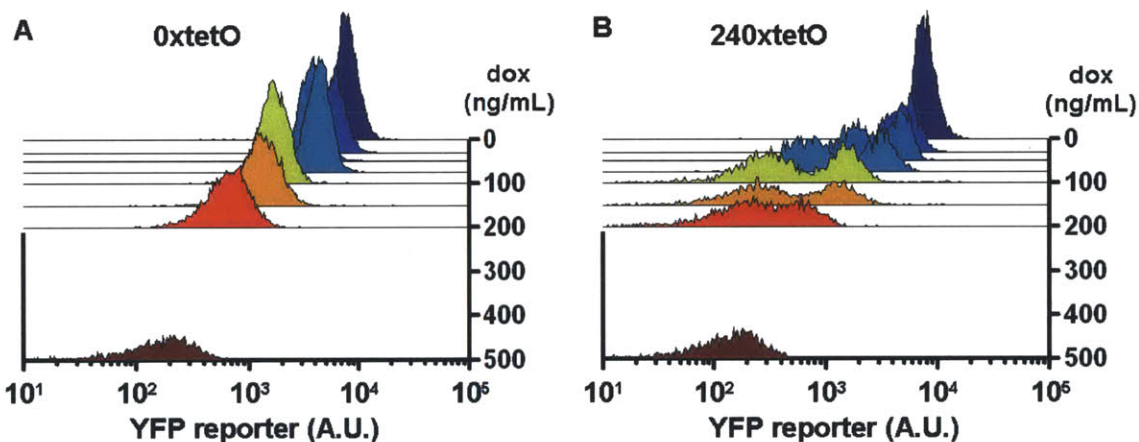


**Figure 20. Sigmoidal-like response for *med9* and *spt8* gene disruption mutants.** The extent of sigmoidal-like response is not related with the function of Med9p, one of the subunits of yeast mediator and Spt8p, one of the subunits of yeast SAGA complex. tTA was driven by constitutive *MYO2* promoter at the *ADE2* locus and titrated by doxycycline. Activation of the 7xtetO promoter driving YFP at the *URA3* locus was monitored by flow cytometry in the absence or presence of 240xtetO array integrated at the *HIS3* locus. To compare the extent of sigmoidal-like response, the YFP expression for the wild type and mutant strain was normalized such that the YFP expressions at 0 ng/mL doxycycline level were identical. In the presence of 240xtetO array, the (A) *med9* and (B) *spt8* disruption mutants exhibit nearly identical responses as compared to the wild type strain. Dots represent experimental data and solid line the model fits. Error bars represent s.d. of biological triplicates.

Interestingly, when we deleted the *SIN4* – a global transcriptional regulator and member of the Tail module of Mediator (Sin4/Gal11 complex) that has both positive and negative roles on transcription regulation through direct interaction with activators and repressors (Lewis & Reinberg, 2003), in the absence of the tetO array, *sin4* $\Delta$  reduces reporter expression as expected given the known interaction of VP16 with the Tail domain (Myers et al, 1999;



Park et al, 2000). However, it unexpectedly generates a bimodal gene expression pattern *only* in the presence of the tetO array (Figure 21). While it is difficult to speculate on mechanism without further studies, this is additional evidence that repeated decoy sites can have qualitative effects on gene expression in different genetic backgrounds. A genome-wide analysis of repeat-dependent effects using the yeast deletion library might shed further light on mechanism and the role of chromatin remodeling on both binding to the tetO array and the promoter.



**Figure 21. Sigmoidal-like response for *sin4* gene disruption mutant**

(A) Flow cytometric analysis of YFP response of the 7xtetO promoter in a *sin4* background shows a graded response. (B) The response becomes bimodal upon integration of a 240xtetO array at the *HIS3* locus.

## References

Han, M., and Grunstein, M. (1988). Nucleosome loss activates yeast downstream promoters in vivo. *Cell* 55, 1137-1145.

Han, S.J., Lee, J.S., Kang, J.S., and Kim, Y.J. (2001). Med9/Cse2 and Gal11 modules are required for transcriptional repression of distinct group of genes. *J. Biol. Chem.* 276, 37020-37026.

Karpova, T.S., Kim, M.J., Spriet, C., Nalley, K., Stasevich, T.J., Kherrouche, Z., Heliot, L., and McNally, J.G. (2008). Concurrent fast and slow cycling of a transcriptional activator at an endogenous promoter. *Science* 319, 466-469.

Karpova, T.S., Chen, T.Y., Sprague, B.L., and McNally, J.G. (2004). Dynamic interactions of a transcription factor with DNA are accelerated by a chromatin remodeller. *EMBO Rep* 5, 1064-1070.

Kitov, P.I., and Bundle, D.R. (2003). On the nature of the multivalency effect: a thermodynamic model. *J. Am. Chem. Soc.* 125, 16271-16284.

Lewis, B.A., and Reinberg, D. (2003). The mediator coactivator complex: functional and physical roles in transcriptional regulation. *J. Cell. Sci.* 116, 3667-3675.

Li, B., Carey, M., and Workman, J.L. (2007). The role of chromatin during transcription. *Cell* 128, 707-719.

Locasale, J.W. (2008). Allovalency revisited: an analysis of multisite phosphorylation and substrate rebinding. *J. Chem. Phys.* 128, 115106.

Mammen, M., Choi, S.K., and Whitesides, G.M. (1998). Polyvalent Interactions in Biological Systems: Implications for Design and Use of Multivalent Ligands and Inhibitors. *Angew. Chem. Int. Ed.* 37,2754–2794.

Molinari, E., Gilman, M., and Natesan, S. (1999). Proteasome-mediated degradation of transcriptional activators correlates with activation domain potency in vivo. *EMBO J* 18, 6439-6447.

Myers, L.C., Gustafsson, C.M., Hayashibara, K.C., Brown, P.O., and Kornberg, R.D. (1999). Mediator protein mutations that selectively abolish activated transcription. *Proc. Natl. Acad. Sci. U. S. A.* 96, 67-72.

Park, J.M., Kim, H.S., Han, S.J., Hwang, M.S., Lee, Y.C., and Kim, Y.J. (2000). In vivo requirement of activator-specific binding targets of mediator. *Mol. Cell. Biol.* 20, 8709-8719.

Salghetti, S.E., Caudy, A.A., Chenoweth, J.G., and Tansey, W.P. (2001). Regulation of Transcriptional Activation Domain Function by Ubiquitin. *Science* 293, 1651-1653.

Swanson, M.J., Qiu, H., Sumibcay, L., Krueger, A., Kim, S.J., Natarajan, K., Yoon, S., and Hinnebusch, A.G. (2003). A multiplicity of coactivators is required by Gcn4p at individual promoters in vivo. *Mol. Cell. Biol.* 23, 2800-2820.

To, T.L., and Maheshri, N. (2010). Noise can induce bimodality in positive transcriptional feedback loops without bistability. *Science* 327, 1142-1145.

Warfield, L., Ranish, J.A., and Hahn, S. (2004). Positive and negative functions of the SAGA complex mediated through interaction of Spt8 with TBP and the N-terminal domain of TFIIA. *Genes Dev.* 18, 1022-1034.

Zawadzki, K.A., Morozov, A.V., and Broach, J.R. (2009). Chromatin-dependent transcription factor accessibility rather than nucleosome remodeling predominates during global transcriptional restructuring in *Saccharomyces cerevisiae*. *Mol. Biol. Cell* 20, 3503-3513.

### Appendix C. Yeast strains used in this study

Strain	Relevant Genotype	Parent Strain	Reference
Y1	MATa <i>trp1-1 can 1-100 leu2-3, 112 his3-11, 5 ura3 GAL+</i>	W303	Laboratory collection
Y3	MATa <i>ade2-1 trp1-1 can 1-100 leu2-3, 112 his3-11, 5 ura3 GAL+</i>	W303	Laboratory collection
Y6	MATa <i>ade2-1 trp1-1 can 1-100 leu2-3, 112 his3-11, 5 ura3 GAL+</i>	W303	Laboratory collection
Y136	MATa <i>ade2::MYO2pr-tTA-ADE2</i>	Y3	This study
Y137	MATa <i>his3::240xtetO-HIS3</i>	Y3	This study
Y138	MATa <i>his3::240xtetO-HIS3</i>	Y6	This study
Y125	MATa <i>his3::1xtetOpr-tTA-HIS3</i> <i>leu2Δ::1xtetOpr-venus</i>	Y1	Laboratory collection
Y215	MATa <i>ura3Δ::7xtetOpr-cerulean-kanMX6</i>	Y3	This study
Y216	MATa <i>ura3Δ::7xtetOpr-venus-kanMX6</i>	Y6	This study
Y226	MATa <i>ura3Δ::7xtetOpr-venus-kanMX6</i> MATa <i>ade2::MYO2pr-tTA-ADE2</i>	Y138	This study
Y229	<i>leu2::PGK1pr-tdTomato ura3Δ::7xtetOpr-cerulean-kanMX6</i> MATa <i>ade2::MYO2pr-tTA-ADE2</i>	Y137	This study
Y231	<i>leu2::PGK1pr-tdTomato ura3Δ::7xtetOpr-cerulean-kanMX6</i>	Y3	This study
Y250	MATa <i>met4Δ::TRP1 met30Δ::LEU2</i>	Y3	Laboratory collection
Y561	MATa <i>ura3Δ::7xtetOpr-venus-kanMX6</i>	Y136	This study
Y562	MATa <i>ura3Δ::1xtetOpr-venus-kanMX6</i>	Y136	This study
Y563	MATa <i>leu2::ACT1pr-tdTomato</i>	Y561	This study
Y564	MATa <i>his3::37xtetO-HIS3</i>	Y561	This study
Y565	MATa <i>his3::67xtetO-HIS3</i>	Y561	This study
Y566	MATa <i>his3::113xtetO-HIS3</i>	Y561	This study
Y567	MATa <i>his3::240xtetO-HIS3</i>	Y561	This study
Y568	MATa <i>his3::37xtetO-HIS3</i>	Y562	This study
Y569	MATa <i>his3::67xtetO-HIS3</i>	Y562	This study
Y570	MATa <i>his3::113xtetO-HIS3</i>	Y562	This study
Y571	MATa <i>his3::240xtetO-HIS3</i>	Y562	This study
Y572	MATa <i>leu2Δ::MET3pr-NLS-tTA-YFP</i>	Y215	This study
Y573	MATa <i>his3::67xtetO-HIS3</i>	Y572	This study
Y574	MATa <i>his3::113xtetO-HIS3</i>	Y572	This study
Y575	MATa <i>his3::240xtetO-HIS3</i>	Y572	This study
Y580	MATa <i>trp1::67xtetO-TRP1 ura3::67xtetO-URA3</i>	Y125	This study
Y581	MATa <i>leu2Δ::MYO2pr-Ub-tTA</i>	Y250	This study

	<i>ura3::7xtetOpr-venus-kanMX6</i>		
Y582	MATa/ $\alpha$	Y216, Y231	This study
Y583	MATa/ $\alpha$	Y216, Y229	This study
Y584	MATa/ $\alpha$	Y226, Y229	This study
Y585	MATa <i>rsc2<math>\Delta</math>::LEU2</i>	Y561	This study
Y586	MATa <i>rsc2<math>\Delta</math>::LEU2</i>	Y567	This study
Y587	MATa <i>spt8<math>\Delta</math>::LEU2</i>	Y561	This study
Y588	MATa <i>spt8<math>\Delta</math>::LEU2</i>	Y567	This study
Y589	MATa <i>med9<math>\Delta</math>::LEU2</i>	Y561	This study
Y590	MATa <i>med9<math>\Delta</math>::LEU2</i>	Y567	This study
Y591	MATa <i>sin4<math>\Delta</math>::LEU2</i>	Y561	This study
Y592	MATa <i>sin4<math>\Delta</math>::LEU2</i>	Y567	This study
Y641	MATa <i>ade2::MYO2pr-tTA-3xHA-ADE2</i>	Y562	This study
Y753	MATa <i>ade2::MYO2pr-tTA-3xHA-ADE2</i>	Y561	This study

---

## Appendix D. Plasmids used in this study

Plasmid	Base vector	Relevant information	Construction information
B030	pLAU44	Lau IF 2003 <i>et al.</i>	Gift from D. Sherrat via A. Grossman
B049	JRL2	<i>PGK1</i> promoter-tdTomato	Laboratory Stock ( <i>LEU2</i> integrating vector)
B193	pRS303	67xtetO	XhoI/SacI digestion of 67xtetO from B030 and ligation into XhoI/SacI digested pRS303
B194	pRS303	113xtetO	EcoRI/XbaI digestion of 113xtetO from B030 and ligation into EcoRI/XbaI digested pRS303
B196	pRS303	240xtetO	XhoI/XbaI digestion of 240xtetO from B030 and ligation into XhoI/XbaI digested pRS303
B224	pRS303	7xtetO promoter-NLS-tTA-venus	Laboratory Stock
B340	TOPO (Invitrogen)	67xtetO	XhoI/SacI digestion of 67xtetO from B030 and ligation into XhoI/SacI digested TOPO vector
B341	pRS426	<i>ACT1</i> promoter-tdTomato	BamHI/NotI digestion of <i>ACT1</i> pr-tdTomato (PCR product) from B597 and ligation into BamHI/NotI digested pRS426
B345	pRS313	67xtetO	XhoI/SacI digestion of 67xtetO from B030 and ligation into XhoI/SacI digested pRS313
B346	pRS316	67xtetO	XhoI/HindIII digestion of 67xtetO from B340 and ligation into XhoI/HindIII digested pRS316
B350	pRS426	<i>ACT1</i> promoter-tdTomato-67xtetO	XhoI/HindIII digestion of 67xtetO from B340 and ligation into XhoI/HindIII digested B341
B352	pRS316	15xtetO	BamHI digestion of 15xtetO from B030 and ligation into BamHI digested pRS316
B356	pRS315	67xtetO	XhoI/HindIII digestion of 67xtetO from B340 and ligation into XhoI/HindIII digested pRS315
B359	pRS316	37xtetO	BamHI/EcoRV digestion of 37xtetO from B030 and ligation into BamHI/EcoRV digested pRS316
B553	JRL2	<i>MET3</i> promoter-NLS-tTA-venus	Sall/NotI digestion of B591 and ligation into Sall/NotI digested B049
B583	pFA6a kanMX6	7xtetOpr-venus	PacI/BamHI digestion of 7xtetO, BamHI/AscI digestion of venus and ligation into PacI/AscI digested vector

B584	pFA6a kanMX6	7xtetO promoter- cerulean-kanMX6	Digestion of 7xtetO promoter – cerulean (PacI/AscI) and ligation into PacI/AscI digested pFA6a kanMX6 vector
B585	pFA6a kanMX6	1xtetOpr-venus	PacI/BamHI digestion of 1xtetO, BamHI/AscI digestion of venus and ligation into PacI/AscI digested vector
B586	pRS316	240xtetO	XhoI/XbaI digestion of 240xtetO from B030 and ligation into XhoI/XbaI digested pRS316
B587	pRS303	37xtetO	XhoI/XbaI digestion of 37xtetO from B359 and ligation into XhoI/XbaI digested pRS303
B588	pRS316	127xtetO	XhoI/EcoRI digestion of 127xtetO from B030 and ligation into XhoI/EcoRI digested pRS316
B590	pRS316	<i>MET3</i> promoter	EcoRI/BamHI digestion of <i>MET3</i> promoter (PCR product) and ligation into EcoRI/BamHI digested pRS316
B591	pRS316	<i>MET3</i> promoter- NLS-tTA-venus	BamHI/NotI digestion of NLS-tTA-venus from B224 and ligation into BamHI/NotI digested B590
B592	pRS304	67xtetO	XhoI/SacI digestion of 67xtetO from B030 and ligation into XhoI/SacI digested pRS304
B594	JRL2	MET3pr-NLS- tetR-venus	BamHI/NotI digestion of NLS-tetR-venus (PCR product) and ligation into BamHI/NotI digested B553
B595	pRS316	67xtetO-7xtetO cerulean	EcoRI/NotI digestion of 7xtetO-cerulean (PCR product) and ligation into EcoRI/NotI digested B346
B596	pRS316	7xtetO promoter - cerulean	EcoRI/NotI digestion of 7xtetO promoter-cerulean (PCR product) and ligation into EcoRI/NotI digested pRS316
B597	JRL2	<i>ACT1</i> promoter- tdTomato	Sall/EcoRI digestion of <i>ACT1</i> promoter (PCR product) and ligation into Sall/EcoRI digested B049
B598	pRS426	127xtetO	XhoI/EcoRI digestion from B030 and ligation into XhoI/EcoRI digested pRS426
B647	pFA6a HisMX6	3xHA	Gift from Erin O'Shea
B692	pRS316	6xtetO	XhoI/KpnI digestion of 6xtetO fragment and ligation into XhoI/KpnI digested pRS316