FUNCTIONAL DISSECTION OF RNA POLYMERASE II ACTIVE SITE AND MECHANISM OF ACTION OF TRANSCRIPTION INHIBITOR THIOLUTIN

A Dissertation

by

CHENXI QIU

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	Craig D. Kaplan
Committee Members,	Hays Rye
	Junjie Zhang
	Steve W. Lockless
Head of Department,	Gregory D. Reinhart

December 2017

Major Subject: Biochemistry

ABSTRACT

mRNA synthesis by RNA polymerase II (Pol II) is an essential process in eukaryotes. In my dissertation, I have undertaken two parallel approaches to expand our understanding of mechanism of Pol II, a large twelve-subunit protein complex in budding yeast *Saccharomyces cerevisiae*. First, we develop a high-throughput genetic platform to dissect functions of every residue in a critical Pol II active site domain: the trigger loop (TL). The TL multitasks in catalysis and translocation through its distinct conformational states, alteration of which causes wide-ranging transcription defects *in vitro* and *in vivo*. By establishing the correlation between a set of *in vivo* conditional growth phenotypes and *in vitro* biochemical defects, our genetic data allows us to predict biochemical defects and alteration of TL states in nearly all TL single substitution variants. For example, we provide evidence supporting critical contribution of an intra-TL hydrophobic pocket in stabilizing the off-catalytic TL state, as evidenced by mutations disrupting the pocket confer phenotypes consistent with increased catalysis and infidelity. These data are also consistent with a critical role of this intra-TL pocket in promoting Pol II fidelity. In addition, we show diverse allele-specific genetic interactions among TL and TL surrounding domains, supporting possible contribution of the TL surrounding funnel and bridge helices to TL dynamics and function. Second, we characterize the mode of action of thiolutin, a well-known transcription inhibitor with unclear mechanism of transcription inhibition. Recent studies demonstrated that thiolutin inhibited multiple metalloproteins through Zn^{2+} chelation, but failed to observe direct

thiolutin inhibition of purified RNA polymerases, suggesting additional factors are needed for thiolutin-mediated transcription inhibition or that the inhibition is indirect. We have taken chemical genetics and biochemical approaches to investigate the thiolutin mode of action. While characterizing multiple thiolutin effects *in vivo*, we demonstrate that thiolutin, when activated by DTT and Mn^{2+}, directly inhibits Pol II *in vitro*. We further investigate the nature of the inhibitory species and the property of the inhibited Pol II. We suggest that thiolutin inhibits Pol II through a novel mechanism distinct from most other known RNA polymerase inhibitors. Taken together, we develop a high-throughput phenotypic system to dissect functions of Pol II TL residues and characterize a novel mode of action of thiolutin.

DEDICATION

I dedicate this dissertation to my grandma, who passed away in 2013.

ACKNOWLEDGEMENTS

The past five years have been incredibly exciting to me. My personal and intellectual growth would never have been possible without the tremendous support from many people, some of whom are acknowledged here.

First and foremost, I would like to thank my mentor Dr. Craig Kaplan. He has been my role model, both in science and in life. His extremely high standard in experimental details and many aspects of life have truly inspired me and constantly reshaped my view of life. He has also shown me how a scientist can turn every single bit of life into an experiment, with controls carefully designed to test a diverse range of questions. For example, we have discussed approaches and how controls can be done to optimize the recipe for making chicken soup dumplings. He has also shown me how being a good friend and helping other people can be rewarding. He is a great mentor, a great scientist, a great friend and a great cook. I could not thank him enough for all I have learned from him.

I would like to thank my committee Dr. Hays Rye, Dr. Junjie Zhang, Dr. Steve Lockless for the invaluable and constructive insights over the years, and their great support during my postdoctoral job search. I want to particularly thank Dr. Steve Lockless, for many enlightening scientific discussions and his patience for many of my immature, crazy ideas. He also spent a lot of time helping me with technical details, including debugging

computer programs and some experimental trouble shooting. He has been incredibly supportive in my growth as a scientist, though not as a cook.

I would also like to thank many professors in our department for teaching me critical analysis and sharing valuable resources. I want to thank Dr. David Peterson, Dr. Mary Bryk and their lab members for providing many constructive suggestions on my research in our joint lab meeting. I also want to thank Dr. Dorothy Shippen, Dr. Geoffrey Kapler, Dr. Hays Rye, Dr. Steve Lockless, Dr. James Hu, Dr. Margaret Glasner, Dr. Sarah Bondos, Dr. Rodolfo Aramayo and Dr. James Sacchettini for invaluable insights in journal clubs. I want to thank Dr. Ryland Young for the extremely valuable critical analysis course and many scientific discussions throughout the years.

I would like to thank many people who have directly contributed to my projects. I would like to thank our collaborators Dr. Guillermo Calero (University of Pittsburgh) and Dr. Sing-Hoi Sze (Texas A&M university) for their work on structural and bioinformatics studies. I would also like to thank all the undergraduate students who have contributed to my work.

I would like to thank my past and current colleagues in Kaplan lab. I want to thank my past colleagues Dr. Huiyan Jin and Dr. Indranil Malik for early guidance in my graduate school. I want to thank my current colleagues Tingting Zhao, Yunye Zhu, Dr. Anandhakumar Jayamani and Evelyn Callaway for making my lab life a lot easier.

Among them, my time in Kaplan lab mostly overlaps with Dr. Indranil Malik, who has been great both as a friend and as a colleague. Our scientific and non-scientific discussions have been and will continuously be a big part of my life.

I would like to thank all of my friends for the support. I do not have enough space to have all of their names written here. I want to particularly thank my roommate in the past four years, Anthony Sperber, for our healthy scientific competition and his continuous support in every aspect of my daily life. I also want to thank some of my friends in College Station: Shuli Li, Qiuyan Shao, Charli Baker, Jimmy Trinh, Thomas Christian Snavely, Daniel Browne and Mehmet Tatli. My life would have been much more difficult without them.

I would like to thank my family for everything they have done to help me. My grandma was a brilliant and strongly motivated woman, who tried to help me in every aspect of my life, including science. She was incredibly patient listening to all of my work, even without any prior knowledge of biochemistry and genetics. My grandma passed away during my second year in graduate school after we failed to treat her Non-Hodgkin's Lymphoma. I want to thank her for her complete trust on me and her strong belief that the compound I was working on (thiolutin) could save her life, although we obviously did not poison her with thiolutin (We tried to save her life with Rituxan from Roche). I want to thank my parents for daily Wechat messages, weekly Facetime video chats and almost yearly visits from thousands of miles away. In the end, I would like to thank my

wonderful and beautiful wife, Yu Liu. She has been always with me every single time when I run into troubles in research and in life. We have been through so much that I could not express my gratitude in words. She is extremely supportive of what I enjoy doing instead of what makes us rich. Her dedication to our relationship and our marriage means a lot to me. I thank her, and I love her.

CONTRIBUTORS AND FUNDING SOURCES

My thesis work was supervised by a dissertation committee consisting of Dr. Craig D. Kaplan (chair), Dr. Hays Rye, Dr. Junjie Zhang from the Department of Biochemistry & Biophysics and Dr. Steve W. Lockless from the Department of Biology.

Olivia C. Erinne performed one round of the TL screen for deep sequencing. Ping Cui amplified the Sloning TL library and prepared TL amplicons for deep sequencing using non-emulsified regular PCRs. Jui M. Dave, Huiyan Jin, Nandhini Muthukrishnan and Leung K. Tang contributed to the manual traditional screen of the Sloning TL mutant library. Huiyan Jin also performed a subset of phenotyping experiments and the primer extension experiments in Chapter II. Paul J. Vandeventer constructed most of the BH-related mutant strains. Sabareesh Ganesh Babu and Kenny C. Lam together performed two rounds of TL screens for deep sequencing, and contributed to the a subset of phenotyping experiments in Chapter II. Ralf Strohner and Jan Van den Brulle contributed to the synthesis of the Sloning TL library. Dr. Sing-Hoi Sze wrote the codon-based alignment program and performed the alignment and fitness score calculation.

Dr. Craig D. Kaplan performed the Pol II bubble template transcription experiments, the first repeat of thiolutin treated Pol II transcription experiments with high DTT and also repeated other biochemical experiments in Chapter III. Dr. Craig D. Kaplan also performed the forward genetic screens with Emily Pavlovic in Chapter III. Dr. Indranil

Malik performed the *in vivo* thiolutin transcription shutoff experiments. Scott Ugo performed a pilot manual screen of the Variomics library. This screen led to isolation of several mutants (*TRR1**, *PRO3**, *YAP6** and *YRR1**). Amber LaPeruta constructed and phenotyped the first repeat of the *trr1*Δ and *snq2*Δ strains, and also contributed to the standardization of the UV-Vis experiments.

This work is funded by National Institutes of Health through grant R01GM097260 and by Welch Foundation through grant A-1763 to Dr. Craig D. Kaplan.

NOMENCLATURE

AAPs	Na-aroyl-N-aryl-phenylalaninamides
BH	Bridge Helix
BP	Biological processes
CC	Cellular components
Cor	Corallopyronin
E.coli	*Escherichia coli*
ePCR	Emulsion PCR
GO	Gene ontology
GOF	Gain-of-function
GSH	Glutathione
GSSG	Glutathione disulfide
LOF	Loss-of-function
Lpm	Lipiarmycin A3
MccJ25	Microcin J25
MF	Molecular Function
Mja	*Methanocaldococcus jannaschii*
MPA	Mycophenolic acid
MSA	Multiple sequence alignment
msRNAP	Multi-subunit RNA polymerase
Myx	Myxopyronin

NAC	Nucleotide addition cycle
NIR	Nucleotide interacting region
NTP	Nucleoside triphosphate
pEMAP	Point mutant epistatic miniarray profile
Pfu	*Pyrococcus furiosus*
Pol I	RNA polymerase I
Pol II	RNA polymerase II
Pol III	RNA polymerase III
PCD	Programmed cell death
Rif	Rifamycin
Rip	Ripostatin
RNAP	RNA polymerase
Sal	Salinamide A
Sce	*Saccharomyces cerevisiae*
Sor	Sorangicin
Spt	Suppressor of Ty
Stl	Streptolydigin
TAP	Tandem-affinity tag
Taq	*Thermus aquaticus*
TGT	Tagetitoxin
TL	Trigger Loop
TSS	Transcription start site

TABLE OF CONTENTS

Page

LIST OF FIGURES

Page

LIST OF TABLES

Page

CHAPTER I

INTRODUCTION AND LITERATURE REVIEW

Transcription is a conserved and essential process that converts genetic information from DNA into RNA. In prokaryotes, a single RNA polymerase synthesizes all cellular RNAs, whereas in eukaryotes distinct RNA polymerases synthesize classes of RNA: RNA polymerase I (Pol I) synthesizes most rRNAs, RNA polymerase II (Pol II) synthesizes nuclear mRNAs and RNA polymerase III (Pol III) synthesizes 5S rRNAs, tRNAs.

All RNA polymerases (RNAPs) transcribe RNAs in a conserved three-step process: initiation, elongation and termination. Each step has many common and distinct factors that control the RNA polymerase specificity and functional states. My main thesis work is divided into two parts: The first involves detailed functional dissection of the conserved Pol II active site and its communication with adjacent Pol II domains in *Saccharomyces cerevisiae (Sce)* (Chapter II). The second investigates the mode of action of thiolutin, a small molecule transcription inhibitor (Chapter III). Here, I will first review the highly conserved RNA polymerase substrate selection mechanisms, followed by discussion of evolutionary conservation and divergence within and surrounding the Pol II active site. Finally, I will review the modes of actions for diverse classes of transcription inhibitors, with the particular focus on the direct RNA polymerase inhibitors and their inhibitory hotspots.

Conserved mechanism of transcription elongation

Pol II elongation is a fast and accurate process

Given the essentiality of gene expression, there are minimally two stringent requirements for Pol II elongation. First, Pol II elongation has to be fast enough to satisfy the demands of cell growth and activity. Depending on different types of measurement in different species, the average elongation rate is in the range of 1-5 kb/min[1], with some early measurements on *Drosophila* Pol II on the lower end (1.1-1.5 kb/min)[2-5] and some recent measurements on human Pol II slightly faster (3.8-4.3 kb/min)[6-11]. Pol II transcription rate can also vary over the template, at least in metazoans[8-13]. Second, Pol II transcription elongation has to be accurate enough to faithfully convey the genetic information. In fact, transcription misincorporation rate *in vivo* used to be too low to quantify by RNA-seq[14-16]. Recent advances in the barcoding strategy for mRNAs and cDNAs dramatically expanded the error detection limit, allowing transcription errors to be distinguishable in yeast and *C. elegans*[15,16]. The C->U error is consistently higher than other error types and appears to be in the range of 10^{-5} to 10^{-4} in *E.coli*, yeast and *C. elegans*[14-16], while other types of errors appear to vary depending on species but mostly in the range of 10^{-6} or lower[15,16]. In addition to the minimal requirements to be fast and accurate, accumulating evidence has supported control of Pol II elongation in the regulation of gene expression[12,13,17]. Pol II transcription elongation is extensively regulated by various factors to coordinate with multiple co-transcriptional processes, as evidenced by the fact that altered Pol II catalytic rate could have wide-ranging consequences on various cellular processes *in vivo*[12,13,17].

Together, the efficient Pol II elongation balances the speed and fidelity, and coordinates with multiple other cellular processes[18,19]. The fine balance and coordination are contributed, at least in part, by a highly conserved and efficient substrate selection mechanism, mediated by two mobile and multifunctional active site domains: the Bridge Helix (BH) and the Trigger Loop (TL).

Multi-subunit RNA polymerases (msRNAPs) shared a conserved "crab-claw" like architecture, with active site BH and TL lying in the center and connecting various structural motifs (clamp, cleft, NTP channel, RNA exit channel *etc*.) together to define the basic functions for efficient transcription (**Figure1-1**, reviewed in [20] and references therein). The structure and function for the different motifs have been reviewed elsewhere[20-23], and some of them that are targeted by various inhibitors will be further reviewed in a later section. BH and TL are dynamic domains in the center of the msRNAPs (**Figure1-1D**), facilitating multiple critical steps during transcription elongation while retaining the ability to communicate with other domains and external factors. Below I will introduce basic mechanism of transcription elongation, along with the multiple conserved functions of the TL and BH.

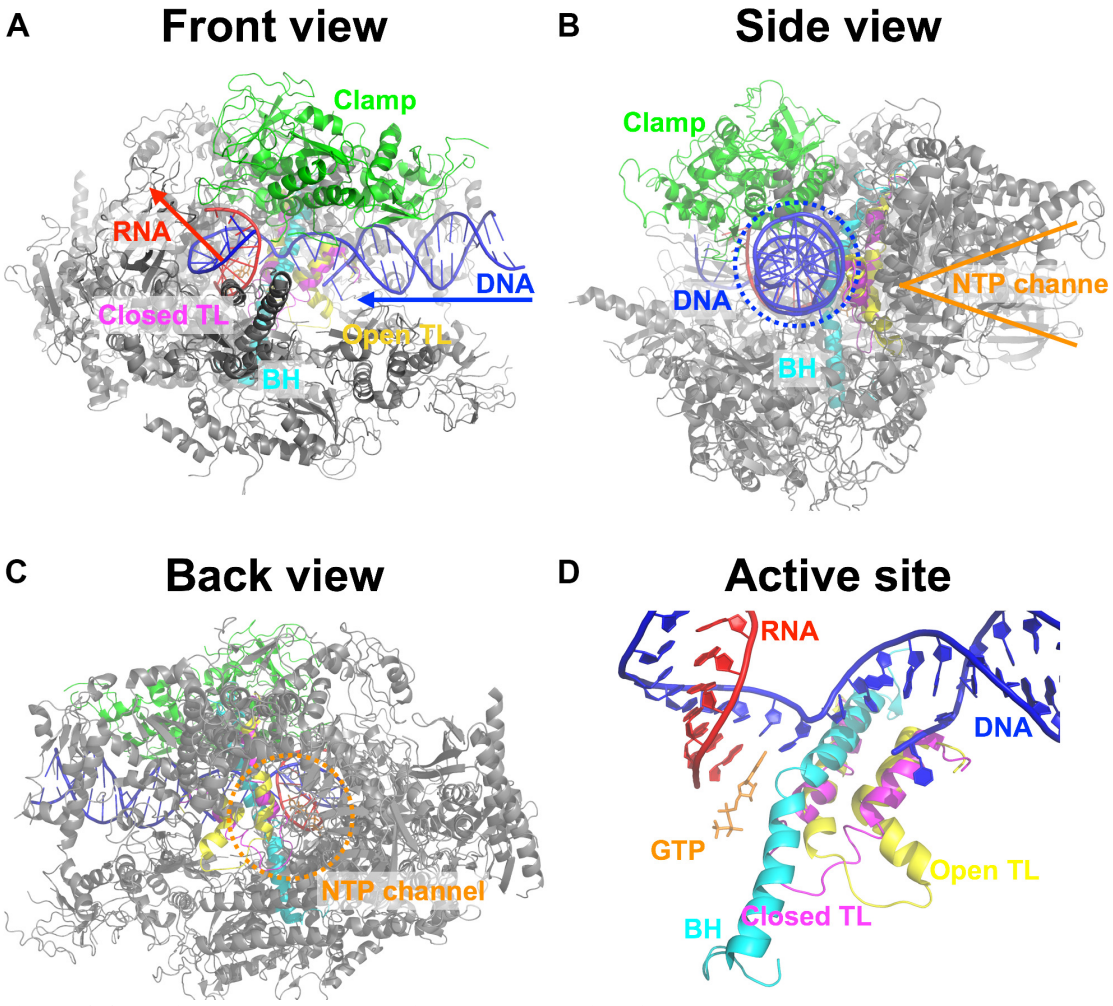


Figure 1-1. RNA polymerase II architecture and active site.
The overall crab-claw like architecture of RNA polymerase II elongation complex (PDB: 2E2H) with a matched substrate GTP bound in the active site. (A-C) Views from different angles to show the overall architecture, with the DNA (blue), RNA (red), GTP (orange) with mobile domains clamp (green), bridge helix (BH, cyan) and two different trigger loop states (TL, magenta for closed state and yellow for open state) colored. Red arrow indicates the direction for the growth of the RNA chain, blue arrow indicates the downstream to upstream direction of the template strand within the double stranded DNA. The open TL (PDB: 5C4J) is modeled to the structure by alignment of the Rpb2 domain.

The conserved trigger loop multitasks in nucleotide addition cycles

RNAPs, including Pol II, transcribe through iterative nucleotide addition cycles, with the TL playing multiple roles through its dynamic nature (**Figure1-2**). Within each cycle, Pol II selects the matched substrate nucleoside triphosphate (NTP) using a set of well-defined interactions among the TL, NTP, the nascent RNA, and template DNA (details below)[24-33]. Binding of the matched substrate appears to shift the TL conformational equilibrium from an open state to a closed state[24-26,34,35]. TL closing allows a substrate NTP to be captured in the active site. When fully closed, TL is hypothesized to promote catalysis of the phosphodiester bond between the NTP and 3' end of the RNA[24,27,35]. The reaction of phosphodiester bond formation releases the by-product pyrophosphate, which has been proposed to promote TL opening[34,36,37]. An open TL is required for Pol II translocation to the next position on the template[18,29,34,38-41], allowing the next round of nucleotide addition cycle. Changes of extensive TL interactions with itself or surrounding domains are involved and presumably promote the conformational shifts between the closed and open states[24,25,41,42]. Different TL conformations in between may also support off-catalytic pathway states such as pausing[35,38,43], backtracking[44,45] and intrinsic cleavage[46-50].

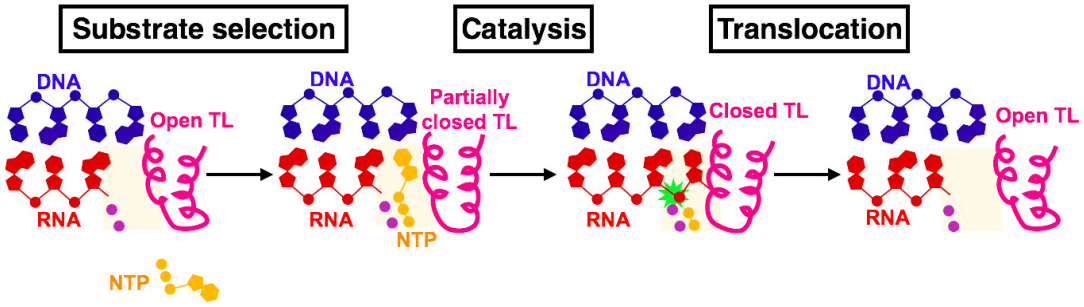


Figure 1-2. Multiple TL functions in NAC are supported by its mobile nature.
Three steps (substrate selection, catalysis, and translocation) in each nucleotide addition cycle (NAC) are labeled with different states of the TL indicated. Two purple circles in the active site represent the two Mg^{2+} ions involved in catalysis.

The TL's primary function in NAC is kinetic selection of the correct substrates. The TL mainly comprises three regions: an N-terminal nucleotide interacting region (NIR, *Sce* Rpb1 1078-1085), a loop (*Sce* Rpb1 1086-1090) and a C-terminal helical region (*Sce* Rpb1 1091-1106). Several residues in the TL NIR are critical in recognizing different components of NTPs such that only correct substrates are selected (**Figure1-3**). First, L1081 (*Sce* Rpb1 residue number is used unless otherwise specified) forms hydrophobic contacts with bases of matched substrates[24], suggesting a function in indirect recognition of substrates positioned by correct base-pairing. Second, a group of residues, including the TL residues Q1078, N1082 and the non-TL residue N479, appear to form an interconnected network of interactions, allowing recognition of 2', 3' OH and favoring incorporation of NTPs vs 2'dNTPs (or 3'dNTPs). This network is observed in structures and supported by genetic studies[24,27,31,45]. Third, H1085 interacts with ß-phosphate, and has been proposed to act as a general acid to protonate the ß-phosphate and to facilitate S_N2 attack on the 3' hydroxyl group of the RNA[24,51]. While this proposed mechanism was supported by molecular dynamics modeling studies[30], several lines of experimental data are inconsistent with this mechanism and suggested additional complexity[46,47,52,53]. In addition, the viability of non-ionizable substitution H1085Q suggests that the H1085 mediated acid-base catalysis is not essential[31]. To reconcile the experimental data, a very recent report proposes that H1085 may act as a positional catalyst instead of an acid-base catalyst[53]. In Chapter II, I will discuss the comprehensive evaluation of the mutational sensitivity of H1085 and discovery of an unexpectedly viable H1085L mutant, providing additional evidence to argue against the acid-base model of H1085 function. In short, the

TL binds the substrates through a set of well-defined interactions and appears to specifically promote the catalysis for correct substrates.

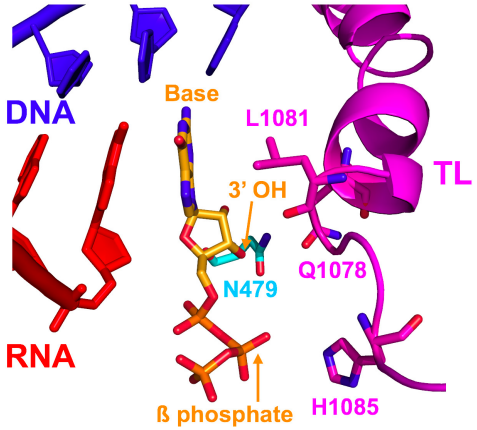


Figure 1-3. Pol II TL-NTP interactions
The details for closed TL (magenta) interaction with a bound GTP (orange) are shown in the cartoon view, with the critical residues and NTP shown in sticks. PDB: 2E2H.

Mutations that are proposed to alter TL dynamics cause specific defects in catalysis, fidelity and translocation[28,29,31,41,54], and can be roughly divided into two classes. First, mutations in the TL NIR broadly confer lethality in yeast, and the viable ones widely confer reduced elongation rates *in vitro*[31]. The second class of mutants confers increased elongation rate *in vitro*[28,29,31], and a prototype of this class, E1103G, also has compromised fidelity and translocation[18,28,29,54,55]. Different TL states are critical for distinct activities: the closed state is catalytically active and the open state allows translocation. E1103G appears to bias the TL dynamic balance towards the closed state[28,29], consistent with the fast elongation, compromised translocation and infidelity. Finally, the two mutant classes confer specific phenotypes that correlate well with decreased and increased activities *in vivo*[19,31] (will be reviewed in Chapter II), and we termed them as loss-of-function (LOF) or gain-of-function (GOF) mutants, respectively. The distinct phenotypes between classes allow us to develop a genetic assay to distinguish mutant classes at a high-throughput scale. In Chapter II, I will discuss the phenotypic profiles of nearly all TL single substituted mutants, and many more TL mutants that we propose alter TL function by shifting or altering TL dynamics. Together, the TL balances transcription speed and fidelity through its dynamics, disruption of which leads to highly distinct classes of mutants with specific *in vitro* and *in vivo* defects.

Context dependence and evolutionary divergence of the trigger loop

Despite the extremely high TL conservation, a number of observations suggest that the conserved TL among RNA polymerases can have distinct functions[56], likely due to

distinct environments surrounding the TL among different RNAPs. First, highly conserved TLs are not entirely interchangeable between yeast Pol I and Pol II[56]. Replacing the Pol II TL (Rpb1 1076-1106) with analogous sequences from Pol I causes lethality in yeast[56], although the nucleotide interacting residues between them are absolutely conserved. The mutant with a truncated Pol I TL swapped into Pol II (Rpb1 1076-1103) is viable but confers phenotypes consistent with LOF[56]. In contrast to Pol I, Pol III TL is more closely related to Pol II (6 mutations from Pol II) than is Pol I (11 mutations from Pol II), and appears to be compatible in the Pol II context[56]. In chapter II, I will describe a model, based on the different conformations of the TL surrounding funnel helices among Pol I, II and III, to reconcile this unique TL incompatibility between Pol I and Pol II.

Surprisingly, substitution of the hyper conserved E1103 residue with a glycine (E1103G) confers distinct and opposite effects in Pol I or in Pol II[56]. As discussed above, Pol II *rpb1* E1103G was extensively characterized by various assays and confers increased elongation rate with compromised translocation *in vitro*[18,28,29,57]. In sharp contrast, the analogous mutation to E1103G in Pol I (E1224G) confers reduced elongation rate both *in vitro* and *in vivo*[56]. E1103G was proposed to bias the TL dynamics towards the active, closed state in Pol II[28,56], presumably by disrupting the conserved interactions in the TL C-terminus required to stabilize the open TL state. Given the conservation of both the E1103 position and surrounding residues in Pol I and Pol II, it would be expected that E1224G also destabilizes the open TL in Pol I, though this has not been directly tested.

The E1103G induced fast elongation in Pol II is the sum of increased catalysis and apparent decreased translocation, suggesting that catalysis is the rate limiting step in Pol II elongation. If E1224G similarly alters catalysis and translocation in Pol I, a distinct outcome in the overall elongation rate would suggest a different rate-limiting step on translocation in Pol I[56]. Further experiments are in need to test this model. Together, distinct behavior of the same TL swapped into different RNAPs suggests critical roles of the TL surrounding environment in impacting TL function.

Sensitive control of TL dynamics and function by extensive residue-residue interactions within and surrounding the TL

Extensive observed residue-residue interactions within or surrounding the TL are critical for the proper TL function, as suggested by multiple observations. First, RNAP structures with different TL states reveal common and distinct residue-residue interactions, and simulation of the TL closing dynamic process further suggests changes of critical contacts in the intermediate states[24,25,42]. Second, some mutations in non-NIR TL residues alter TL function, and are hypothesized to bias the TL conformational cycle, likely through gain or loss of critical interactions[28,29,41]. Consistent with this hypothesis, several mutants in the TL surrounding domains confer similar *in vitro* increased elongation rates and *in vivo* phenotypes with the TL GOF mutants, suggesting that residues in the TL surrounding domain are also critical for maintaining a balanced TL[19,58-60]. Third, the TL function is dependent on context (discussed above), suggesting critical roles of environment for proper TL function. Together, this evidence collectively

suggests critical contributions from the residues within and without the TL for proper TL function, likely through extensive residue-residue interactions.

Contributions of residues outside of the TL suggest possible pathways of allosteric control of Pol II active site functions. Evidence from different species has shown that many RNAP domains (BH, F-Loop, RNA exit channel, funnel helix, clamp *etc*.) could either directly or allosterically impact the TL function[49,59-63]. For example, deletion of Rpb9, a small Pol II subunit distant from the active, confers *in vitro* and *in vivo* defects consistent with GOF[64-66], suggesting an Rpb1 helix, the funnel helix that interacts with both open TL and Rpb9, likely stabilizes the TL in open state[66]. This allosteric pathway from a distant subunit Rpb9 to the active site TL appears to play critical roles in maintaining the Pol II transcription fidelity[64-66]. In addition, Pol II interacting factors such as the elongation factor TFIIS, can directly insert into the Pol II active site, with the open TL state being required for this insertion and TFIIS dependent cleavage[28,29,67]. Together, given the sensitivity of TL to small changes and given the large number of interactions TL makes in different states, it is therefore critical to comprehensively evaluate many single-substituted TL mutants and the functional relationship among them. In Chapter II, I will describe a high-throughput phenotypic system to dissect almost all possible single-substituted TL mutants, and allele-specific genetic interactions among TL and several TL surrounding domains.

Diverse modes of actions for transcription inhibitors

Different classes of transcription inhibitors

Highly diverse transcription inhibitors exist in nature, as microbes compete against each other with natural products to target essential processes, such as transcription. Transcription inhibitors can be roughly classified into three major classes. The first class consists of direct RNA polymerase inhibitors. These inhibitors target small but essential pockets in RNA polymerases to inhibit its function, and may also be used to probe transcription mechanisms, as they may act on distinct conformations or steps in transcription. In RNA polymerases, there are four major pockets that have been shown to be targeted by direct inhibitors: the active site, RNA exit channel, switch regions and the NTP uptake channel. Inhibitors targeting these regions (here I term them inhibitory hotspots) will be separately discussed in a later section, with the particular focus on their modes of actions and the evidence supporting them.

The second class of inhibitors intercalates into DNA and blocks transcription[68-70]. This class of inhibitors often has multiple effects besides transcription inhibition, because they generally induce DNA damage and multiple stress response pathways. For example, actinomycin D arrests DNA replication, transcription, causes DNA damage, and induces apoptosis[68-73]. Therefore, transcription inhibition is generally only one of the multiple effects for this class of inhibitors and will not be further discussed in this review.

The third class inhibits transcription indirectly but specifically by targeting a transcriptional regulatory or accessory protein, which can include kinases that function in eukaryotic transcription, Pol II general transcription factors, or nucleotide biosynthetic enzymes that produce the NTP substrates for transcription[74-78]. Inhibition of transcription-associated kinases leads to alteration of Pol II functional states[79]. Inhibition of the general transcription factor TFIIH leads to transcription initiation defects[78]. Depletion of substrate NTPs by inhibiting related biosynthetic enzymes could in theory affect transcription elongation, but can also induce expression of specific genes to counteract with its elongation effect. For example, MPA was thought to induce global elongation defects, but recent work has shown that alteration in transcription elongation in yeast is not the major determinant of MPA sensitivity[31,80]. In yeast, MPA specifically induced expression of *IMD2* gene, which encodes for an enzyme substituting for the MPA inhibited *IMD3* and *IMD4*[81,82]. The ability to induce the *IMD2* expression through change of transcription start sites plays an essential role in conferring MPA resistance in yeast[82,83].

Together, all three classes of inhibitors perturb transcription in ways that can be of great value for the researchers and clinicians, and the common and distinct values for each class are discussed below (**Figure1-4**).

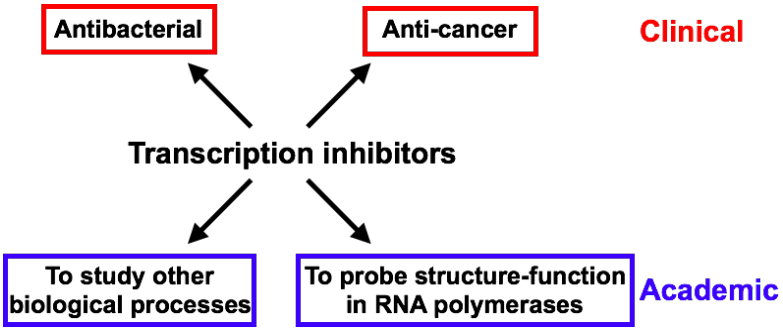


Figure 1-4. Values of studying transcription inhibitors

Values of studying transcription inhibitors

First, many transcription inhibitors have great potency as antibacterial antibiotics. The rationale of antibacterial treatment is straightforward. Despite the high conservation in the overall enzymatic architecture, bacterial RNA polymerase has regions that are highly conserved among bacteria but are distinct from eukaryotes, making them an ideal drug target for antibacterial treatment[84-86]. Given the essentiality of gene expression, inhibition of bacterial RNA polymerases can be extremely potent. In fact, bacterial RNA polymerase is a FDA proven drug target for various antimicrobial treatments (reviewed in [85] and references therein). For example, bacterial RNA polymerase inhibitor Rifampicin is the first-line drug for treating various bacterial infections, including but not limited to tuberculosis (reviewed in [85] and references therein).

Second, transcription inhibitors are emerging therapeutics for anti-cancer treatments. Cancer is a collection of diseases driven by uncontrolled cell proliferation, requiring high levels of ribosome and protein synthesis, which together are supported by all three cellular RNA polymerases. Many transcription inhibitors are found to specifically induce programmed cell death (PCD) in cancer cells, though further characterization reveals multiple distinct mechanisms are involved (**Figure 1-5**) (reviewed in [87]). Two examples are briefly reviewed below.

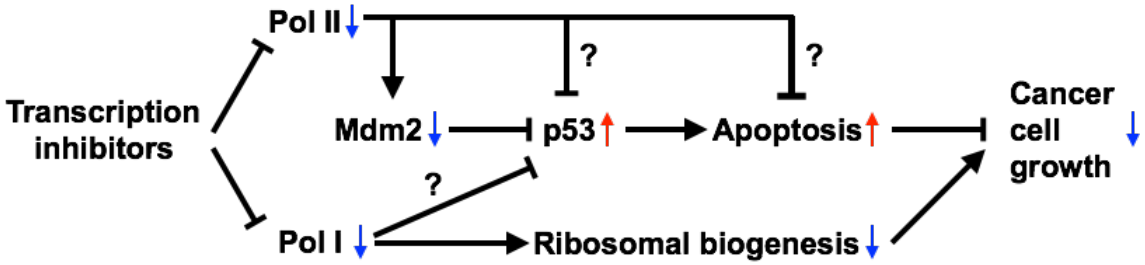


Figure 1-5. Multiple distinct mechanisms of cancer inhibition by transcription inhibitors. Transcription inhibitor inhibits cancer cell growth through a cascade of inhibition (blunt-end arrows with) and activation (arrows) signals. The increased or decreased activities after transcription inhibitor treatment are indicated with red and blue arrows, respectively.

Several Pol II global transcription inhibitors (DRB, roscovitine and α-amanitin) induce PCD in cancer cells through multiple p53 dependent and independent mechanisms (**Figure 1-5**) (reviewed in [87] and references therein). One example is through Mdm2-p53 pathway. Global transcription inhibition specifically down-regulates the short-life transcripts and proteins, including the E3 ligase Mdm2 that targets p53 for degradation[88-91]. Mdm2 down-regulation by Pol II inhibitors appears to be the cause for the subsequent p53 accumulation[92-95], though one report contradicts this hypothesis[96]. p53 generally functions as a transcription factor, and how p53 accumulation leads to PCD in the presence of transcription inhibition remains unclear (reviewed in [87]). It has been increasingly appreciated that p53 confers transcription independent functions in apoptosis (reviewed in [87] and references therein). For example, it was proposed that α-amanitin induces p53 accumulation and translocation into the mitochondria to induce apoptosis[97]. Finally, it has been also suggested that transcription inhibitors may induce PCD in p53 independent pathways, but the exact mechanism remains unclear (reviewed in [87] and references therein).

Selective Pol I inhibition preferentially induces PCD in many cancer cell types (reviewed in [87,98,99]). Increased Pol I transcription and ribosomal biogenesis have been well known as hallmarks of cancer for some time, but it was not clear until recently that specific Pol I inhibition can be explored as a potential therapeutic for treating cancer[98]. Pol I transcription inhibition, either by genetic inactivation or small molecule inhibition of necessary and specific Pol I initiation factors, was effective to inhibit the growth of

several cancer types[98-100]. Although the exact mechanism remains elusive, it appears that selective Pol I inhibition causes cancer inhibition through multiple pathways such as nucleolar disruption, cell cycle arrest and p53-dependent PCD[98-101]. In addition, several FDA approved anti-cancer drugs have been also found to function at least partially through inhibiting Pol I transcription[98,102-107], and many more selective Pol I transcription inhibitors are promising new drugs currently under development for cancer treatment (reviewed in [98,99]).

Third, many transcription inhibitors have been widely used as molecular tools to understand other biological processes in eukaryotes, including budding yeast *Saccharomyces cerevisiae*, pathogenic fungi *Candida albicans* and dinoflagellates[108-111]. For example, experiments monitoring downstream biological processes after perturbing transcription often reveal functional insights. A prominent example is the use of transcription inhibitors in mRNA stability studies. The stability of gene-specific RNA transcripts has been monitored after shutting down mRNA synthesis using various transcription inhibitors or the temperature sensitive Pol II allele *rpb1-1*[108-111]. In addition, genetic screens for mutants sensitive or resistant to transcription inhibitors often lead to discovery of functionally relevant processes. As an example, MPA and 6-AU, inhibitors that deplete cellular NTP levels, were utilized to screen for their sensitive mutants[112-116]. Many mutants involved in transcription, chromatin remodeling and other biological processes have been discovered from these types of screens[114,116]. These studies are not the main focus of my thesis and will not be discussed further.

Fourth, studies of transcription inhibitors have greatly contributed to the understanding of the basic mechanism of transcription. A paradigm for this is the use of rifampicin, an *E.coli* RNA polymerase inhibitor, to study transcription mechanism. Early on, rifampicin resistance was used as a phenotypic marker to genetically map the genes encoding RNA polymerase in *E.coli*[117-120]. Later, with the development of DNA sequencing, identification and characterization of the rifampicin resistant alleles contributed to our understanding of RNA polymerase structure and function prior to the full determination of RNA polymerase structure by X-ray crystallographic methods. For example, the mode of action for rifampicin, which inhibits the RNA extension after synthesis of the first phosphodiester bond through steric clashes, was proposed in 1978[121], almost 25 years prior to structural observation[122]. Subsequently, tens of rifampicin resistant mutants were sequenced and mapped to a tight pocket[123]. The evidence, together, revealed the intimate connection between the resistance pocket and the enzymatic active site even though crystal structures were unavailable, and also provide extensive quality validations for the later development of the structural studies.

With significant advances in crystallography, the binding sites of many transcription inhibitors have been explicitly mapped to small and essential pockets on various RNAPs. These structural observations, when coupled with biochemical studies of the inhibited RNAPs and genetic studies on the relevant regions, provide valuable connections between structure and function. For example, the observed rifampicin binding channel is functionally connected to RNA exiting[121,122]. In the following section, I will only focus

on the direct inhibitors that target RNA polymerases, and the structure-function relationship revealed from studying the modes of action for these inhibitors.

Hotspots for direct RNA polymerase inhibitors

The multi-step function and the partially conserved architecture of RNAPs make it a unique target for natural products. First, many RNAP functional steps can be hypothetically targeted by inhibitors. In fact, many steps, including promoter binding, open complex formation, early extension of first a few bases and elongation, are known targets of distinct inhibitors (discussed below). Second, while RNAPs have a conserved structural framework and active sites, there is structural diversity in the fine details, allowing for specificity/selectivity of natural product antibiotics (as discussed above). For example, the rifampicin binding pocket is highly conserved among bacterial RNAPs but not into the eukaryotic RNA polymerases, so rifampicin confers high potency against bacteria but low toxicity to humans. Due to clinical interest, many of the identified transcription inhibitors specifically target bacterial RNA polymerases, though inhibitors targeting eukaryotic polymerases (*e.g.* α-amanitin selectively inhibits eukaryotic Pol II) are also available.

There are a few inhibitory hotspots for the known RNAP inhibitors. They are the RNA exit channel, active site TL/BH, switch regions, and NTP uptake channels. The distribution for the inhibitors with available inhibitor/RNAP complex structures is shown on an *Tth* RNAP elongation complex structure (**Figure 1-6, Table 1-1**). The

tagetitoxin binding site is still under debate and falls into a unique category by itself[124-127] (**Figure 1-6**).

The inhibitor binding pockets are in general well correlated with their alteration of RNAP biochemical properties and highly informative on the function of the inhibited region. Here I will briefly discuss the structure, function and evolutionary conservation for each inhibitory hot-spot, followed by particular focuses on the known relevant inhibitors in the region. For each inhibitor, I will start with the current understanding of the mode of action. I will then discuss the structural and functional alterations induced by different inhibitors, and the mutations that confer resistance to inhibitors. There are two major rationales for discussing the resistant mutants in the context of the structural and functional alterations: (1) In the absence of co-crystal structures, resistance and cross-resistance with other inhibitors can be highly informative on the binding pocket of the inhibitors; (2) When co-crystal structures are present, the resistance not only tests the structural observation, but also reveals important insights into the functional relationship (allostery *etc*.) among the residues, because not all resistant mutants affect residues that directly interact with relevant inhibitors. Finally, I will discuss two inhibitors with unresolved modes of action, tagetitoxin and thiolutin. I will discuss the existing biochemical and genetic characterizations, along with complexities in studying them.

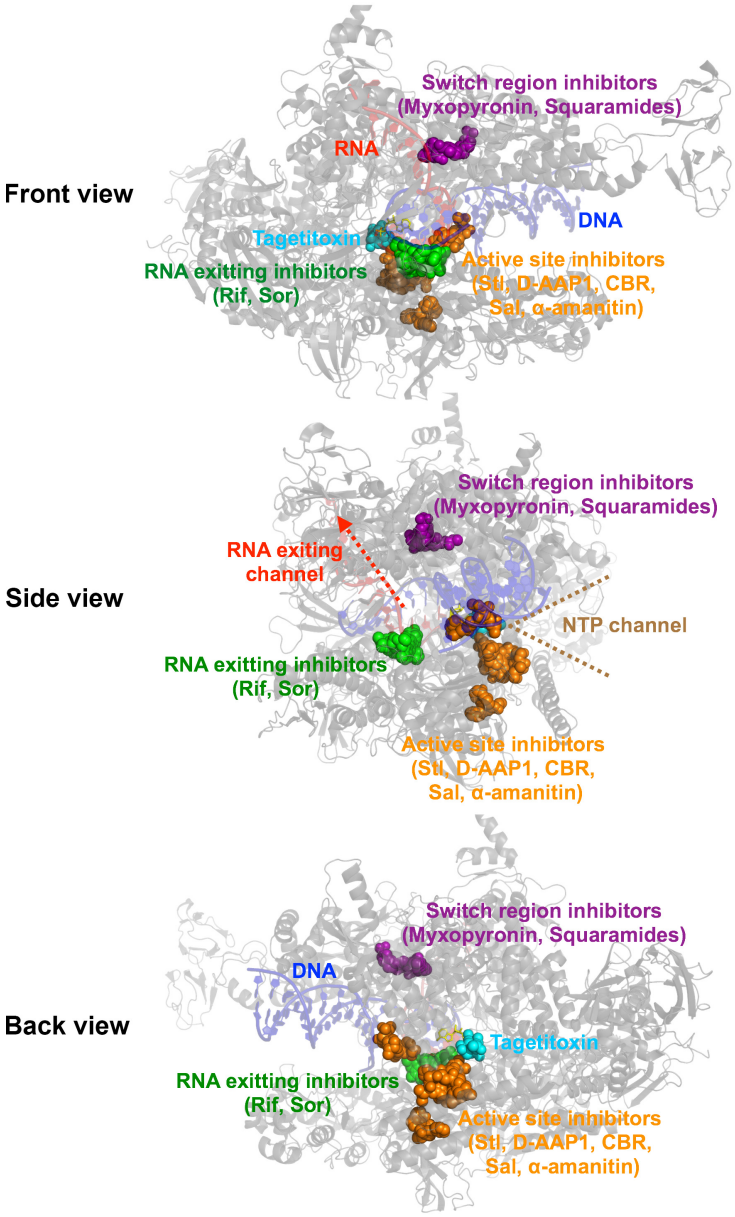


Figure 1-6. Distribution of the RNAP inhibitors' binding sites
Overall architecture of the transcription elongation complex for *Tth* RNAP (PDB: 2O5J) is shown in half transparent cartoon view (RNAP in grey, DNA in blue and RNA in red). Binding sites of different inhibitors are mapped to the elongation complex through simple homology alignment using ß subunit. α-amanitin comes from a *Sce* Pol II structure (PDB: 3CQZ) aligned to *Tth* RNAP using the Rpb2 subunit alignment to ß. Depending on both structural distribution and functional alteration, inhibitors can be classified into active site inhibitors (orange), RNA exit channel inhibitors (green), switch region inhibitors (purple) and unresolved tagetitoxin (cyan). Whether Tagetitoxin binds to the NTP uptake channel (shown, PDB: 2BE5) or the active site trigger loop (model unavailable, not shown) remains controversial and will be discussed further.

Table 1-1. PDB files used in this chapter

PDB	Summary	Reference
1I6V	*Taq* RNAP with rifampicin	Campbell et al, Cell, 2001
2O5J	*Tth* RNAP elongation complex. This structure has RNA that can be compared to the rifampin binding site.	Vassylyev et al, Nature, 2007
4KN4	*E.coli* RNAP with benzoxazinorifamycins	Molodtsov et al, J Med Chem, 2013
2A68	*Tth* RNAP with rifabutin	Artsimovitch Cell 2005
2A69	*Tth* RNAP with rifapentin	Artsimovitch Cell 2005
4KMU	*E.coli* RNAP with rifampin	Molodtsov et al, J Med Chem, 2013
1YNN	*Taq* RNAP with sorangicin	Campbell, EMBO J, 2005
1ZYR	*Tth* RNAP with Streptolydigin	Tuske et al, Cell, 2005
2A6H	*Tth* RNAP with Streptolydigin published on the same day with 1ZYR. The binding pocket is consistent with 1ZYR.	Temiakov Mol Cell 2005
4MEX	*E.coli* RNAP with Salinamide A	Degen et al, Elife, 2014
4XSX	*E.coli* RNAP with CBR 703	Bae et al, PNAS, 2015
4XSY	*E.coli* RNAP with CBR 9379	Bae et al, PNAS, 2015
4XSZ	*E.coli* RNAP with CBR 9393	Bae et al, PNAS, 2015
4ZH2	*E.coli* RNAP with CBR 703	Feng et al, Structure, 2015
4ZH3	*E.coli* RNAP with CBRH16-Br	Feng et al, Structure, 2015
4ZH4	*E.coli* RNAP with CBRP18	Feng et al, Structure, 2015
5UHE	*Mtb* RNAP with D-AAP1	Lin et al, Mol Cell, 2017
3CQZ	*Sce* Pol II bound with α-amanitin, refined from an early dataset 1K83	Kaplan et al, Mol Cell, 2008
2E2H	*Sce* Pol II elongation complex with a closed trigger loop. This structure is used to show the conservation of different inhibitor binding sites.	Wang et al, Cell, 2006
4MQ9	*Tth* RNAP with GE23077	Zhang et al, Elife, 2014
4OIN	*Tth* transcription initiation complex with GE23077	Zhang et al, Elife, 2014
3DXJ	*Tth* RNAP with Myxopyronin	Mukhopadhyay et al, Cell, 2008
3EQL	*Tth* RNAP with Myxopyronin	Belogurov et al, Nature, 2009
4YFK	*E.coli* RNA with Squaramide compound 8	Molodtsov et al, J Med Chem, 2015
4YFN	*E.coli* RNA with Squaramide compound 14	Molodtsov et al, J Med Chem, 2015
4YFX	*E.coli* RNA with Myxopyronin B	Molodtsov et al, J Med Chem, 2015

a) Active site inhibitors

Multi-subunit RNA polymerases (msRNAPs) have a highly conserved active site. As reviewed above, the TL/BH conformation cycling is critical for catalysis and translocation in nucleotide addition cycles. Therefore, interfering with the TL/BH motion is an extremely effective approach to inhibit catalysis and translocation. In addition, despite the high sequence conservation within TL and BH, evolutionary divergence in the TL/BH surrounding domains allows targeting of individual or a subset of RNAPs. The high essentiality and distinct evolutionary divergence makes the msRNAP active site a unique target for various natural products, as discussed below.

Streptolydigin

Streptolydigin (Stl) binds to a tight pocket near the RNAP active site, and locks the mobile BH and TL into inactive conformations[128,129]. Two Stl bound *Tth* RNAP crystal structures, published contemporaneously, showed an almost identical binding pocket and Stl orientation[128,129]. Stl binds to residues from BH, TL and two additional loops (named as "Stl pocket" by Tuske et al) from the ß subunit (**Figure 1-7**)[128,129]. Stl inhibits multiple bacterial RNAPs but not eukaryotic Pol I, II and III, likely due to the conservation of the two ß loops within bacteria but not in eukaryotes[130-132].

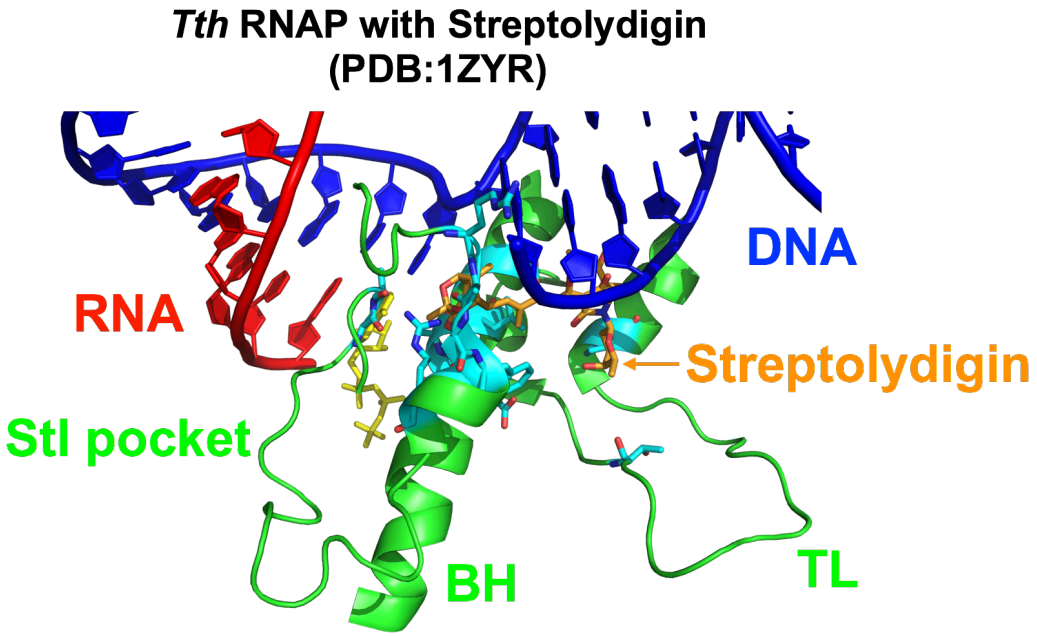


Figure 1-7. Streptolydigin binding site on the *Tth* RNAP.
DNA (blue), RNA (red), NTP (yellow) and the Stl surrounding domains (green) are shown in the cartoon view. Stl (orange) and the potential Stl interacting residues are shown in sticks (cyan).

As discussed above, biasing BH/TL to different states may have distinct effects, leading to either GOF or LOF in various RNAP activities[19,28,29,31,57,59,60]. The Stl-locked TL conformation is the catalytically inactive "open" state, and the Stl-locked BH is in a straight conformation. The Stl-bound BH was the first straight BH observed in bacterial RNAPs[128,129], which had previously been observed to be kinked in apo-RNAP structures lacking nucleic acids[129]. These differences suggested that Stl may stabilize the straight BH state in *Tth* RNAP. Together, the Stl locks RNAP active site in a particular state (straight BH, open TL), where TL is trapped in a state that does not favor catalysis.

Genetics and biochemical studies are consistent with the structural observations. First, most Stl resistant mutants universally overlap with the observed Stl binding pocket. In addition, two mutants (ß' A791G and S793P) do not appear to directly interact with Stl but confer high Stl resistance[129], and were proposed to bias the BH to the alternative kinked state that is not accessible to Stl[129]. Interestingly, the analogue to S793P in archaeal *Methanocaldococcus jannaschii (Mja)* RNAP (S824P) appears to confer increased catalytic or elongating activity[60]. In Chapter II, I will further discuss the analogous mutant in *Sce* Pol II (T834P), which indeed confers fast elongation *in vitro* and phenotypes consistent with GOF *in vivo* to yeast RNA polymerase.

Second, the biochemical defects of Stl inhibited RNAP are consistent with loss of TL function. First, Stl alters nucleotide addition, pyro-phosphorolysis, translocation, intrinsic and factor-dependent cleavage[128,132] but does not alter substrate NTP binding affinity[128]. In addition, Stl does not inhibit but induces the residual elongation activity of

TL deleted (*TL*Δ) RNAP[128]. Whether and how Stl stimulated the activity remains unclear, but these data also suggest the TL-dependence for Stl inhibition. Overall, the distribution of the Stl resistant mutants and the Stl induced inhibitory effects are consistent with the structural observation that Stl inhibits RNAP by locking BH/TL in a specific state that causes partial loss of TL function.

Interestingly, Stl appears to bias translocation dynamics by specifically stabilizing the post-translocation state[34,128,129], consistent with critical contribution of different BH/TL states in translocation. As reviewed above, different TL conformations appear to correlate with distinct translocation states. The observed Stl stabilization of post-translocation state is consistent with the Stl trapped open TL and straight BH conformations stabilizing post-translocation state. This observation is also consistent with multiple subsequent observations that biasing TL towards the catalytically active closed state inhibits translocation, likely by stabilizing pre-translocation states (reviewed above)[18,28,29,35,41]. In short, the Stl inhibitory effect is consistent with the model that open TL and straight BH may stabilize the post-translocation state, but further experiments are in need to test this model.

α-amanitin

α-amanitin is an eukaryotic-Pol II selective inhibitor and traps the active site BH and TL[28,133,134]. Similar to Stl, α-amanitin traps the TL in the inactive open state, and the BH appears to be locked in the straight state. However, in contrast to the Stl binding pocket,

α-amanitin binds to residues from TL, BH, the funnel helix and the Rpb2 link domain (**Figure 1-8**). Consistent with the structural data, most of the Pol II α-amanitin resistant mutants are mapped to residues near the observed α-amanitin binding site[28,133,135-137]. The α-amanitin binding pocket is universally conserved among eukaryotic Pol II but distinct from eukaryotic Pol I, Pol III and bacterial RNAPs, consistent with the selective inhibitory activity on Pol II at lower concentration.

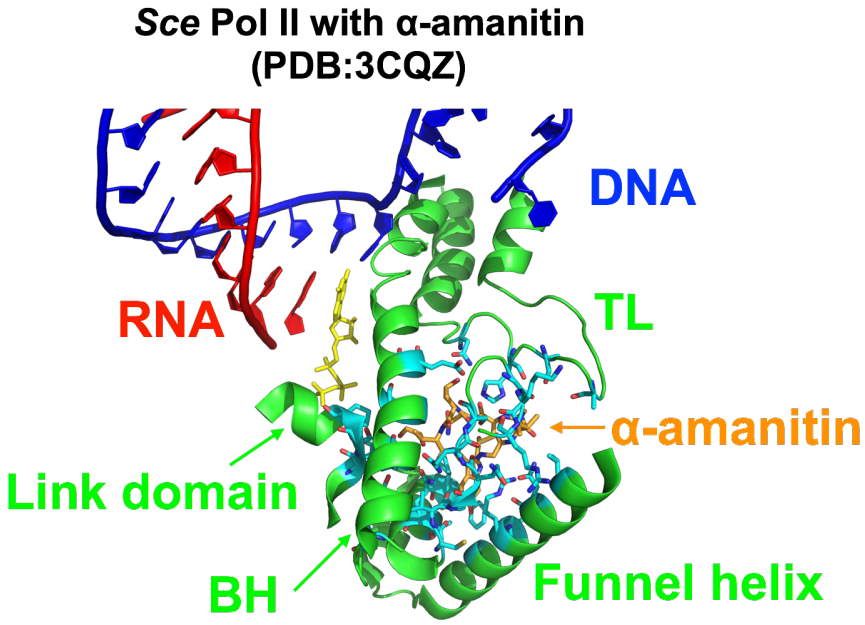


Figure 1-8. α-amanitin binding sites on the *Sce* Pol II.
DNA (blue), RNA (red), NTP (yellow) and the α-amanitin surrounding domains (green) are shown in the cartoon view. α-amanitin (orange) and the potential α-amanitin interacting residues are shown in sticks (cyan).

The inhibitory activities conferred by α-amanitin are consistent with a partial loss of TL function. First, α-amanitin inhibits multiple Pol II activities, including nucleotide addition, pyro-phosphorolysis, translocation, pause-release and TFIIS-mediated cleavage[138-140], but does not alter the substrate binding affinity[139]. Second, the α-amanitin inhibition is much stronger for the matched nucleotides than the unmatched ones[28], suggesting that α-amanitin targets the TL substrate selection mechanism[28]. Third, α-amanitin inhibits human Pol II translocation[141], but it has not been explicitly tested whether α-amanitin inhibits translocation through stabilization of pre- or post-translocated states, although one report suggested that the apparent post-translocated Pol II elongation complex might be resistant to α-amanitin[141]. Together, α-amanitin appears to trap Pol II BH/TL in a way similar to the Stl effects on bacterial RNAPs, but whether they trap BH/TL in the exact same state and whether they alter translocation similarly are still open questions.

Salinamides

Salinamides (Salinamide A and Salinamide B) are structurally similar bacterial RNAP inhibitors[142,143]. Their modes of action were not well characterized until very recently[143], possibly due to poor membrane permeability and low clinical interests. The prototype Salinamide A (Sal) binds to a tight pocket formed by the BH N-terminus, link domain and fork loop (**Figure 1-9**)[143], and inhibits transcription elongation through modulating the BH conformation[143]. The high conservation of the Sal binding pocket among bacteria

but not eukaryotes is in agreement with lack of Sal inhibitory activity on eukaryotic Pol I, II and III.

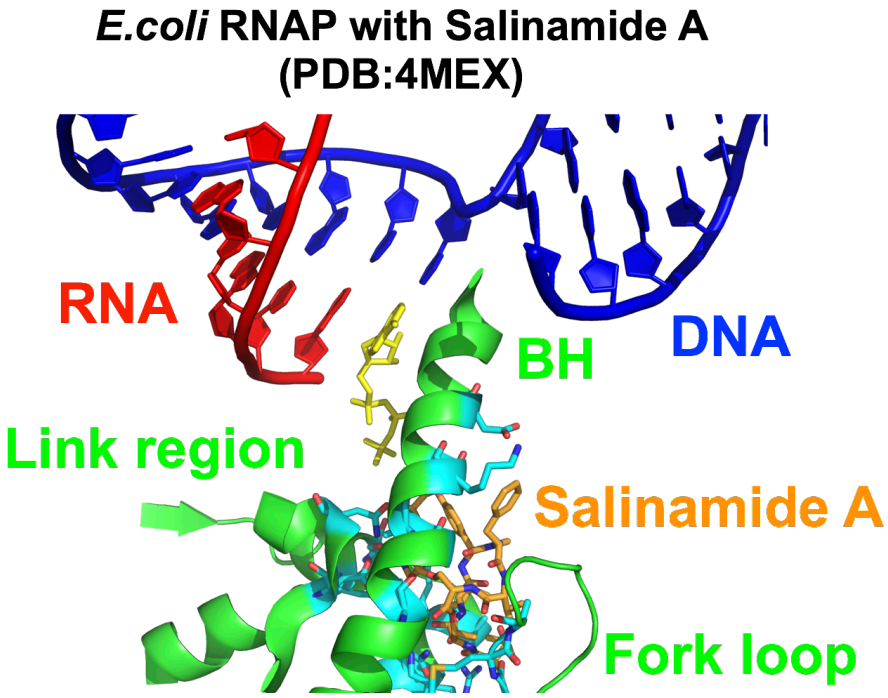


Figure 1-9. Salinamide A binding sites on the *E.coli* RNAP.
DNA (blue), RNA (red), NTP (yellow) and the Sal surrounding domains (green) are shown in the cartoon view. Sal (orange) and the potential Sal interacting residues are shown in sticks (cyan).

The mode of inhibition of Sal has both similarities and differences with inhibition by Stl and α-amanitin. Similar to Stl and α-amanitin, Sal inhibits both transcription initiation and elongation, specifically through altering catalysis but not substrate NTP binding[143]. In addition, Sal inhibits pyrophosphorolysis. These Sal inhibitory activities are consistent with interference with BH/TL conformation. However, in contrast to Stl and α-amanitin, Sal does not directly bind to the TL and does not require TL for its inhibition[143], as evidenced by similar Sal inhibition on both WT and *TL*Δ RNAP. It should be noted that Sal may still indirectly impact TL conformation through altering BH conformation, although TL is not essential for Sal inhibition.

Isolation of Sal resistant mutants from a drug efflux pump deficient strain leads to important insights into the Sal mode of action[143]. First, the isolated resistant mutants are tightly clustered in the Sal binding pocket, and all the mutants with high level resistance are involved in direct interaction with Sal. It is worth noting that none of the Sal resistant mutants were from the trigger loop, consistent with the TL independence of mode of Sal action, although we cannot rule out the possibility that the mutagenesis could be unsaturated. Second, Sal resistant mutants do not confer cross-resistance with many other known RNAP inhibitors, including the BH/TL trapping inhibitors Stl reviewed above[143]. In fact, some Sal resistant mutants not only do not confer resistance, but also confer hyper-sensitivity to Stl and another active site inhibitor CBR703 (reviewed below). A mutant conferring hyper-resistance to one inhibitor but hyper-sensitivity to another inhibitor generally indicates that the two inhibitors target two distinct but

functionally related steps or conformational states. This suggests that Sal binds to a distinct BH state from those bound by Stl and CBR703, although the difference is not entirely clear from the structures[128,129,143-145]. The connections between Sal and CBR703 will be further discussed below.

CBR compounds

CBR compounds are a series of synthetic bacterial RNAP inhibitors originally derived from a large synthetic compound library[146]. In 2015, two groups independently reported a series of *E.coli* RNAP structures bound with several CBR compounds, and unambiguously revealed an identical site for various CBR derivatives (**Figure 1-10**)[144,145]. The CBR compounds bind to a pocket that consists of residues from the ß region D, ß regions D/E spacer and the ß' BH N-terminus and the nearby F loop[144,145], and were proposed to allosterically inhibit the active site TL folding (**Figure 1-10**)[144,146,147]. The conservation of CBR binding pocket, especially the F loop and ß subunit residues are consistent with the selective activity of CBR compounds on a subset of RNAP[144,145].

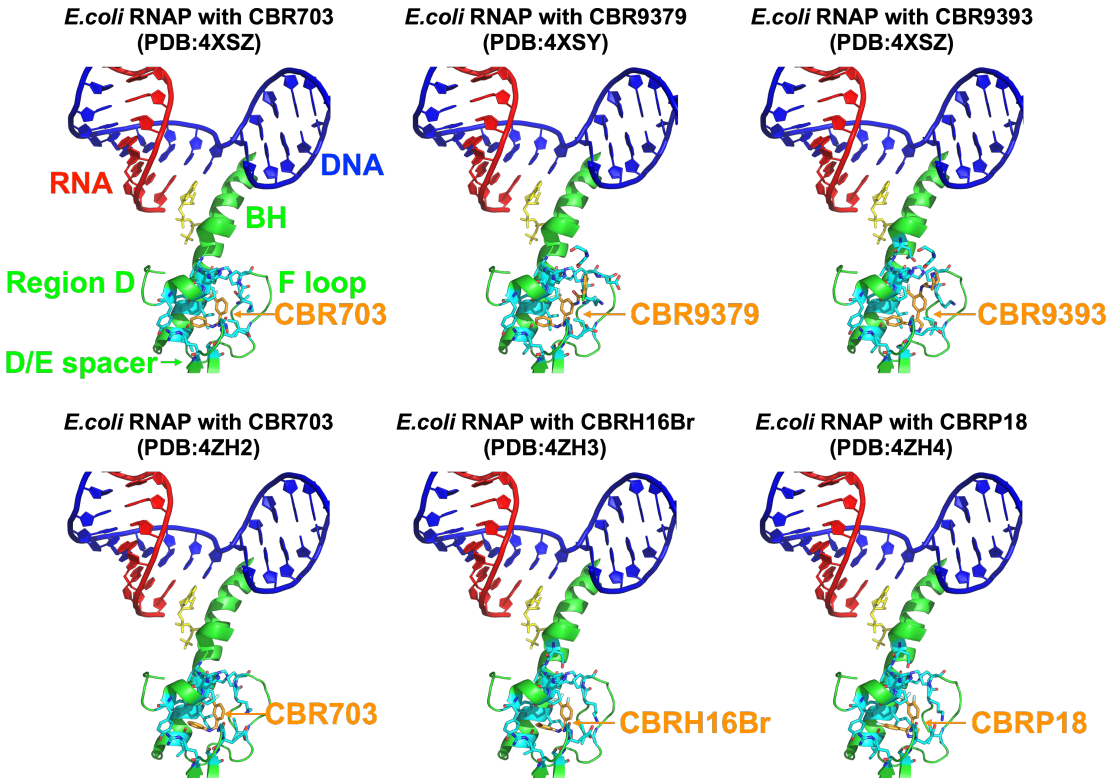


Figure 1-10. Various CBR compounds bind to the same site on the *E.coli* RNAP.
DNA (blue), RNA (red), NTP (yellow) and the CBR compound surrounding domains (green) are shown in the cartoon view. CBR compounds (orange) and the potential CBR compound-interacting residues (cyan) are shown in sticks. CBR compounds bind to the same site with slightly different interactions and orientations.

Genetic and biochemical studies not only validate the observed CBR binding sites, but also reveal important functional insights into the region. All isolated CBR resistant mutants[145,146] map to residues that either directly interact with CBR compounds or are near to residues that do[144,145]. The spontaneous resistance rate in *E.coli* for CBR703 is substantially lower than other known RNAP inhibitors, consistent with the CBR703 target region being highly essential[145]. Interestingly, two mutants (β′P750L and β′F773V) have severe growth defects that are suppressed by CBR703[146], which could be partially explained by CBR703 suppressing their translocation defects[147]. However, the exact reasons for their growth defect and dependence on CBR703 remain unresolved and fascinating. Finally, extensive functional studies of residues in this pocket suggest their conformational changes are allosterically coupled with the active site TL[49,59-63,148,149], and CBR indeed inhibits TL closing[144] while promoting forward translocation[147], a behavior reminiscent of Stl effect to maintain TL in open conformation. Together, these data are consistent with the hypothesis that CBR compounds allosterically inhibit the active site function from a distant site.

Interestingly, CBR inhibited multiple RNAP activities in two distinct modes: one is TL dependent, the other is TL independent[144]. Structural, genetic and biochemical data altogether suggest that CBR binding triggers a series of conformational changes propagating from the binding pocket to inhibit the active site TL folding, likely through the BH[144,147]. The inhibition of TL folding or alteration of other domains could lead to defects for specific functions. The nonessential nature of TL for *E.coli* RNAP allows

explicit tests of the TL dependence for any CBR inhibitory activities[144,145]. First, CBR inhibition of the nucleotide addition and pyrophosphorolysis are not fully TL dependent[144,145], as CBR compounds can still inhibit the *TL*Δ RNAPs. However, CBR inhibition of RNAP intrinsic cleavage is completely TL dependent[144], as *TL*Δ RNAP is fully resistant to CBR inhibition on the intrinsic cleavage activity. The close proximity of CBR and Sal binding sites and their TL-independent inhibition of specific RNAP activities suggests similar modes of actions between them, but further experiments are needed to rigorously test this hypothesis.

Na-aroyl-N-aryl-phenylalaninamides (AAPs)

AAPs are a series of synthetic compounds recently discovered from a large-scale screen for specific *Mycobacterium tuberculosis* (*Mtb)* RNAP inhibitors that do not inhibit other bacterial RNAPs and human Pol I, II and III[150]. The prototype D-AAP1 binds to a pocket near the BH N-terminus of *Mtb* RNAP (**Figure 1-11**). The D-AAP1 binding pocket is highly similar to the CBR pocket in *E.coli* RNAP, suggesting similar modes of action between them[144,145,150]. The sequence divergence in the pocket residues are consistent with lineage-specific inhibitory effects of both D-AAP1 (for *Mtb*) and CBR compounds (for Gram-negative bacteria). *Mtb* RNAP can be co-treated with both rifampin and D-AAP1, suggesting potential clinical values. However, further biochemical characterization is in need to validate the proposed mode of action of D-AAP1.

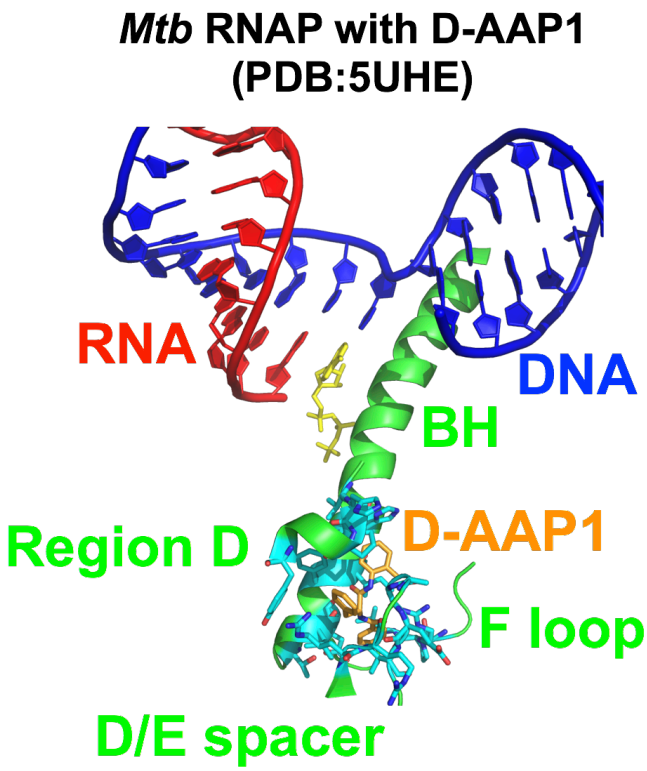


Figure 1-11. D-AAP1 binds to *Mtb* RNAP in a pocket similar to the CBR binding pocket in *E.coli* RNAP.
DNA (blue), RNA (red), NTP (yellow) and the D-AAP1 surrounding domains (green) are shown in the cartoon view. D-AAP1 (orange) and the potential D-AAP1 interacting residues (cyan) are shown in sticks.

b) RNA exit channel inhibitors

RNA has to properly exit the RNAP active site once it is synthesized. Structural studies reveal a primary and conserved RNA exit channel adjacent to the active site, releasing synthesized RNAs from a site opposite to the DNA entry groove. The RNA exit channel undergoes conformational changes and coordinates with various domains to maintain the upstream end of the transcription bubble while separating the DNA:RNA hybrid[63]. This dynamic process is critical for RNAP activity and is an inhibitory hotspot for natural products. The first-line anti-tuberculosis drug rifamycin family inhibits the RNA extension through causing steric clash with RNAs in the exit channel[121,122,150]. The modes of actions of RNA exit channel inhibitors (rifamycin and sorangicin) are discussed below.

Rifamycin (Rif)

Rifamycin is a family of compounds that are proven anti-bacterial drugs. Due to high clinical interests, rifamycin family has been extensively studied (reviewed in [85] and references therein). Many Rif compounds have been co-crystallized with various bacterial RNAPs, and showed an unambiguous binding site deep in the RNA exit channel just proximal to the active site (**Figure 1-12**). As discussed above, Rif compounds cause steric clashes with RNAs extending from the i-2 position, and thus specifically inhibiting extension of RNAs that are over 2-3 nts (**Figure 1-12**)[122,151]. The detailed interactions differ in various rifamycin derivatives, and the extent of the binding appears to correlate well with the potency (reviewed in [85,152]). Finally, it is worth noting

that a very recent report directly crystallized the Rif bound *Mtb* RNAP on a series of nucleic acids scaffolds, and directly demonstrated the steric clashes from a clinically relevant species[150].

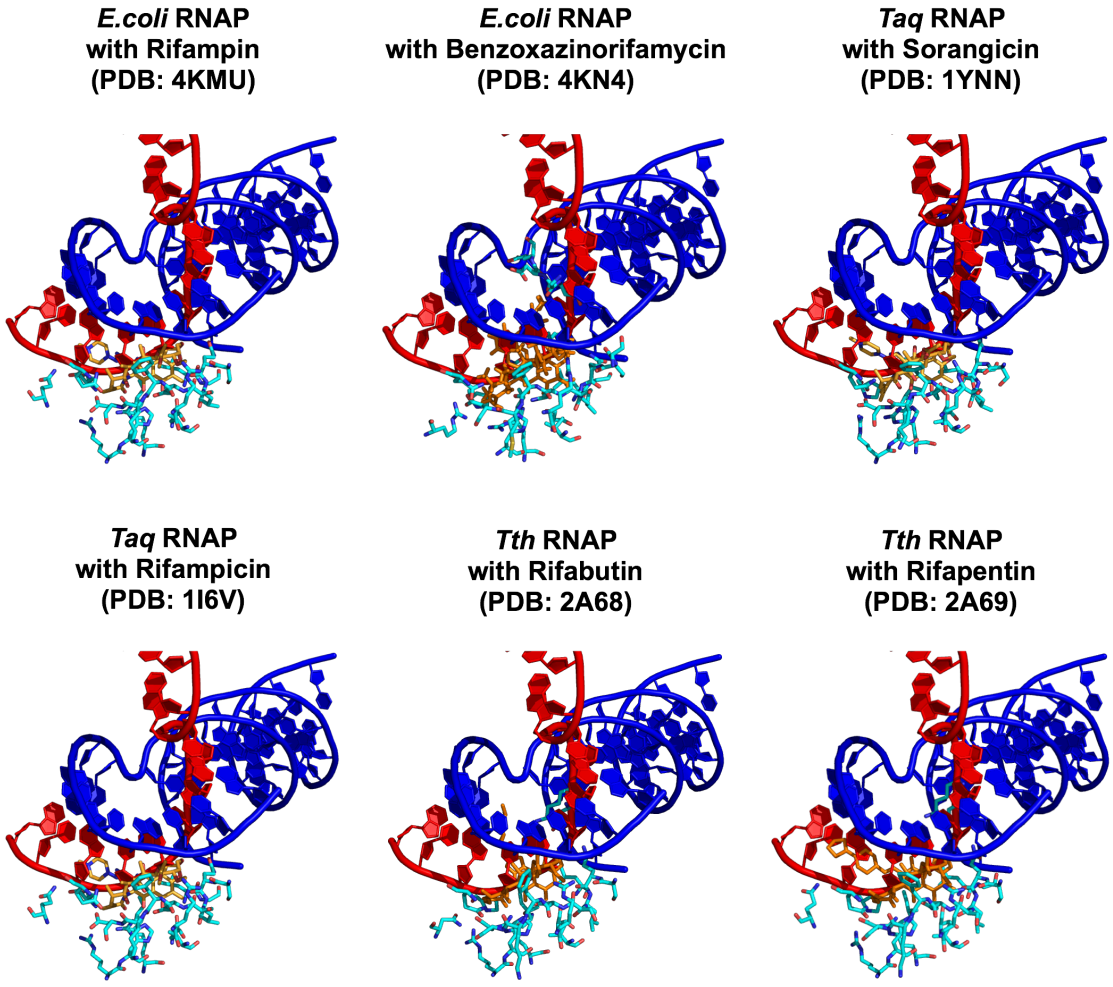


Figure 1-12. Rifamycin family compounds bind to various RNAP in the similar pocket in the RNA exit channel, causing steric clashes with the extending RNAs.
DNA (blue), RNA (red) are shown in the cartoon view. Rifamycin compounds (orange) and the potential rifamycin interacting RNAP residues (cyan) are shown in sticks.

Extensive evolutionary and genetic data support the structural observation and also reveal important insights. First, high conservation of the binding pocket among bacteria but not eukaryotes are consistent with the high Rif antibacterial potency with low cytotoxicity[122]. Second, most of the isolated Rif resistant mutants are tightly clustered near the Rif binding pocket (reviewed in [153] and reference therein). Many *Mtb* Rif resistant mutations confer fitness defects and were proposed to require a secondary suppressor mutation to outcompete the wild-type strain[154]. In fact, a study sequenced the whole genomes for a set of clinical or laboratory evolved Rif resistant strains, and identified many compensatory mutations from different RNAP subunits that do not directly interact with Rif[155]. Although the exact mechanism remains unclear, the distribution of these suppressors could reveal important structural and functional connections in RNAP, in addition to the clinical values discussed therein[155].

Rif compounds cause highly specific biochemical defects that are in agreement with the structural observations. Rif does not inhibit RNAP-promoter interaction, substrate interaction and formation of the first phosphodiester bond[121,156], and the elongating RNAP with already sufficiently long (>3 nts) RNA product is also resistant to Rif[157]. In contrast, Rif specifically inhibits formation of second or third phosphodiester bond in RNAs[121]. The Rif inhibitory effect is highly specific and distinct from the active site inhibitors, and is in complete agreement with the structural observation that Rif causes steric clashes with the RNAs that are more than 2 nucleotides long. An alternative model was proposed that Rif may allosterically impact the active site and cause decreased Mg^{2+}

affinity in the active site[124], but other experiments contradict these conclusions[158]. In short, the current model remains that Rif causes steric clashes with the extension of RNA at a specific length (2-3 nts), and the available biochemical data are consistent with the model.

Sorangicin (Sor)

Sorangicin binds to the same site (**Figure1-12**)[159] and causes similar biochemical defects with Rif[159-161]. Therefore, it is as expected that there is extensive cross-resistance between Sor and Rif compounds[159,161-164]. The resistance spectrum for Sor is significantly narrower than that for Rif (reviewed in [152] and references therein). In other words, all Sor resistant mutants are also resistant to Rif, whereas not all the Rif resistant mutants are resistant to Sor. The narrower spectrum for Sor was reconciled by the more flexible nature of the Sor molecule compared to Rif[152], thus providing the structural basis to adapt to the altered binding pocket in the resistant mutants. Ho et al also pointed out that similar cases have been observed in other compounds[165].

c) Clamp and switch region inhibitors

msRNAPs have a conserved and flexible clamp domain that can swing open or closed[166-170]. The clamp motion is intimately connected to various RNAP functions, such as initiation[166-170], elongation[35,63] and regulation by other factors[126,171-173]. In a single-molecule FRET assay looking at different RNAP clamp states, most *E.coli* RNAP clamps are open in the free solution, closed in an early step during initiation and are

proposed to remain closed most of the time during elongation[170]. In *E.coli* RNAP, transient clamp closure is critical for the initial promoter melting[169], but the melting extension into the downstream DNA to form a complete transcription bubble does not appear to require a particular clamp conformation[169]. In addition, the clamp has been hypothesized to remain closed during elongation to maintain high processivity[166,168,170]. However, several lines of evidence suggest that the clamp may retain some extent of flexibility in response to special circumstances, such as pausing[38] or allosteric communications between the RNA exit channel and the active site[63]. Clamp motion was also proposed to be linked to the active site BH/TL dynamics in *E.coli* RNAP[35], but this hypothesis remains to be further tested.

Five switch regions control the clamp motion. These switch regions (named switch 1-switch 5) are located at the base of the clamp and serve as hinges to control the clamp motion[166,168,170]. Although the five switch regions are universally conserved from bacteria to humans, the specific residues and surrounding environment can differ[153]. The critical function, high conservation in bacteria and the distinctions from the eukaryotes together make the RNAP switch regions another inhibitory hotspot. Several inhibitors have been discovered to target the switch regions so far (reviewed in [153]). Here I will review the binding sites, biochemical defects and the proposed mechanism linking the binding site to the specific functional defects.

Myxopyronin (Myx), Corallopyronin (Cor) and Ripostatin (Rip)

Myxopyronin (Myx), Corallopyronin (Cor) and Ripostatin (Rip) are all highly potent antibacterial transcription inhibitors sharing the same mode of action[174], through inhibition of RNAP clamp motion. Myx and Cor are structurally highly similar, but they are distinct from Rip (**Figure 1-13A**). Here I will first review the studies on Myx, the most extensively characterized among the three, while discussing their similarities.

Myx binds to a pocket near the switch regions 1 and 2 and appears to inhibit the clamp motion (**Figure 1-13B,C,D**)[170,174]. First, two *Tth* and one *E. coli* RNAP/Myx complex structures showed almost identical binding pockets though slightly different orientations[174-176]. Despite the overall high conservation in switch regions 1 and 2, three non-switch residues in the Myx binding pocket are not conserved in eukaryotic RNAPs, consistent with the lack of Myx activity on them[174,175]. In agreement with the observed Myx pocket, all the mutants with strong Myx resistance are from residues in close proximity to the Myx binding site[174,175]. Second, the Myx binding appears to lock the clamp in the partially or fully closed state. A single-molecule FRET system was developed to monitor different clamp states, and Myx appears to shift the majority of RNAPs to the partially or fully closed states from open states[170].

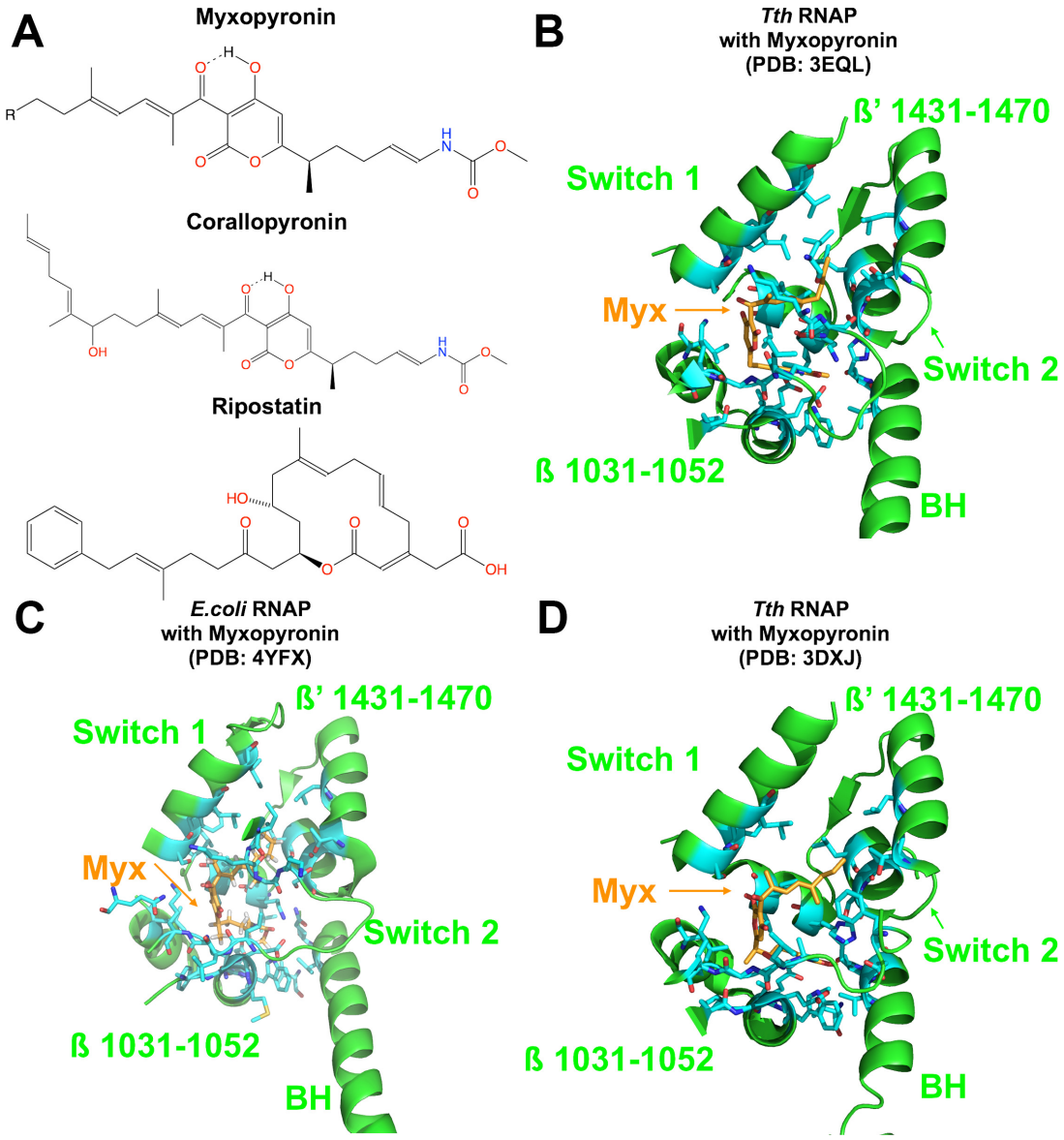


Figure 1-13. Myxopyronin binds to the similar sites in the *Tth* and *E.coli* RNAPs.
(A) Structures of Myx, Cor and Rip.
(B-D) The Myx surrounding domains (green) are shown in the cartoon view. Myx (orange) and the potential Myx interacting residues (cyan) are shown in sticks.

Myx inhibited *E.coli* RNAP shows highly specific biochemical defects that are distinct from those exhibited by other classes of inhibitors. First, Myx inhibits both abortive initiation and transcription of full-length transcripts, but the inhibition relies on the addition of Myx prior to RNAP/DNA binding[174]. In other words, template DNA binding protects the RNAP from Myx inhibition, suggesting that the Myx binding site is only accessible prior to RNAP binding to template DNA, or that Myx inhibits a functional step prior to DNA binding, or both. Second, Myx specifically inhibits RNAP binding to the double-stranded promoter DNA in the -11 to +15 region, but does not inhibit RNAP binding to the upstream region (-40 to -12)[174]. Promoter DNA from -11 to +15 is the region to be loaded to the RNAP cleft and melted to create transcription bubbles[174], and this process was recently shown to require apparent transient closure of the clamp[169]. In addition, Myx does not inhibit RNAP binding to the various single-strand DNAs, including non-template single-strand DNA, a fork-junction template, an artificial transcription bubble, or a gapped DNA templates[174]. Finally, RNAP transcription on the artificial transcription bubble, tailed templates and rolling-circle-transcription template bypasses the step of promoter opening and is not inhibited by Myx. Together, the Myx specifically inhibits promoter opening and formation of the transcription open complex, a critical step prior to transcription elongation.

Two lines of evidence further support the hypotheses that Myx inhibits RNAP through interfering clamp function. First, a Myx^R mutant in the switch region suppresses all the Myx induced biochemical defects listed above[174]. Second, several mutations in the

switch 2 region mimic the effect of Myx on promoter open complex formation[175]. These observations are consistent with a dependence of the switch region in Myx inhibition. RNAP structure with Cor or Rip is not available so far, but the almost complete overlapping cross-resistance and the similar biochemical effects suggest that Myx, Cor and Rip share the same target region and similar inhibitory mechanism[170,174].

Squaramides

Squaramides are a family of synthetic compounds originally identified in a high-throughput screening for transcription or translation inhibitors[177], and they were subsequently found to specifically inhibit *E.coli* RNAP *in vitro*[177]. Two squaramides (Squaramide 8 and Squaramide 14) are shown to bind to a pocket almost identical to the Myx binding site, although details differ (**Figure 1-14**). Several squaramide resistant mutants were isolated from *E.coli*, and all of the mutations were mapped to the switch regions in the pocket[177], consistent with the observed pocket being critical for a wide range of squaramides. The same pocket also suggests that squaramides may behave similarly to Myx, Cor and Rip and lock the clamp in the closed state, but biochemical characterization of squaramide inhibited RNAPs is needed to further test this idea.

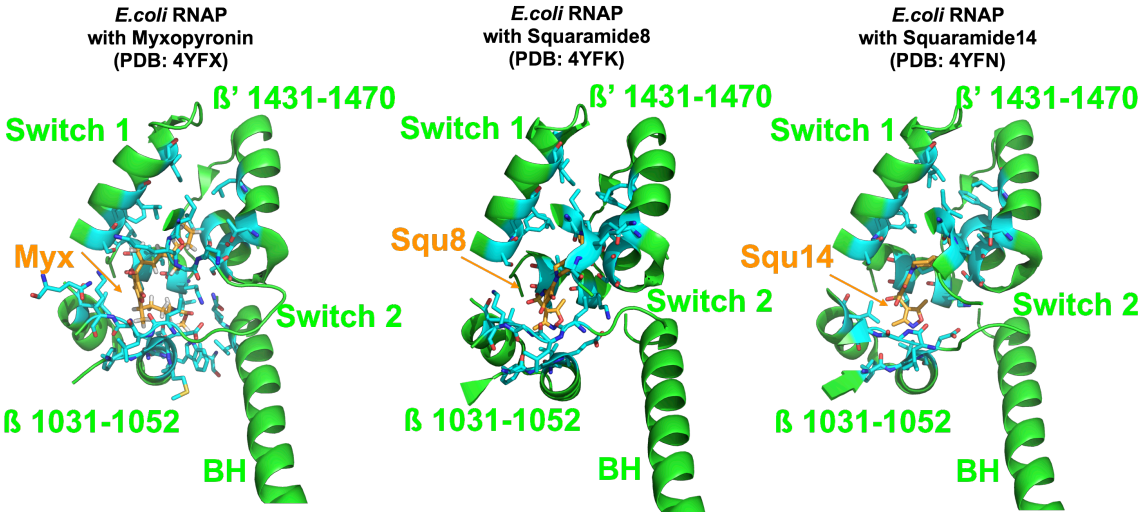


Figure 1-14. Squaramides bind to a similar pocket to the Myx binding sites in *E.coli* RNAPs.
The Myx and Squaramides surrounding domains (green) are shown in the cartoon view. Myx (orange), squaramides (orange) and the potential Myx or squaramides interacting residues (cyan) are shown in sticks.

Lipiarmycin A3 (Lpm)

Lipiarmycin is a mixture of antibiotics originally isolated from several *Actinomycete* species and has many variations on its naming[153,178-185]. The major component from two species[183,184], named as Lipiarmycin A3 or fidaxomicin, inhibits both Gram-positive bacterial and Gram-negative bacterial RNAPs and was hereafter referred to as Lpm[153]. One report[186] uses a Lpm derivative that has minor structural differences with Lpm, but the inhibition appears to be similar to Lpm. Lpm appears to bind to switch regions to inhibit clamp motion, in a way distinct from the closed clamp trapping inhibitors Myx, Cor and Rip. The mode of action of Lpm is discussed below.

Lpm appears to bind to a site distinct from the Myx, Cor and Rip pocket[153,169,187]. Although an RNAP co-crystal with Lpm is not yet available, the Lpm resistance spectrum, assessed from a broad range of bacterial species, reveals a tight region encompassing the RNA exit channel and switch regions 2 and 3[153]. The Lpm resistance region only partially and minimally overlaps with Myx, Cor and Rip, and no significant cross-resistance was observed among them[153].

The Lpm inhibited RNAP has both similar and distinct biochemical defects from those inhibited by Myx, Cor and Rip. Similar to Myx, Cor and Rip, the Lpm specifically inhibits promoter binding and opening[153], and the inhibition was only observed if RNAP was treated prior to template DNA binding[180,181,188]. The critical order-of-addition requirement is consistent with the classical behavior of the switch region inhibitors[153,187],

as discussed above. However, in contrast to Myx, Cor and Rip, Lpm appears to trap the clamp in a specific open or partially open conformation[187]. Lpm inhibited RNAP could not initiate promoter melting at the -10 position[169], an effect distinct from other inhibitors that traps the clamp closed (Myx, Cor and Rip). The different effect between Lpm and other closed clamp trapping inhibitors may represent distinct functional properties of RNAP with open or closed clamps. Together, the Lpm inhibits RNAP through limiting the clamp to a partially or fully open state and is highly distinct from all other known clamp inhibitors so far.

d) NTP uptake channel

msRNAPs have a conserved funnel-shaped NTP uptake channel (**Figure1-1**)[166,168]. Despite the conservation of the overall shape, residues in the NTP uptake channel differ from bacteria to human, thus making it a potential target to specifically inhibit bacterial RNAPs[84,86,189]. Micro-peptide Microcin J25 and natural product tagetitoxin are two best known inhibitors that have been proposed to inhibit the NTP uptake channel or the substrate entry/exit site. I will discuss Microcin J25 here below, but will discuss tagetitoxin later in the section on transcription inhibitors with unclear or complicated modes of action, since the exact binding site of tagetitoxin remains controversial[125-127,190].

Microcin J25 (MccJ25)

MccJ25 is a micro-peptide consists of 21 amino acids, forming an uncommon lassoed tail structure[191-194]. MccJ25 inhibits the growth of some gram-negative bacteria through RNAP inhibition[195,196]. Although a RNAP/MccJ25 complex structure is not yet available,

extensive genetics, biochemical and biophysical data are consistent with a model that MccJ25 binds to and obstructs the RNAP NTP uptake channel[197,198].

First, MccJ25 inhibits both abortive initiation and elongation, and the inhibition can be reversed by excess of NTPs, suggesting that MccJ25 is a partial competitive inhibitor with the NTPs[198]. In other words, the binding of MccJ25 and NTPs are mutually inhibitory but not mutually exclusive. Second, most of the isolated MccJ25 resistant mutants are clustered in the NTP uptake channel, and MccJ25 was shown to directly bind to the NTP channel in a FRET assay[198]. Third, MccJ25 inhibits the GreA and GreB dependent cleavage activities through interference with their binding, which is partially through the NTP uptake channel[198]. Fourth, MccJ25 inhibits processive pyrophosphorolysis and backtracking, both of which require the reaction products to release or occupy from the NTP channel[197,198]. Finally, the single molecular optical trapping experiments reveal that MccJ25 increases the pausing frequency but does not alter the pause-free elongation velocity[197]. Together, these data are consistent with the model that MccJ25 binds to and obstructs the NTP binding channel. MccJ25 also has additional effects on ROS induction *in vivo*[199-202], but these effects are unlikely to be related to transcriptional inhibition and will not be further discussed in this review.

e) Transcription inhibitors with unclear or complicated modes of action

Given the academic and clinical value of transcription inhibitor studies, discovering novel transcription inhibitors and characterizing identified compounds with unclear

modes of action are important goals. Many new transcription inhibitors are being discovered using different approaches, including but not limited to affinity based[203] and *in silico* screens[189]. Further characterization of these primary hits is necessary and in progress. In addition, several known transcription inhibitors are still being investigated and the modes of actions appear complicated[85]. Tagetitoxin (TGT) and thiolutin are two examples of these. Many studies reveal partially contradictory results on the two compounds, adding confusion about their possible modes of actions. Here I will briefly review the existing studies on both compounds, attempt to reconcile the results and propose possible experiments to further test their modes of action.

Tagetitoxin (TGT)

Tagetitoxin (TGT) has been known for decades to inhibit transcription, and has a highly selective inhibitory spectrum for a subset of RNAPs. TGT inhibits bacterial RNAP[204], plant chloroplast RNAP[205] and eukaryotic Pol III[206], but not eukaryotic Pol I, Pol II[206], or various viral RNAPs[207].

The exact binding site of TGT remains controversial[125-127,190]. In one report, a TGT bound *Tth* RNAP crystal structure revealed TGT binding to the NTP uptake channel (**Figure1-15**)[190]. Based on the structural and biochemical data, the authors proposed that TGT may act as an uncompetitive inhibitor to stabilize a particular inactive intermediate state during substrate loading[190]. However, this model does not explain why some long paused RNAP complexes are resistant to TGT[125,206]. To reconcile the results, Artsimovitch et al hypothesized that a particular TL state may be required for TGT

inhibition, and the open TL state in a paused complex may not be accessible to TGT, thus conferring resistance[125]. Artsimovitch et al further tested *in silico* whether a compatible TGT binding site exists on a *Tth* RNAP elongation complex with altered TL conformations, given that TL is a dynamic domain in elongation[125]. One TL adjacent TGT-compatible site was identified if the TL position was moved to a partially closed but slightly distorted state[125]. Resistant mutants have been isolated from both proposed binding pockets and residues distant from both pockets[125,190,208], suggesting that TGT inhibition may be sensitive to allosteric effects. Therefore, the exact TGT binding site remains unclear, although the recently revised TGT structure may be helpful for carefully revisiting the structural and modeling studies[209].

Despite the controversy about the TGT binding site, accumulating evidence has suggested TL dependence for TGT inhibition. First, TGT fails to inhibit multiple RNAP activities if the TL is absent[34,208]. Second, TGT appears to stabilize the closed or partially closed TL in a system where the TL can be monitored using engineered disulfide bond formation[35]. Third, TGT inhibits translocation in multiple assays[34,190,208], consistent with the behavior of a trapped TL. Together, the properties of the TGT-inhibited RNAP are consistent with the TL trapping model, but cannot rule out an allosteric effect imposed from the observed TGT binding site in the NTP channel[190]. It should be noted that the current data neither explicitly reveal the binding site of TGT nor rule out the possibility that TGT may bind to the free and elongating RNAP at different sites. More experiments are needed to further distinguish the two proposed binding sites.

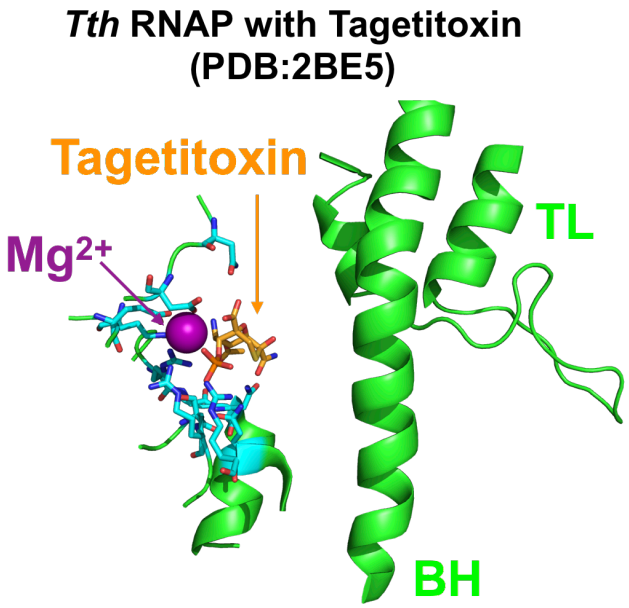


Figure 1-15. The proposed TGT binding pocket in the NTP uptake channel in *Tth* RNAPs.
The TGT surrounding and the active site BH, TL domains are shown in the green cartoon view. TGT (orange) and the potential TGT interacting residues (cyan) are shown in sticks. The proposed TGT trapped Mg^{2+} is shown in purple sphere view.

Thiolutin

Thiolutin is a well-known transcription inhibitor and belongs to the dithiolopyrrolone family, which consists of diverse compounds with a bicyclic structure and an intra-cyclic disulfide bond formed between two ene-thiols. Dithiolopyrrolones have very broad-spectrum inhibitory activities against many Gram-positive and Gram-negative bacteria, yeast and human cancer cell lines (reviewed in [210] and references therein). Thiolutin and holomycin are the two best-characterized dithiolopyrrolones regarding possible modes of actions, and they are structurally highly similar, with only one methyl group difference (**Figure1-16**). The mode of transcription inhibition by thiolutin and holomycin appears to be complicated and controversial, and recent evidence collectively suggests that multiple proteins are targeted by thiolutin and holomycin through reduction of the intra-molecular disulfide bond followed by Zn^{2+} chelation mechanisms (will be revisited later)[211-214]. Here I will first review the studies on the complicated mode of transcription inhibition, and I will discuss how Zn^{2+} chelation by thiolutin and holomycin is consistent with their multiple modes of actions but does not satisfactorily explain some early data on their transcription inhibition. Finally, I will attempt to reconcile the controversial results into a unified model. The history, chemical properties, biosynthesis and total synthesis for dithiolopyrrolones are comprehensively reviewed elsewhere[210] and will not be further discussed here.

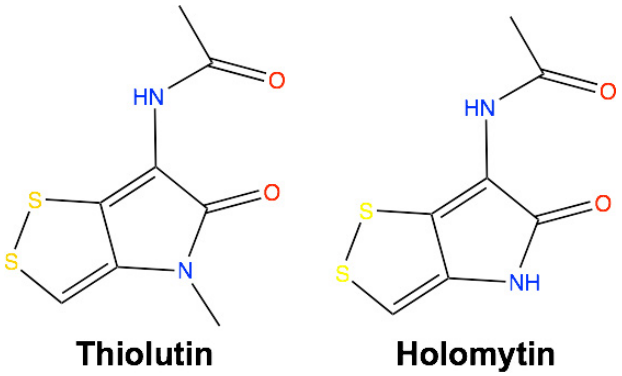


Figure 1-16. Thiolutin and holomycin structures

Thiolutin was originally revealed as a transcription inhibitor for both yeast and *E.coli* using substrate incorporation assays *in vivo* [215-217]. In brief, the activities for DNA replication, transcription, and translation were quantified by assaying incorporation of radioactively-labeled substrates (thymidine, uridine, and leucine, respectively) into trichloroacetic acid-precipitable macromolecular DNA, RNA and proteins[215,217]. It was found that thiolutin inhibited all three processes, but transcription inhibition appeared to be immediate and faster than DNA replication and translation inhibition, suggesting that transcription inhibition was the direct and primary effect *in vivo*. Similar results were observed for holomycin using similar assays in *E.coli*, although holomycin did not appear to arrest yeast growth in two independent strains[164] (also verified in our hands in Chapter III). Thiolutin has been commonly used to inhibit transcription to study mRNA stability in yeast, fungi, and in one report, a dinoglagellate[108-110,218,219]. While it appears that thiolutin can inhibit transcription in some eukaryotes and prokaryotes, and holomycin can inhibit transcription in prokaryotes, further *in vitro* characterization has revealed complexity in their possible modes of actions, as discussed below.

First, thiolutin was observed to directly inhibit RNA polymerases from yeast, but not other species *in vitro*, despite the conserved *in vivo* transcription inhibition from multiple assays and species[215,217]. Thiolutin inhibition of all three partially purified yeast RNA polymerases *in vitro* was distinct from that of the Pol II specific inhibitor α-amanitin[216]. Thiolutin inhibition appeared to be independent of specific DNA templates, as the inhibition was similar when the native salmon sperm DNA and synthetic poly(dA-dT)

were used. In addition, thiolutin was only inhibitory when RNA polymerases were pre-treated with thiolutin prior to template addition. In other words, DNA-bound RNA polymerases were resistant to thiolutin inhibition[216]. This critical order-of-addition is reminiscent of the switch region inhibitors Myx, Cor, Rip and Lpm behavior (discussed above). Two models were proposed based on the observed specific order-of-addition requirements. One, the RNA polymerases may be only accessible to thiolutin before template binding. The alternative model was that thiolutin may only inhibit the polymerase-DNA interaction, but not initiation or elongation of RNA synthesis. The two models are mutually non-exclusive and can now be further tested using fully purified RNA polymerases under experimental setups that specifically target each stage of transcription, as has been shown for other inhibitors (reviewed above). In my chapter III, I will discuss my work utilizing multiple biochemical assays to further characterize this mode of inhibition in detail.

Despite the observed inhibition in multiple *in vivo* assays[215,217], it was surprising that all the reports failed to observe thiolutin or holomycin inhibition of the *E.coli* RNA polymerase *in vitro*[164,213,220,221]. In addition, similar *in vivo* inhibition with lack of *in vitro* support was observed in *S. typhimurium*[222] and five distinct *Actinomycete* strains[221]. Together, this evidence collectively suggested that thiolutin may confer different modes of transcriptional inhibition in prokaryotes and eukaryotes. However, conclusions should not be made without comprehensive assessment of the experimental

parameters in different transcription assays. In chapter III, I will discuss my thesis work dissecting the necessary co-factors for thiolutin inhibition *in vitro*.

Second, it has been controversial whether thiolutin inhibits transcription initiation or elongation *in vivo*. The kinetics for inhibiting ß-galactosidase synthesis after IPTG induction is a classical experiment for predicting the targeted step of inhibitors. A transcription initiation inhibitor, when sufficiently added at a given time, presumably stops transcription initiation of all RNA polymerases but allows the elongating and completed transcripts to be further processed to functional ß-galactosidase (residual activity). In contrast, a transcription elongation inhibitor only allows the already completed transcripts to contribute to functional ß-galactosidases, thus resulting in different residual activities. In one report, Khachatourians et al. compared the thiolutin inhibition kinetics with rifampin (transcription initiation inhibitor) and chloramphenical (translation initiation inhibitor), and suggested that thiolutin targeted a step exclusively between transcription initiation and translation initiation, likely transcription elongation or termination[217]. In contrast, Sivasubramanian et al. did a similar experiment with rifampicin (transcription initiation inhibitor) and streptolydigin (transcription elongation inhibitor), and suggested that thiolutin appeared to inhibit transcription initiation[220].

It should be noted that these classical experiments, although elegantly designed, have limitations. The inhibitors with more than one effect may complicate the output. Although Khachatourians et al. did careful controls to show that thiolutin does not inhibit translation and ß-galactosidase *in vitro*[217], variables in other relevant biological

processes could not be ruled out. In fact, more recent reports have suggested that thiolutin inhibits multiple additional processes, including mRNA stability and proteasome degradation[214,223]. Alterations in these processes could complicate the readout in the classical experiments. Currently, we can directly monitor Pol II occupancy on genes *in vivo* using chromatin IP enabling assessment of initiation and elongation upon drug treatment[1,80,224] in yeast (Chapter III). An initiation inhibitor is predicted to specifically decrease the Pol II occupancy at the 5' end if elongating polymerases are allowed to clear the gene while initiation is blocked. In contrast, an elongation inhibitor should stop elongating Pol II across the gene, presumably eliciting Pol II degradation across the gene.

Two very recent reports nicely illustrate that reduced thiolutin and holomycin alter Zn^{2+} homeostasis by directly chelating Zn^{2+} (**Figure1-17**)[213,214], consistent with our results obtained from a different approach (Chapter III). Zn^{2+} caused specific redshifts for the reduced thiolutin/holomycin)[213,214], suggesting the reduced thiolutin and holomycin chelate Zn^{2+} through the reduced thiolate groups. In addition, a high-resolution MS/MS experiment was done in negative ion mode to characterize an apparent holomycin/Zn^{2+} complex[213]. The MS/MS result was consistent with Zn^{2+} binding to reduced holomycin in a 1:2 ratio, with extremely high affinity (~10pM range). In addition, the reduced holomycin can be re-oxidized when exposed to the air[212]. This behavior is consistent with redox-cycling compounds[225,226], which are known oxidative stressors. Thiolutin can be also reduced[214], but the redox cycling has not been directly tested. In chapter III, I

will describe our work in characterizing thiolutin induced apparent oxidative stress and functions of the yeast thioredoxin pathways in thiolutin resistance.

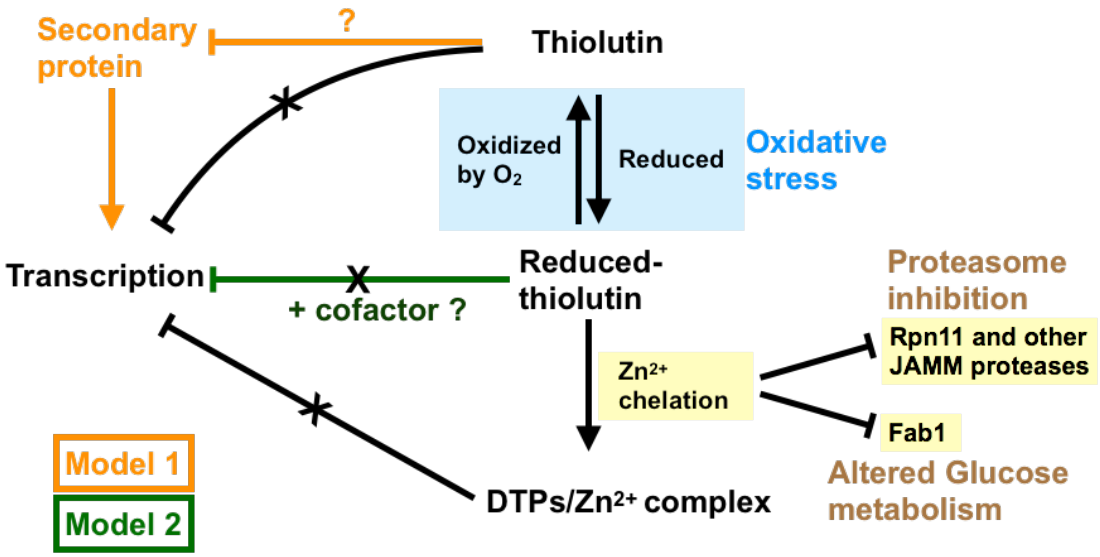


Figure 1-17. Two proposed models for thiolutin mediated transcription inhibition based on published data.

Tight Zn^{2+} binding suggested that thiolutin/holomycin could inhibit Zn^{2+} dependent enzymes through Zn^{2+} chelation. The inhibition of several Zn^{2+} dependent enzymes was directly validated *in vitro*. For example, the zinc-dependent fructose bisphosphate aldolase FbaA was inhibited by reduced holomycin at ~25μM, a relevant concentration for growth inhibition[213], and FbaA inhibition is consistent with an early report that thiolutin altered glucose metabolism[227]. In addition, thiolutin inhibited proteasome activity through chelating Zn^{2+} from Rpn11, a deubiquitinating enzyme with a JAB1/MPN/Mov34 domain, and an essential subunit of the 19S proteasome[214]. Lauinger et al. further showed that reduced thiolutin also inhibits another JAB1/MPN/Mov34 domain–containing metalloprotease. Interestingly, both groups failed to observe inhibition of RNA polymerases through Zn^{2+} chelation, which they interpreted as evidence for transcription inhibition being a secondary consequence of dithiolopyrrolones. The contradictory results from *in vitro* and *in vivo* transcription assays in *E.coli* and accumulating evidence on other effects have now cast doubt on initial proposals that thiolutin primarily and directly inhibits transcription.

I propose two models to reconcile the apparent universal observation that thiolutin inhibits transcription *in vivo* but not *in vitro*, except for one early report[216]. First, thiolutin may inhibit a secondary protein, which is present *in vivo* and in the original partially purified RNA polymerase fractions[216] but not others[164,213,214,220,221]. Second, thiolutin may act as a prodrug that requires activation *in vivo*. Recent work has advanced our understanding of the structure-activity relationship in dithiolopyrrolones and

revealed the possibility of reductant and metal being the co-factors[213,214]. Although pre-reduced thiolutin failed to inhibit transcription, the requirement for additional co-factors should not be ruled out. In Chapter III, I will describe a multi-pronged approach to investigate the mode of action of thiolutin, leading to our discovery that both appropriate doses of reductant and Mn^{2+} are necessary co-factors for direct thiolutin-mediated transcription inhibition.

CHAPTER II

HIGH-RESOLUTION PHENOTYPIC LANDSCAPE OF THE RNA POLYMERASE II TRIGGER LOOP*

Overview

The active sites of multisubunit RNA polymerases have a "trigger loop" (TL) that multitasks in substrate selection, catalysis, and translocation. To dissect the *Saccharomyces cerevisiae* RNA polymerase II TL at individual-residue resolution, we quantitatively phenotyped nearly all TL single variants *en masse*. Three mutant classes, revealed by phenotypes linked to transcription defects or various stresses, have distinct distributions among TL residues. We find that mutations disrupting an intra-TL hydrophobic pocket, proposed to provide a mechanism for substrate-triggered TL folding through destabilization of a catalytically inactive TL state, confer phenotypes consistent with pocket disruption and increased catalysis. Furthermore, allele-specific genetic interactions among TL and TL-proximal domain residues support the contribution of the funnel and bridge helices (BH) to TL dynamics. Our structural genetics approach incorporates structural and phenotypic data for high-resolution dissection of transcription mechanisms and their evolution, and is readily applicable to other essential yeast proteins.

Introduction

RNA polymerase II (Pol II) synthesizes all eukaryotic mRNA. Structural studies of *Saccharomyces cerevisiae* (*Sce*) Pol II have illuminated mechanisms of transcription[20,21,24,166,167,228], especially RNA synthesis. RNA synthesis occurs through iterative nucleotide addition cycles (NACs): selection of correct substrate nucleoside triphosphate (NTP), catalysis of phosphodiester bond formation, and enzyme translocation to the next template position. These critical steps in NAC appear to be coordinated by a critical, conserved domain within the Pol II active site: the trigger loop (TL).

TL functions are underpinned by its mobile and flexible nature (**Figure 2-1A**). The primary function of the TL is kinetic selection of correct NTP substrates while balancing transcription speed and fidelity, and this function is highly conserved based on studies of RNAPs from *Escherichia coli* (*Eco*)[41,46], *Thermus aquaticus* (*Taq*)[229], the archaeons *Pyrococcus furiosus* (*Pfu*)[230] and *Methanocaldococcus jannaschii* (*Mja*)[59], and eukaryotic Pol II from *Sce*[28,29] and human[231]. In a simplified two-step model, correct NTP binding appears to facilitate TL movement such that a bound, matched NTP shifts the TL from the "open" state to the "closed" state[24-26,34,35], allowing capture of the matched NTP in the Pol II active site and promotion of phosphodiester bond formation[24,27,35]. The subsequent release of the byproduct, pyrophosphate, allows a conformational shift of the TL from the "closed" state back to the "open" state[34,36,37]. TL opening has been proposed to be critical for enzyme translocation relative to the DNA

template, an essential step for the next nucleotide addition cycle[18,29,34,38-41]. Furthermore, additional TL states have been implicated in transcriptional pausing from studies in *E.coli*[35,38,43], backtracking from structural observations[44,45], and, although controversial, intrinsic cleavage[46-50]. Thus, distinct TL conformations or interactions are linked to different functions in transcription, with delicate control of TL dynamics promoting proper transcription elongation while possibly incorporating signals from the rest of Pol II or Pol II bound factors[35,63,67,232,233].

Genetic and biochemical studies have revealed TL functions in the NAC. First, the nucleotide interacting region (NIR, Rpb1 1078-1085) discriminates matched rNTPs from 2'-dNTPs and non-complementary rNTPs[28,29]. NIR substitutions in residues observed to interact with rNTPs widely conferred lethality. Where viable, substitutions reduced catalytic activity *in vitro* and were termed as partially loss-of-function (LOF)[28,31,41,46,229,230]. Second, a TL C-terminal mutant E1103G, conferred increased catalytic activity *in vitro*, which we termed gain-of-function (GOF)[28,29,57]. Fast kinetics experiments revealed that E1103G may bias TL dynamics towards the catalytically active "closed" state[29], consistent with infidelity and compromised translocation in addition to increased catalysis[18,28,29,54,55]. Furthermore, we previously described a set of Pol II TL mutants with broad and distinct alterations to transcription *in vivo*, thus conferring allele-specific phenotypes (**Table 2-1**) that correlate with decreased or increased activity[19,31] *in vitro.* Various genetic interactions (suppression, exacerbation, and epistasis) have also been observed among TL substitutions, suggesting a complex

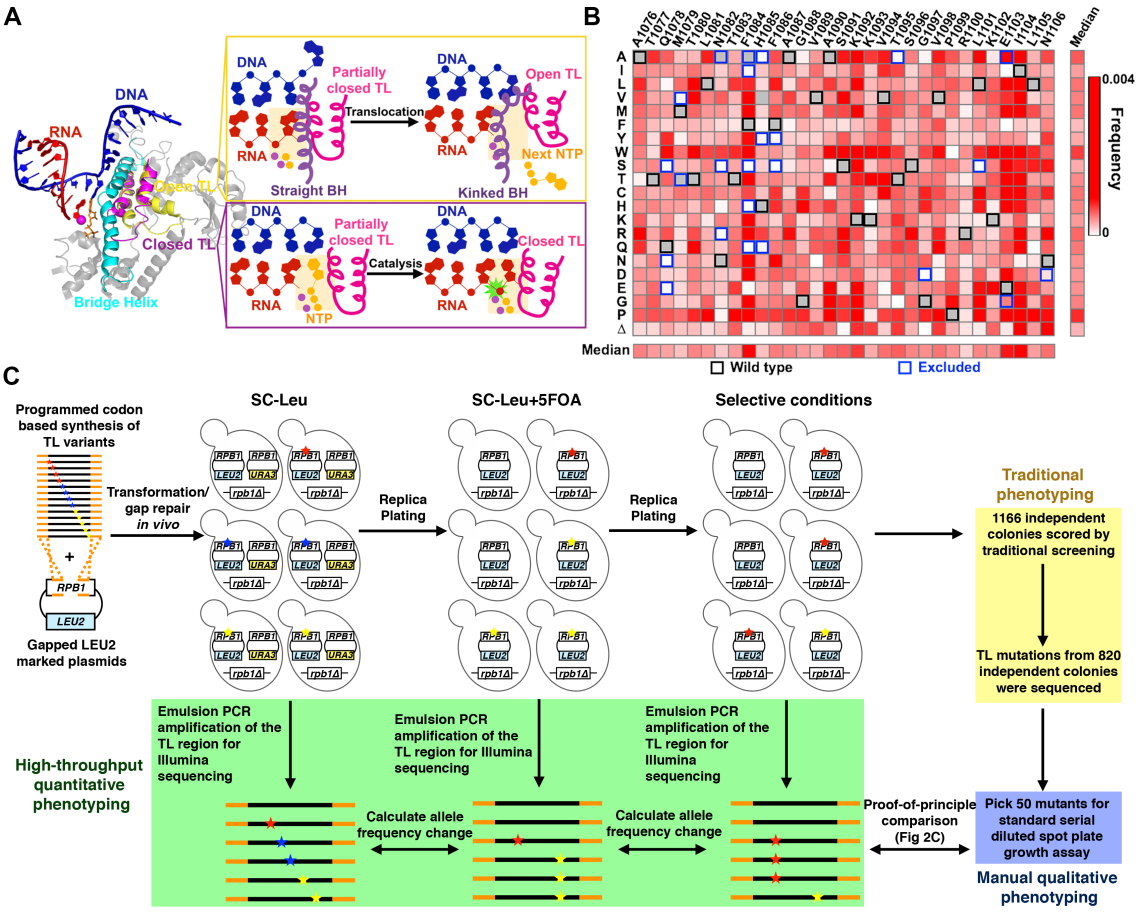


Figure 2-1. Establishment of a high-throughput platform for phenotyping comprehensive TL single variant library.

(A) Multiple TL functions are underpinned by its mobile nature. Structures of open (PDB:5C4J) and closed TLs (PDB:2E2H) are shown in the context of surrounding domains. Template DNA (blue), RNA (red), Bridge Helix (cyan), Closed TL (magenta) and Open TL (yellow) are shown in cartoon rendering. The open TL has been proposed to allow Pol II translocation while the closed TL has been shown to facilitate catalysis (right panel).

(B) Mutational coverage of the TL variant library is shown as a heatmap illustrating the allele frequencies of single substitution variants (WT amino acids and positions labeled on x axis; amino acid substitutions on y axis). The WT amino acids for each position are noted with black boxes, and mutants excluded from library synthesis are noted using blue boxes.

(C) Schematic representation of experimental approach. Stars of different colors represent distinct substitutions. The TL variant library PCR amplicon (encoding Rpb1 amino acids 1076-1106) flanked by *RPB1 TL* flanking sequence (orange) was co-transformed with a linearized *LEU2 CEN* plasmid containing an *rpb1* gene with the TL deleted, allowing construction of full-length *RPB1* (with TL variants) by *in vivo* homologous recombination. Heterozygous Leu^+ transformants were replica-plated onto SC-Leu+5FOA to select against the WT *RPB1* (*URA3 CEN*) plasmid and to create TL variant pools. TL variant pools were subsequently replica-plated to different selective conditions for either traditional individual colony screening or high-throughput phenotyping using deep sequencing. For the latter, replica-plated colonies were pooled for genomic DNA extraction, and the TL region was amplified by emulsion PCR to prepare templates for deep sequencing.

functional network within the Pol II TL[31]. Finally, context dependence for TL residue function has been observed, wherein analogous mutations in a conserved TL residue showed opposite effects in *Sce* Pol I and Pol II, suggesting different rate limiting steps for the two enzymes[56]. Together, the intricate intra-TL functional network and the context dependence of TL properties suggest importance of the extensive residue-residue interactions within and outside the TL.

Table 2-1. Plate phenotypes employed for the screening Pol II alleles *in vivo*.

Phenotype	Affected Gene/Reporter Allele; Pol II mutant class affected	WT growth	Mutant growth
Sensitivity to 5FOA	**Detects ability of *rpb1 LEU2* plasmid to complement *rpb1*Δ**	**Resistant to drug. *RPB1 LEU2* plasmid complements *rpb1*Δ**	**Sensitivity to drug (Partial or no complementation of *rpb1*Δ by *rpb1 LEU2*)**
Suppressor of Ty (Spt$^-$)	**lys2-128∂; reports on chromatin defects and start site selection. Specific class of GOF Pol II mutants.**	**Lysine auxotroph (Lys$^-$)**	**Lysine prototroph (Lys$^+$)**
Mycophenolic acid sensitivity (MPA^S)	***IMD2* expression required for resistance; reports on start site selection. Specific classes of GOF and LOF Pol II mutants.**	**Resistance to drug**	**Sensitivity to drug for GOF mutants, relative resistance for LOF mutants.**
Modulation of transcriptional readthrough at *gal10*Δ*56* (Gal^R)	***gal10*Δ*56*; likely reports on termination, mRNA processing and initiation. It is found widely in LOF Pol II mutants, some GOF.**	**Moderate sensitivity to galactose (Gal^S)**	**Resistance to galactose (Gal^R)**

The possible multifunctional nature of each TL residue complicates interpretations of functions if interpretations are based on a limited number of mutants. This is because the phenotype of any given mutant could result from removal of the wild type side-chain or additional functions of the substituted residue. Furthermore, different substitutions may have distinct effects on particular TL conformations[31,234]. In the TL, different substitutions in the same residue can confer distinct phenotypes, so limiting mutational analyses to a single substitution at a particular position may mislead about residue function[29,31]. Deep mutational scanning is an emerging technique for studying large sets of mutants by assessing the enrichment or depletion of variants after a strict selection process[235]. Different selection approaches have been designed such that a specific protein property (sensitivity to substitutions[236], thermo-stability[237], protein stability[238] *etc.*) can be studied. Notably, our established genetic phenotypes (**Table 2-1**) were well correlated with altered transcription elongation rates *in vitro* and specific transcription defects *in vivo*[19,31], thus providing a powerful phenotypic framework for studying TL function. In this work, we have defined the fitness and phenotypic landscape of the conserved, essential *Sce* Pol II TL. We have found three distinct classes of transcriptionally defective TL mutants that are associated with differential stress response profiles, allowing the determination of functional contributions of each TL residue. We have examined the mechanisms by which proximal Pol II domains communicate with the TL, while identifying examples of inter-residue epistasis, which are the likely drivers of incompatibility of RNAP evolutionary variants when placed in the Pol II context.

Results

Strategy for studying *in vivo* effects of TL variant library

A comprehensively mutagenized TL variant library (Rpb1 1076-1106), excepting some previously well-characterized variants[28,31], was synthesized using the Slonomics technology[239,240] and validated by deep sequencing (**Figure 2-1B**). Synthesis conditions were such that single substitution mutants would predominate. Our TL mutant library showed an even distribution of substitutions across all positions and substitution types (**Figure 2-2A,B**), with generally very low frequencies for excluded mutants, as expected (**Figure 2-1B**). We first sought evidence that the measured allele frequencies reflected the real allele frequency distribution because PCR fidelity for highly similar amplicons is often compromised by template switching[241,242]. We spiked in five excluded single substitution variants (H1085Y, H1085Q, F1086S, G1097D, E1103G) as controls. Double mutant variants comprised of these single substitution spike-in variants would not be present in our library, but if observed they would presumably be the result of template switching between spike-ins. We prepared TL amplicons from a subset of conditions using both standard PCR and emulsion PCR (emPCR), which can suppress template switching[241,242]. First, double mutants derived from spike in controls were found at a significantly lower frequency than the relevant single substitution variants; Second, emPCR further suppressed the template switching frequency for all possible double mutants derived from spike-in single variants (**Figure 2-3A, left**), at about 2.5-fold on average (**Figure 2-3A, right**). We conclude that template switching is likely not

extensive in our reactions but further reduction by emPCR led us to employ emPCR for our studies.

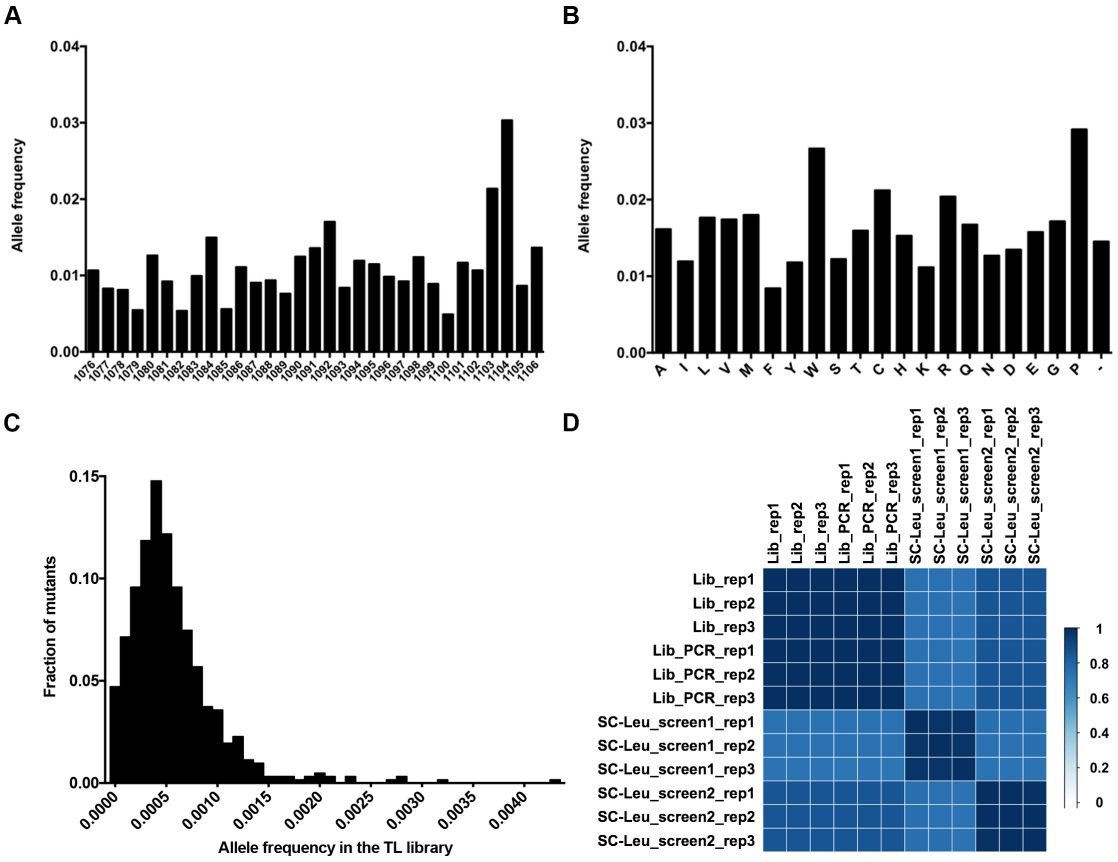


Figure 2-2. TL variant library composition and screening reproducibility.
(A) Fraction of TL substitutions at each position of the TL (Rpb1 1076-1106). Allele frequencies were determined by deep sequencing of the TL variant library, and calculated by the number of reads from all the variants at a position divided by the total number of mapped reads.
(B) Fraction of TL substitutions for codons encoding specific amino acids. The allele frequency for each substitution was determined by deep sequencing of the TL variant library, calculated by the number of reads for variants substituted at a particular substitution divided by the total number of mapped reads.
(C) Distribution of allele frequencies for the detected TL single substitution variants.
(D) The TL library is robust to PCR amplification and yeast transformation. Pearson correlation coefficients calculated between different libraries are shown as a heatmap. TL library (Lib), PCR amplified TL library (Lib_PCR) and two yeast pools independently transformed with TL library (SC-Leu_screen1 and SC-Leu_screen2) were amplified and sequenced in triplicate (rep1, rep2 and rep3), and pairwise Pearson correlation analyses were performed between different sequencing libraries.

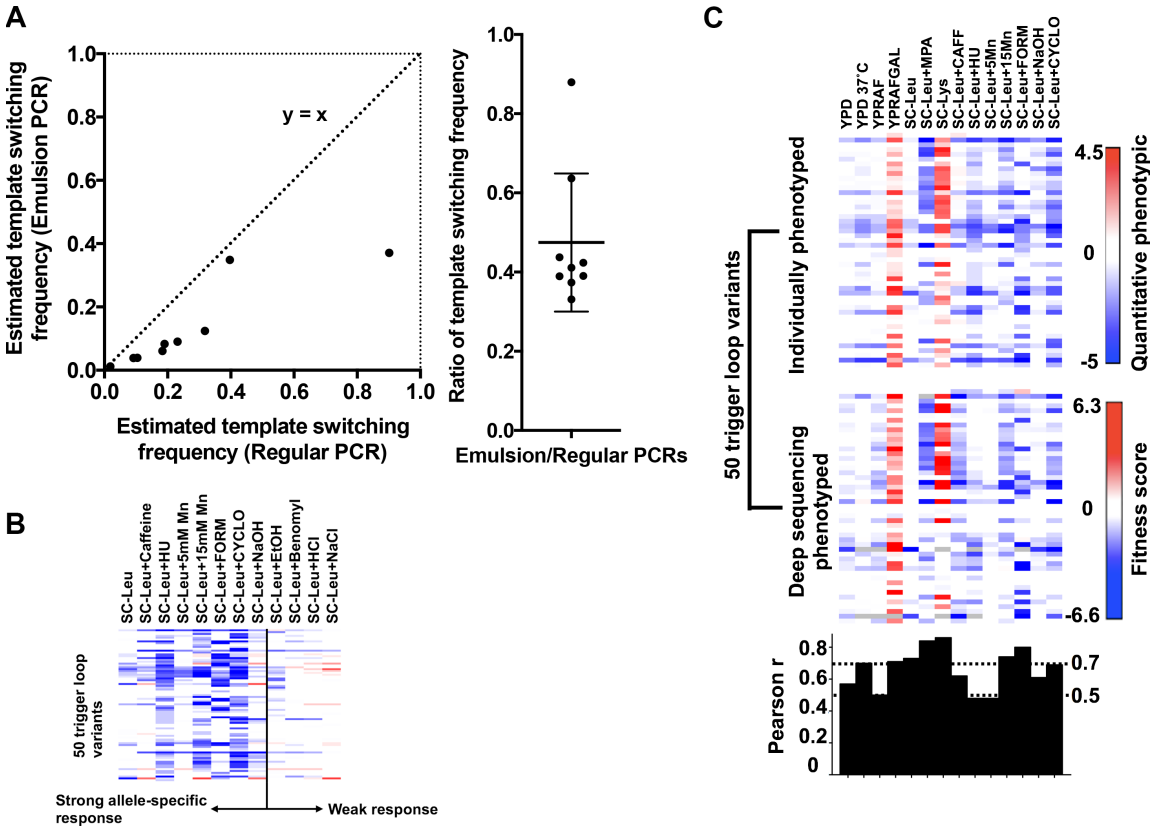


Figure 2-3. Quality controls for the TL high-throughput phenotyping approach.
(A) Comparison of estimated template switching frequencies in regular and emulsion PCR conditions. Template switching was estimated by the ratio ($Freq^{Double}$) / ($Freq^{Single1} \times Freq^{Single2}$) for all the possible double mutants combined from five spiked-in single mutants.
(B) Additional growth conditions were employed to increase resolution for distinguishing similar TL alleles. Growth scores for 50 individually isolated TL mutants (y axis) under 12 growth conditions (x axis), as determined by standard serial dilution plate phenotyping (**Figures 2-4 and 2-5**), are shown as a heatmap. Positive values shown in red indicate increase in allele frequency relative to WT and negative values in blue indicate decrease in allele frequency relative to WT.
(C) High-throughput quantitative phenotyping results are consistent with individual phenotyping of variants. Top heatmap shows qualitative growth scores (as in **Figure 2-3B**) of 50 individually phenotyped TL variants on the y axis (**Figures 2-4, 2-5 and 2-6**) with selective conditions on the x axis. Deep sequencing results for the same mutants using median of fitness defects from three independent high-throughput screens are shown in the middle panel. Pearson r calculated to show the correlation between each condition from the two datasets is shown in the bottom panel.

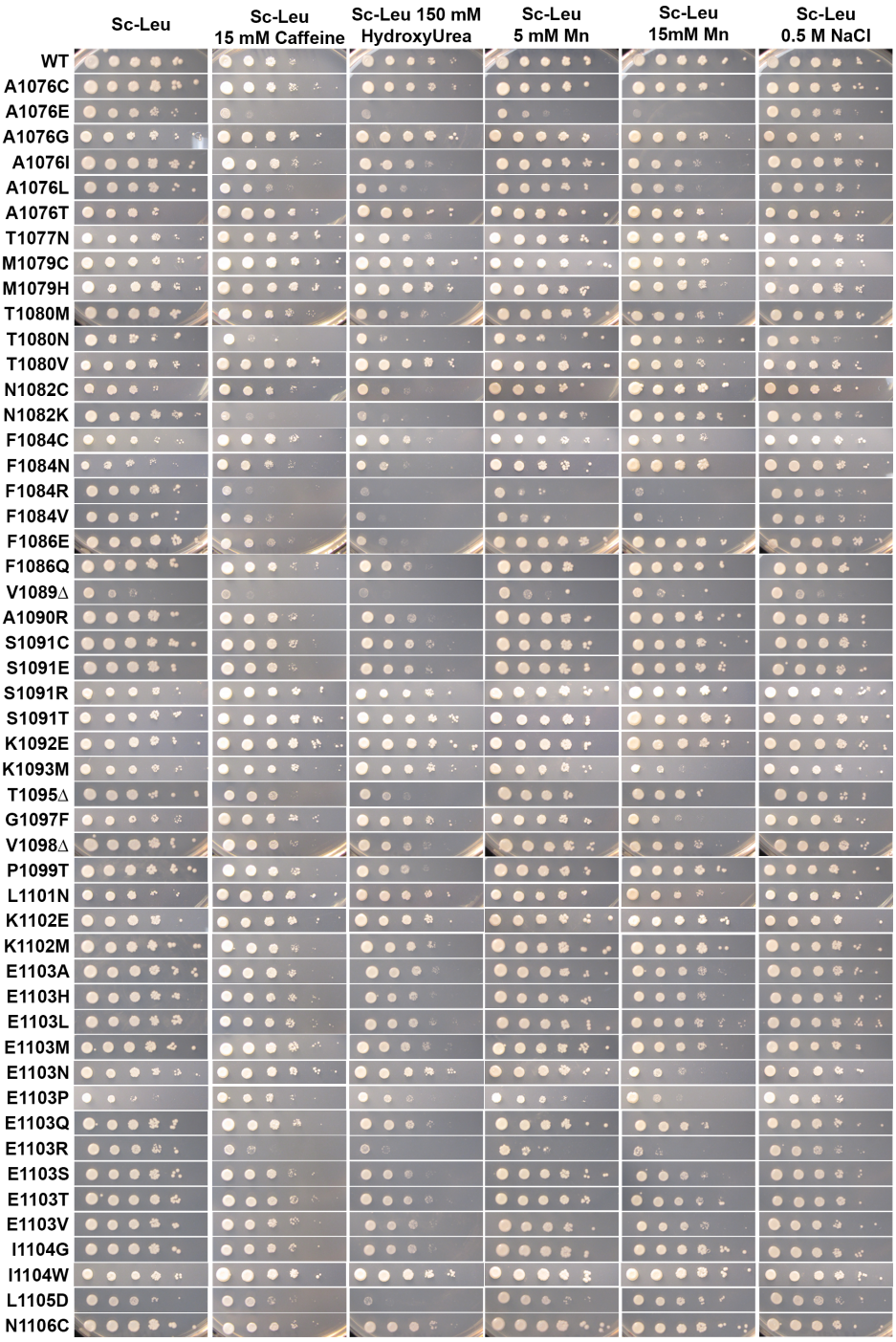


Figure 2-4. Screening for allele-specific stress conditions by standard plate phenotyping of 50 isolated TL variants.

10-fold serial dilutions of saturated cultures of the 50 TL variants were plated on the indicated conditions, including 15 mM caffeine, 150 mM hydroxyurea, 5 mM Mn^{2+}, 15 mM Mn^{2+} and 0.5 M NaCl.

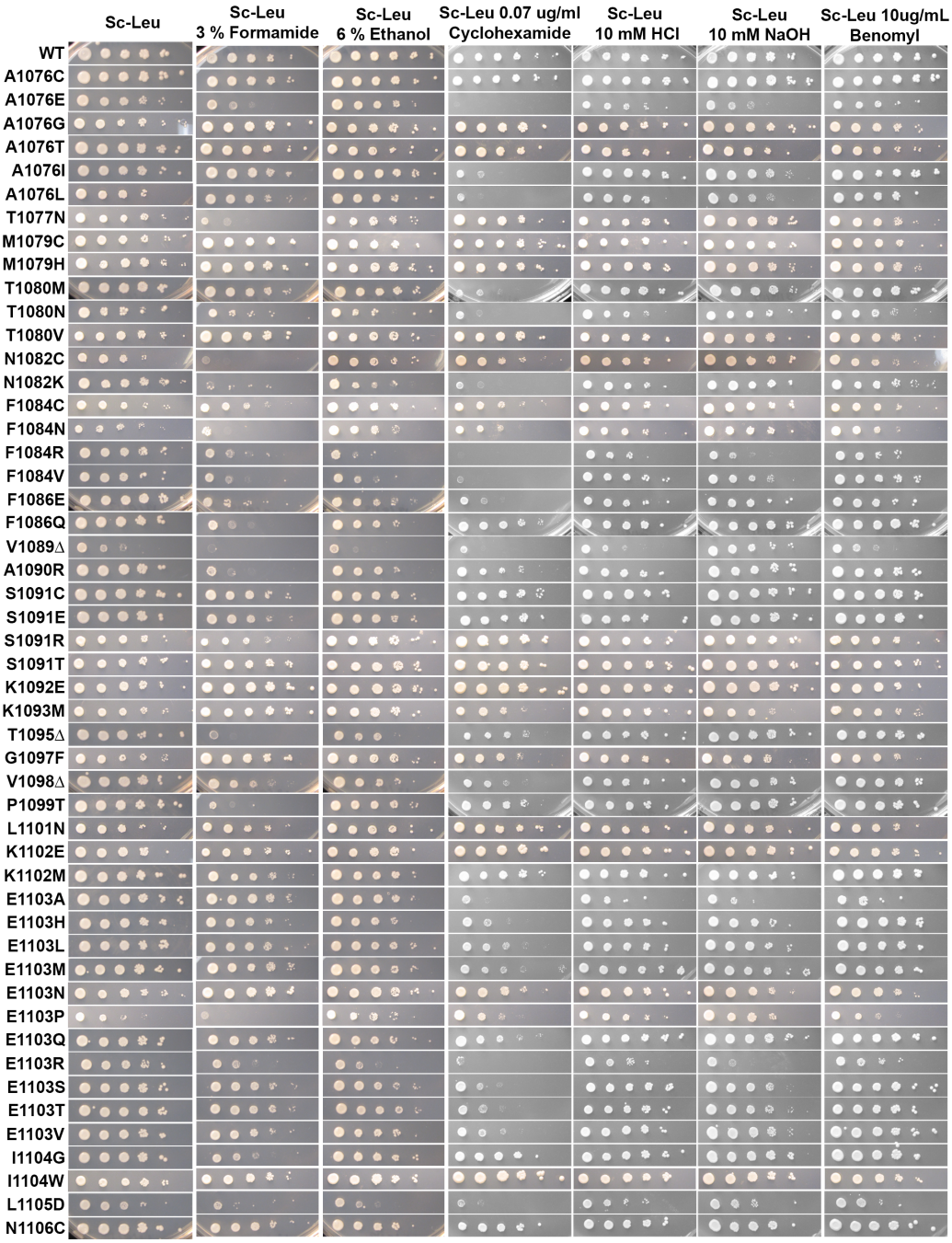


Figure 2-5. Screening for additional allele-specific stress conditions for 50 isolated TL variants. 10-fold serial dilutions of saturated cultures of the 50 TL variants were plated on the indicated conditions, including 3% formamide, 6% ethanol, 0.07 µg/mL cycloheximide, 10 mM HCl, 10 mM NaOH and 10 µg/mL benomyl.

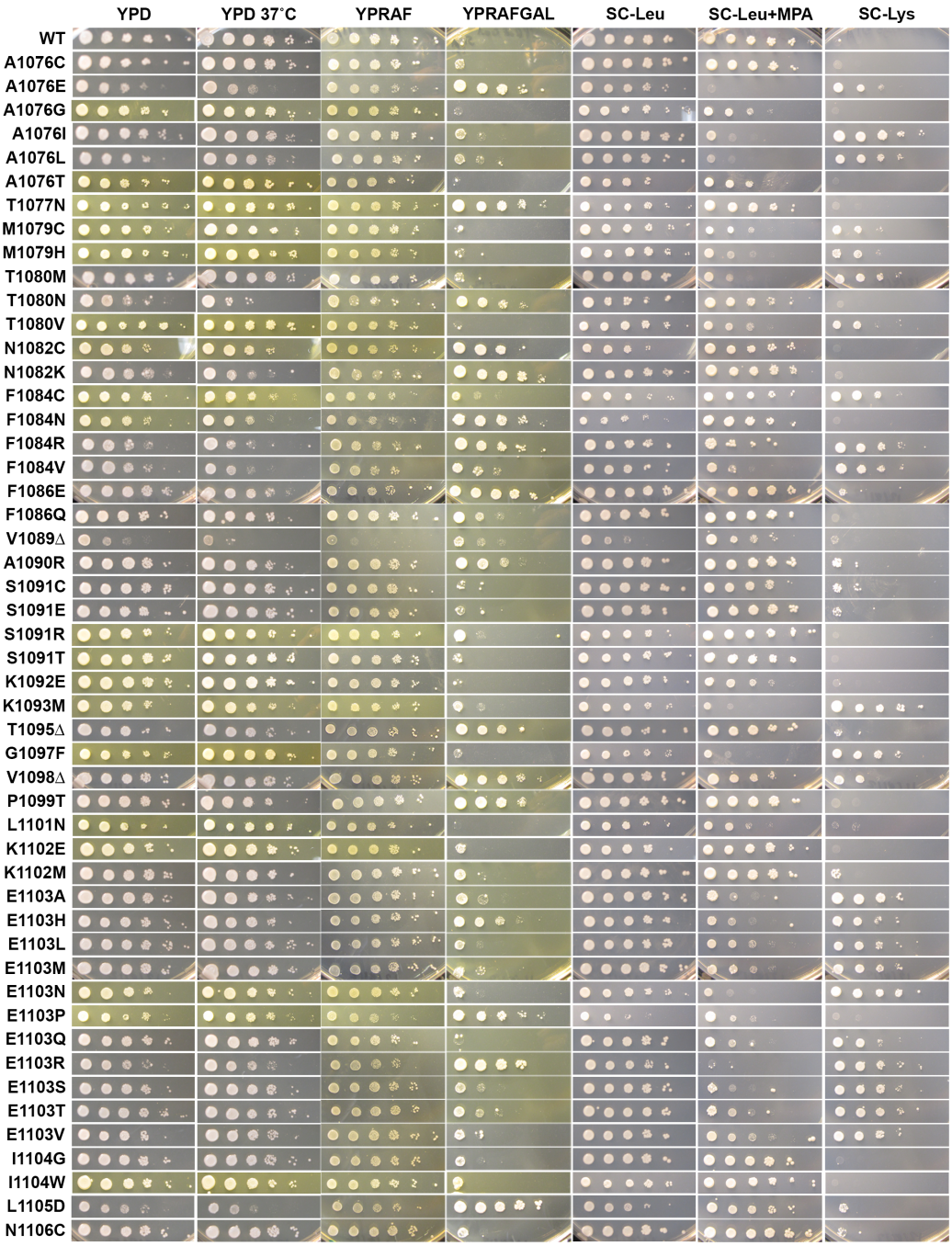


Figure 2-6. Transcription-related phenotypes of 50 isolated TL variants.
Gal^{R}, MPA^{S} and Spt^{-} phenotypes of the 50 TL variants were assessed as a control for the high-throughput phenotyping.

We have developed an experimental pipeline to examine mutations in an essential gene using a plasmid shuffling strategy, and have applied it to study the TL variant library (**Figure 2-1C**). To validate our pipeline and to isolate novel TL alleles, we performed a traditional genetic screening for mutants with transcriptional defects (**Table 2-1**). We have shown previously that these phenotypes correlate with Pol II biochemical activity *in vitro*[19,28,31]. Transcription-related phenotypes employed include, first, the Suppressor of Ty (Spt^-) phenotype, derived from a transposon insertion into the 5′ end of the *LYS2* gene (*lys2-128∂* allele)[243,244]. The transposable element insertion renders wild-type cells Lys^-. A subset of Pol II TL mutants allow expression of a normally silent promoter within the transposable element to express a truncated but functional *LYS2* transcript, conferring the Spt^- phenotype by allowing cells to become Lys^+. Spt^- mutants in the TL correlate with biochemical GOF phenotypes and their related genetic interaction and gene expression signatures[19,31,244]. Second, we employed suppression of the galactose-induced toxicity conferred by the *gal10Δ56* allele of *GAL10*, (Gal^R)[245,246]. *gal10Δ56* contains a deletion in the major *GAL10* polyadenylation signal, allowing transcription readthrough and interference with the downstream *GAL7* gene[245,246]. This readthrough/interference alters the ratio of metabolic enzymes in the galactose-utilization pathway, causing the buildup of a toxic intermediate, resulting in galactose sensitivity (Gal^S). Mutations in transcription elongation factors and Pol II subunits can alter these transcription defects and suppress *gal10Δ56* galactose sensitivity[19,31,246]. Third, we employed Mycophenolic acid (MPA) sensitivity. Sensitivity to MPA for examined Pol II TL mutants derives from altered transcription initiation at the *IMD2*

promoter[31,80], whose transcription is controlled through use of multiple start sites[82,83], and whose expression is required for cell resistance to MPA[81]. We have linked Pol II catalytic activity to the ability to induce *IMD2*. Increased activity Pol II alleles (GOF) fail to induce *IMD2* in the presence of MPA due to aberrant transcription start site selection[31,80]. By screening for these three transcription-related phenotypes, we isolated 1166 candidate mutants, which included 154 singly-substituted and 386 multiply-substituted variants.

To further distinguish mutants, we examined 50 single substitution variants under various stress conditions to screen for conditions that could induce allele-specific phenotypes (**Figure 2-3B, Figure 2-4, Figure 2-5**). We observed that media containing caffeine, hydroxyurea, $MnCl_2$, formamide, cycloheximide, or NaOH induced allele-specific sensitivity or resistance, while media containing ethanol, benomyl, HCl or NaCl showed fewer allele-specific effects (**Figure 2-3B, Figure 2-4, Figure 2-5**). Therefore, in our high-throughput approach, we phenotyped TL variant library under our established conditions (medium lacking lysine (Spt^-), medium containing MPA (MPA^S) or medium containing galactose (Gal^R)) and appropriate media for the stress conditions empirically determined to discriminate among our pilot alleles. Phenotypic scores were estimated from the change of allele frequency normalized to WT, as is standard in mutational scanning studies[235-238]. Quantitative phenotypic scores of the 50 mutants from the high-throughput phenotyping were consistent with semi-quantitative growth scores

derived from standard phenotyping (**Figure 2-3C, Figure 2-4, Figure 2-5, Figure 2-6**), validating our approach.

The Pol II TL fitness landscape

The TL is highly conserved, especially in the NIR, the loop tip residue (Rpb1 G1088) and for several TL C-terminal residues (**Figure 2-7A**). Highly-conserved residues are predicted to be critical for protein function, thus substitutions during evolution are expected to confer fitness defects and be selected against. We first sought to evaluate general fitness defects of observed TL singly-substituted variants (termed the "fitness landscape"), both in the presence of WT *RPB1* (**Figure 2-7B**) and upon the removal of WT *RPB1* (**Figure 2-7C**). Notably, TL NIR and loop tip substitutions conferred large fitness defects in general, while most perturbations in the similarly conserved C-terminal residues did not confer severe growth defects (**Figure 2-7B,C**). This observation highlights that conservation does not necessarily reflect sensitivity to perturbations, and that the TL fitness landscape can further distinguish extremely highly conserved TL residues, as discussed below:

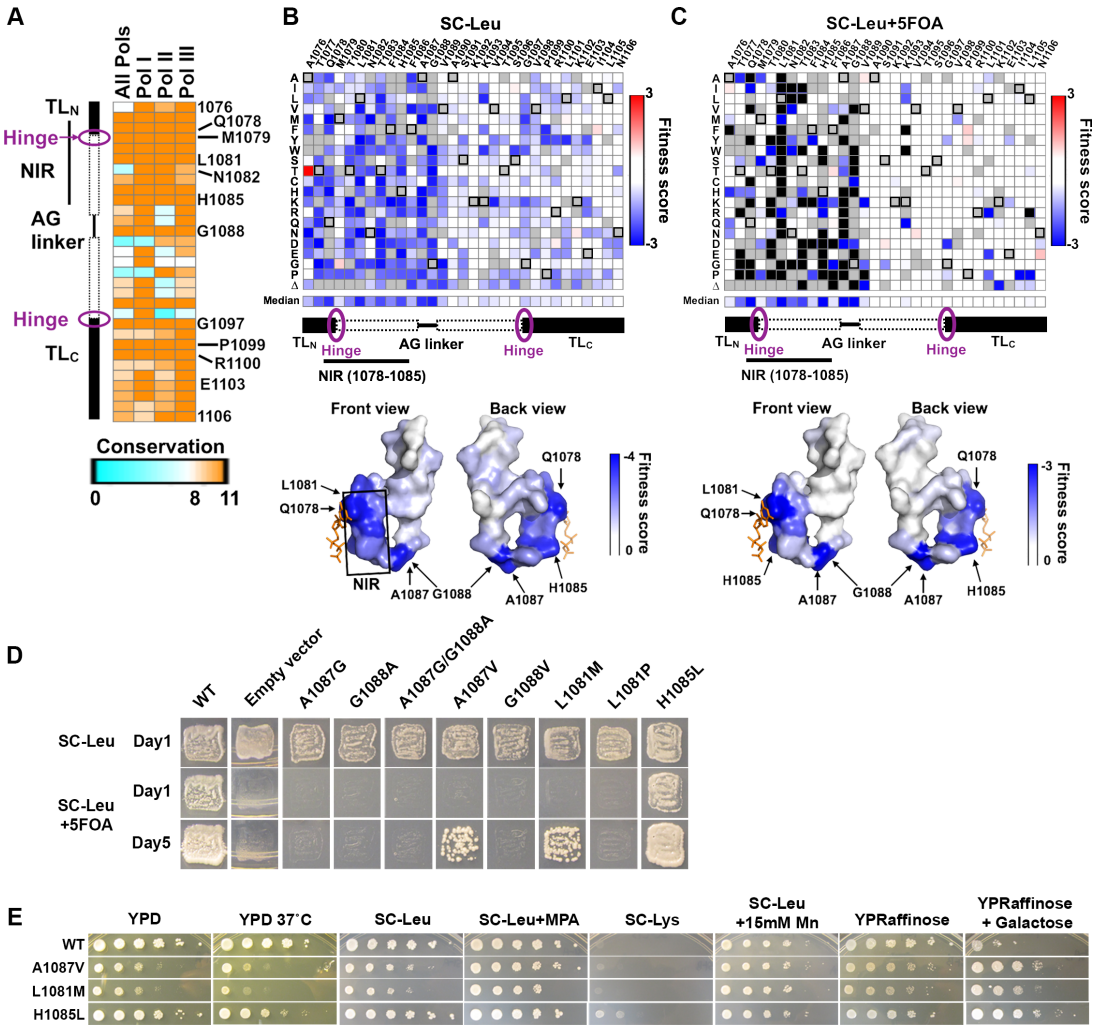


Figure 2-7. The TL fitness landscape distinguishes highly conserved TL residues and reveals high mutational sensitivity in the nucleotide interacting region (NIR) and the Alanine-Glycine linker.
(A) Conservation heatmap of TL residues in eukaryotic RNA polymerases. The conservation scores were extracted from a multiple sequence alignment, including 182 Pol II, 59 Pol I, and 111 Pol III sequences utilizing the conservation metric from Jalview 2.8 version 14.0[247].
(B) Fitness defects of TL variants in the heterozygous state are shown as a heatmap. Unavailable data points are denoted by filled grey squares. WT residues at indicated positions are denoted by black boxes. Surface representation (bottom panel) of the TL structure (PDB:2E2H) is shaded by the median fitness value for all available variants at each position, in a gradient of white (rare defects) to blue (common defects). The position of the matched GTP substrate is shown in orange stick representation.
(C) General fitness defects of TL variants upon removal of WT *RPB1*. Fitness defects predicted to result in lethality shown in black. Surface representation (bottom panel) of the TL structure is shaded by the median fitness value of all available variants at each position, in a gradient of white (rare defects) to blue (common defects).
(D) Complementation abilities of variants in the difficult-to-substitute TL positions (L1081, A1087, G1088) or unexpected TL variants (H1085L) assayed by plasmid shuffling of individual strains. Ability to grow on SC-Leu+5FOA indicates complementation of essential functions of *RPB1*. SC-Leu medium is the control state where WT *RPB1* is present.
(E) Transcription-linked phenotypes of viable substitutions in difficult-to-substitute residues (L1081M, A1087V) or a TL variant with unexpectedly mild fitness defects (H1085L).

First, substitutions in the NIR (Rpb1 1077-1085) generally conferred both fitness defects (**Figure 2-7C**) and apparent dominance (**Figure 2-7B**). Observed fitness defects were consistent with previous observations that several NIR mutants render Pol II slow in elongation *in vitro* and cause fitness defect *in vivo*[28,31]. The observed dominance for many NIR variants was consistent with TL variants being assembled into Pol II complexes that interfere with WT Pol II function, likely through clashes with WT Pol II on genes *in vivo*. Second, substitutions within the alanine-glycine linker (Rpb1 1087-1088) almost universally conferred lethality or severe growth defects. A Pol II structure with a closed TL[24] reveals that A1087 and G1088 are in a tight pocket between the funnel and bridge helices, presumably necessitating small side-chain residues (**Figure 2-8A**). To determine the extent of spatial constraint, we individually assessed the fitness of AG swapping variants, and small hydrophobic valine substitutions (**Figure 2-7D**). Notably, all the swapping variants (A1087G, G1088A and A1087G/G1088A) were lethal (**Figure 2-7D**). While G1088V is lethal, A1087V is severely sick but viable (**Figure 2-7D**), suggesting extremely high, but differential, spatial constraint but differential tolerability for the two residues. This pocket/TL interaction is only observed in the closed TL[24] but not in any of the open states[42], suggesting function in stabilizing the active, closed TL conformation for promoting catalysis. Consistent with disruption of the pocket/TL interaction and the closed TL state, we observed genetically LOF phenotypes for A1087V (Gal^R, slight MPA^R) (**Figure 2-7E**). Finally, substitutions in the conserved C-terminal helix, though not strongly defective in general fitness, are likely to

have transcription defects, based on our prior studies, and were further characterized (discussed below).

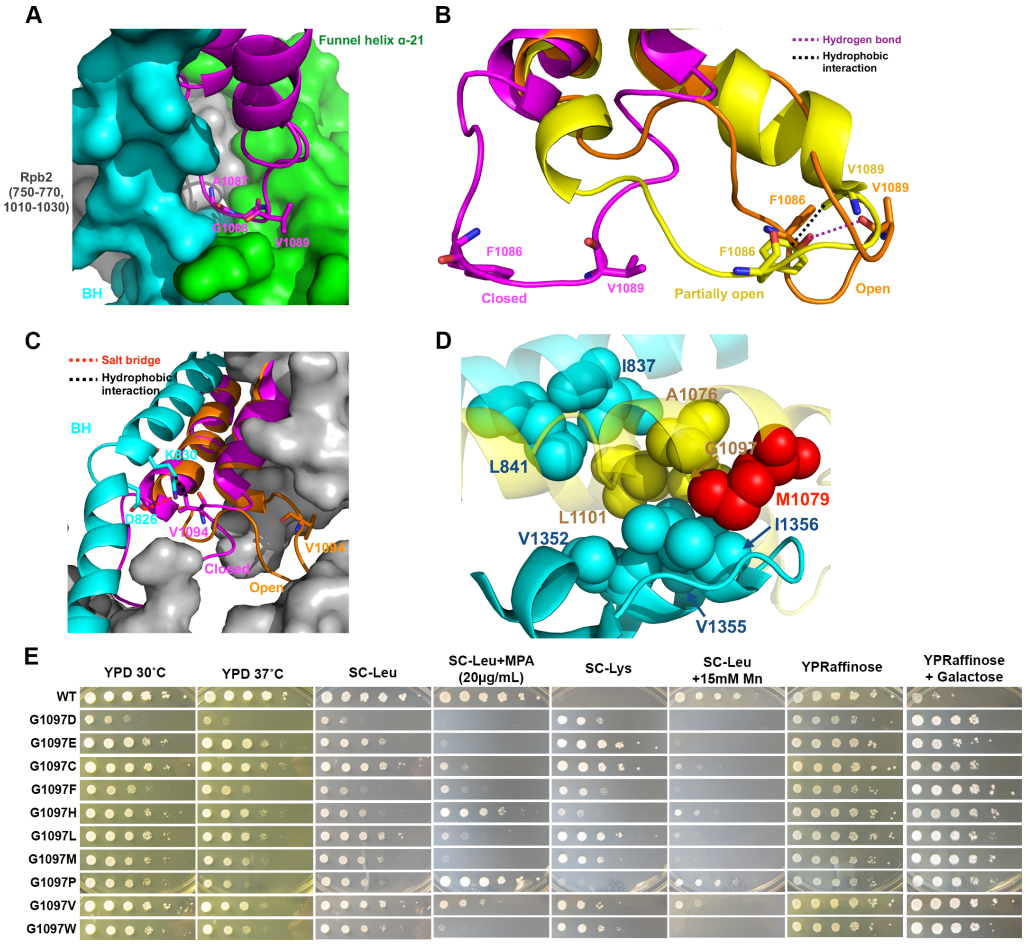


Figure 2-8. Structures of different TL states allow prediction of functionally important residue-residue interactions.

(A) A1087-G1088 linker is highly spatially constrained. The closed TL (magenta) is shown in cartoon (A1087, G1088 in sticks), and TL-proximal domains are shown in surface representation. Rpb2 domains are colored in grey; Bridge Helix (Rpb1 800-860) in cyan; Funnel helix α-21 (Rpb1 700-750) in green.
(B) Change of F1086-V1089 interactions in different TL states. V1089 forms a backbone-backbone hydrogen bond with F1086 in the open TL (orange, PDB: 5C4X), but the side chain flips towards the F1086 for a hydrophobic interaction when the TL is in a less open state (yellow, PDB: 5C4J).
(C) V1094-K830 interaction in the closed TL state. The charged K830 side chain appears to be neutralized by D826 through a salt bridge interaction, and the neutralized K830 side chain interacts with the V1094 side chain.
(D) Observed hydrophobic pocket in the open TL surrounding M1079 (PDB: 5C4J). TL (yellow) and the proximal domains (cyan) are shown in the cartoon representation with the M1079-proximal hydrophobic residues shown in spheres. M1079 is highlighted in red.
(E) Transcription-related phenotypes of G1097 variants.

Novel TL NIR mutants allow mechanistic insights

The TL fitness landscape identified residues highly sensitive to perturbations, while also revealing variants in NIR residues previously known to be difficult to viably substitute. We highlight L1081 and H1085 as two examples. L1081 directly interacts with the nucleobase moieties of matched NTPs[24], and equivalent residues in *Eco, Taq* and *Pfu* RNAPs are important for substrate selection or catalysis[46,229]. L1081 is the most sensitive residue to perturbations among the hyper-conserved NIR. All previously tested L1081 variants were lethal[31], though viable substitutions were identified for all other NIR residues of interest. Furthermore, the GOF allele E1103G can generally suppress lethal substitutions for most NIR residues, but could not for tested L1081 substitutions[31]. In our TL fitness landscape, almost all L1081 variants were indeed predicted to be lethal based on our fitness threshold (**Figure 2-7C**). L1081M conferred a severe growth defect, but was predicted to be just above the viable threshold (**Figure 2-7C**). To validate this prediction, we constructed L1081M for direct analysis, and found that L1081M was indeed viable yet severely sick (**Figure 2-7D**). Furthermore, L1081M conferred Gal^R and slight MPA^R phenotypes, consistent with other LOF mutants (**Figure 2-7E**). Eukaryotic multi-subunit RNA Polymerases share a stringent evolutionary requirement for L at this TL position, while bacterial and archaeal lineages show both M and L variants. Consistent with evolutionary tolerance of variation within bacterial and archaeal lineages, the *Taq* RNAP M1238L variant shows near WT activity for substrate selection and catalysis *in vitro*[229]. The severe growth defect of L1081M highlights

epistasis within *Sce* Pol II and likely eukaryotic RNAP lineages, which imposes a stringent requirement for Leucine at this position.

H1085 interacts with the β-phosphate of the matched NTP[24], and has been implicated in substrate selection, catalysis, intrinsic cleavage and PPi release[47,52]. We previously constructed several H1085 variants (A/N/D/F were lethal, K/R/W/Y caused severe growth defects, Q caused slight growth defect[19,28,31]), suggesting that some polar or positively charged residues, but not a hydrophobic phenylalanine or alanine, could partially complement loss of the histidine[31]. Here, we found that H1085L was viable and healthy in the fitness landscape (**Figure 2-7C**), and validated it with phenotypic analyses of a reconstructed H1085L allele (**Figure 2-7D**). While H1085L conferred slight MPA^R and Gal^R phenotypes, consistent with other LOF mutants (**Figure 2-7E**), it also conferred a slight Spt^- defect, suggesting distinct defects from most other NIR mutants and all known LOF mutants[31]. This observation alters our understanding of the likely bounds of active site chemistry (see discussion).

There are at least three distinguishable TL mutant classes

The overall TL fitness landscape revealed the essentiality of almost all single substitution TL variants in standard growth medium, but could not indicate the nature of transcriptional defects, as we had previously found that both LOF and GOF alleles conferred growth defects. Therefore, we sought to determine the phenotypic outcome of the TL variants for the transcription-related Gal^R, MPA^S and Spt^- phenotypes and a

variety of allele-distinguishing stress conditions (investigated earlier in **Figure 2-3A**). Here, we term this response profile as the "phenotypic landscape", as it distinguishes the TL mutants with presumably distinct transcription defects, in contrast to the general "fitness landscape" described above.

Hierarchical clustering of the phenotypic landscape for 412 TL variants passing fitness filters revealed three major mutant classes with distinct features (**Figure 2-9A, Figure 2-9B**). Class 1 mutants generally conferred a strong Gal^R phenotype yet were Spt^+, and in some cases were also slightly MPA^R relative to WT, consistent with previously characterized LOF mutants. We also identified high formamide sensitivity as a new signature phenotype for Class 1 mutants. Class 2 mutants showed generally weaker Gal^R, slight formamide resistance, and did not confer strong phenotypes otherwise, representing a novel TL mutant class yet to be biochemically characterized. Class 3 mutants generally conferred Gal^R, Spt^- and MPA^S phenotypes, consistent with previously characterized GOF mutants. Mn^{2+} hypersensitivity (Mn^S) was correlated broadly with Spt^- and MPA^S phenotypes, suggesting a relationship among these phenotypes, and consistent with previous *in vitro* biochemical and *in vivo* phenotypic data for a subset of known GOF mutants[248,249]. Notably, our spike-in LOF (F1086S, H1085Q and H1085Y) and GOF mutants (E1103G and G1097D) co-clustered with Class 1 and Class 3 mutants, respectively.

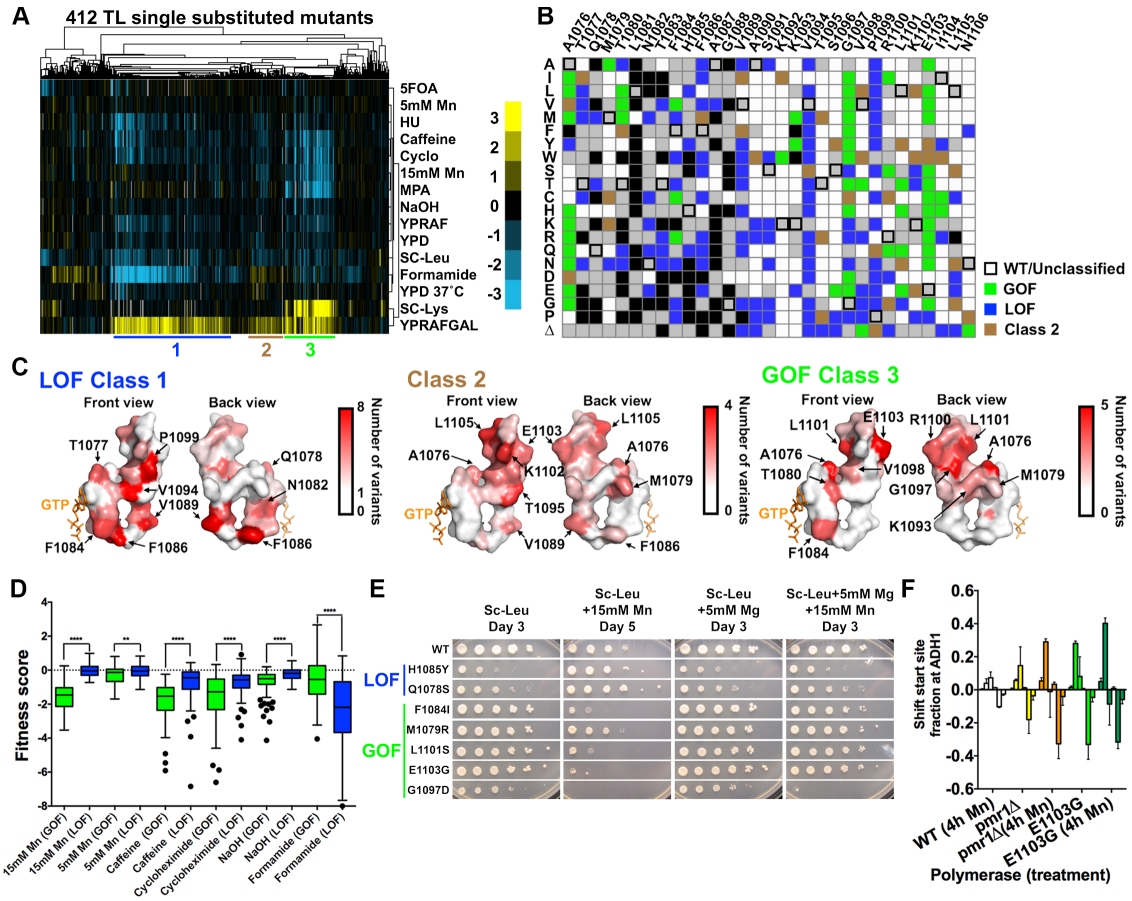


Figure 2-9. Three distinct TL mutant classes, revealed from TL phenotypic landscape, have specific distribution on the TL structure and distinct stress response profiles.

(A) Hierarchical clustering of 412 single TL variants' (x axis) phenotypes (calculated as in **Figure 2-1C**) under 14 different conditions (y axis) reveals distinct mutant classes. Positive (yellow) and negative (blue) fitness scores are shown as a heatmap. Mutant classes (clusters) are annotated by colored lines beneath the heatmap.

(B) Distribution of three major mutant classes is shown in a single substitution variant heatmap. Class 1 (genetic GOF) mutants are shown in green; Class 2 mutants are shown in brown and Class 3 (genetic LOF) mutants are shown in blue.

(C) Distribution of different mutant classes on the TL structure. TL is shown in surface and colored in the gradient from white to red by the number of clustered mutants at each position.

(D) Differential stress responses in genetic GOF and LOF mutants. Genetic GOF mutants are more sensitive to Mn^{2+}, caffeine and cycloheximide, whereas genetic LOF mutants are more sensitive to formamide. **p<0.01, ****p<0.0001 (Two-tailed unpaired t-test).

(E) Differential Mn^{2+} sensitivity and its suppression by Mg^{2+} for selected TL variants representative of mutant classes.

(F) Mn^{2+} effects on different mutants' transcription start sites (TSSs) distribution at *ADH1*, determined by primer extension analysis. TSSs at *ADH1* are distributed in a range of positions and were divided into six bins for quantitation: from upstream (left) to downstream (right). Change of TSSs (normalized to untreated WT) is calculated by the change in TSS fraction for each bin relative to the WT distribution. Average and standard deviation of three experimental replicates are shown as a bar graph with error bars.

Functional contribution of TL residues in different states and substrate-induced TL closing mechanism

The distributions within different mutant classes predict distinct functional contributions of TL residues to TL dynamics. Perturbations predicted to bias the TL towards the active, closed TL state have been shown to result in GOF, whereas destabilization of the closed TL state generally leads to LOF[28,29,31,35,41]. Therefore, distributions of Class 1 (LOF) and Class 3 (GOF) mutants predict alterations to TL dynamics, as follows:

Class 1 (LOF) mutants included most variants from F1086, V1089, V1094 and P1099 (**Figure 2-9C, left**), suggesting important functions of these residues in stabilizing the closed TL. F1086 and V1089 are both proximal to multiple funnel helix residues when TL is closed[24,25], while F1086 was proposed to orient H1085 for correct substrate interaction[25]. Therefore, alteration of these interactions may disrupt the closed TL state and result in LOF. Alternatively, recent Pol II structures with open TL revealed potential function of F1086-V1089 interaction in TL closing dynamics (**Figure 2-8B**)[42]. V1089 forms a backbone-backbone hydrogen bond with F1086 when TL is open, while its side chain flips towards the F1086 to form a hydrophobic interaction when TL is partially closed, suggesting that this side-chain interaction may be important for particular TL states (**Figure 2-8B**), though it was not discussed in previous molecular dynamics studies[25]. Furthermore, V1094 was observed to be proximal to the BH residue K830 in the closed TL state[24]. An interaction between K830 and V1094 side-chains could be counter-intuitive and possibly undervalued. However, neutralization of lysine's positive

charge through ionic interactions (such as D836) can promote hydrophobicity of the lysine side chain[250], supporting the observed K830-V1094 interactions in the TL closed state (**Figure 2-8C**). Most variants in V1094 are LOF (**Figure 2-9B**), consistent with disruption of K830-V1094 interaction and concomitant destabilization of the closed, active TL conformation.

Models for NTP substrate-induced TL closing remain largely untested[24-26,34,35]. A recent Pol II structure[42] exhibiting an open TL state led to explicit implication of a hydrophobic pocket formed by TL residues (A1076, M1079, T1080, G1097 and L1101) and other TL proximal residues (I837, L841, V1352, V1355 and I1356) in substrate-induced TL-folding (**Figure 2-8D**). Q1078 recognition of the 2'-OH of a matched NTP substrate was proposed to promote release of the adjacent residue M1079 from the hydrophobic pocket, triggering TL closing[42,251]. Consistent with disruption of this observed pocket and concomitant destabilization of the inactive open TL state, A1076T, a pocket variant previously isolated as genetically GOF, conferred increased transcription activity *in vitro* (**Figure 2-10B**). Notably, GOF phenotypes were observed for a large number of variants in pocket residues. Among them, we observed almost universal GOF phenotypes for G1097 variants, but not the extreme fitness defects found for the previously observed GOF variant G1097D. We individually phenotyped ten G1097 variants from the traditional screening and confirmed this observation (**Figure 2-8E**). Together, these results are consistent with the hydrophobic pocket stabilizing the inactive, open TL and

providing a plausible mechanism for substrate-induced TL closing. A single residue, M1079, can act as a linchpin for the entire TL through a network of interactions.

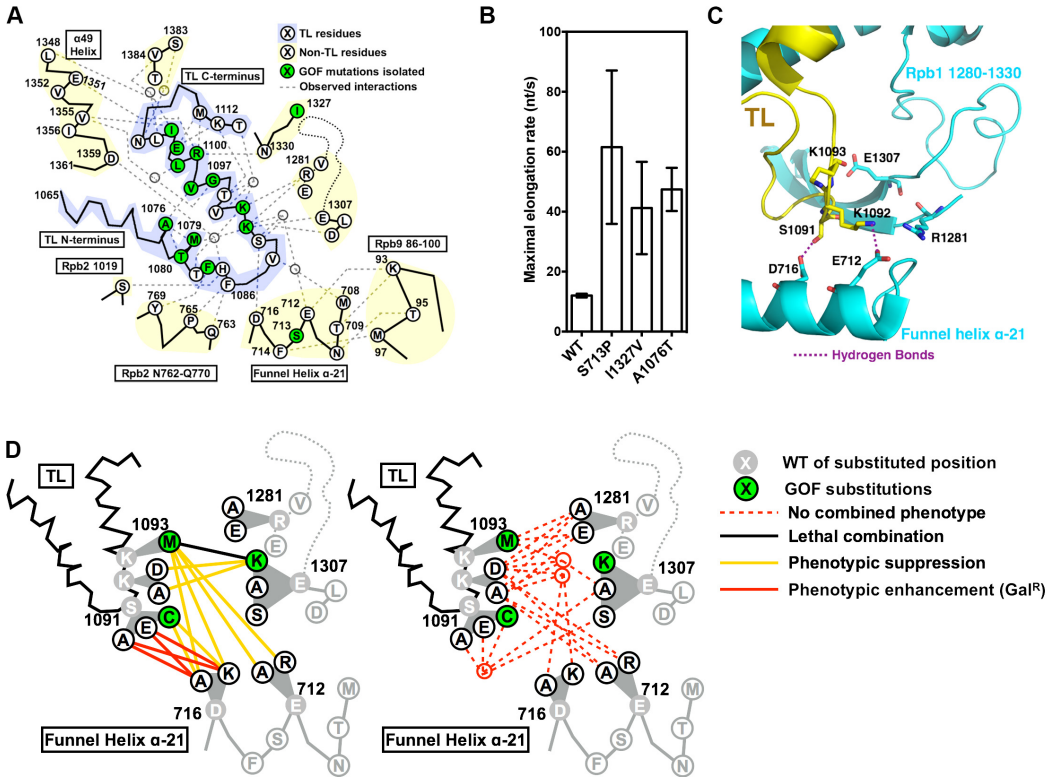


Figure 2-10. Functional contribution of TL tip and Funnel Helix α-21 to proper TL dynamics.
(A) Observed and predicted interactions between TL and TL-proximal domains. TL schematic is shown with residues identified by single-letter amino acid code and positions of interest annotated. Positions of GOF mutants isolated in our screen, along with the positions for a subset of previously isolated TL-proximal GOF mutants, are color coded in green. Observed TL interactions with other Rpb1 domains from structures or simulation studies are shown as grey dashed lines.
(B) Maximal *in vitro* elongation rates (nucleotides/second) of Pol II WT and genetic GOF mutants S713P, I1327V and A1076T.
(C) Observed interactions between open TL tip and TL adjacent charged residues (PDB: 5C4X). Funnel Helix refers to the Rpb1 α-21 alpha-helix.
(D) Genetic interactions between the TL tip and proximal Rpb1 domains. Schematics of the TL and adjacent domains are shown in lines, with positions of interest shown in single-letter amino acid code. Substituted residues are shown in grey, with substituting amino acids shown in white or green filled circles based on single substitution phenotypes (Fig 2-11). Double substitution phenotypes are shown as colored lines connecting the two relevant single substitutions. Some sets of similar interactions were grouped into nodes to reduce complexity in interaction lines.

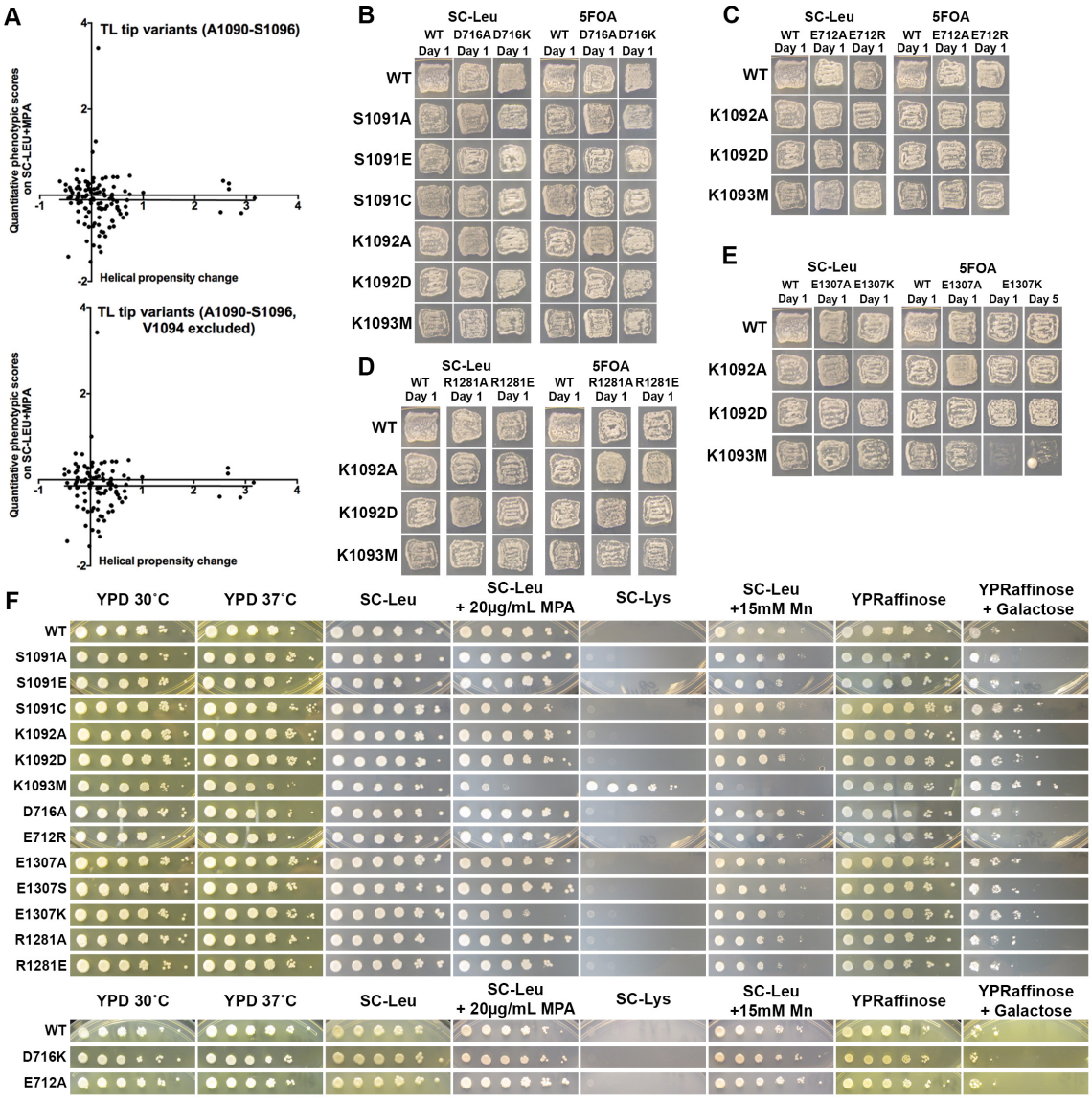


Figure 2-11. Construction and transcription-related phenotypes of the TL tip and nearby charged residue variants.

(A) *x-y* plot showing the lack of correlation between helical propensity change and phenotypic score on MPA, a good indicator of altered transcription activity. 120 variants from the TL tip region (top panel) and 104 variants from the same region but excluding V1094 mutants (bottom panel) are shown, with linear regression fit of the data shown in black lines.

(B-E) Complementation abilities of TL tip (S1091, K1092, K1093) variants, tip proximal D716 (B), E712 (C), R1281 (D), E1307 (E) variants and the corresponding double mutants were determined by plasmid shuffling assays.

(F) Transcription-related phenotypes of TL tip and the TL-proximal charged residue variants. S1091C, K1093M and E1307K confer MPA^S phenotypes, and K1093M additionally confers an Spt^- phenotype, while others alone don't confer any strong transcription-related phenotypes.

Identification of stress conditions that alter transcription *in vivo*

GOF and LOF TL variant classes have distinct phenotypic profiles. In general, compared to LOF variants, GOF mutants are more sensitive to Mn^{2+}, caffeine and cycloheximide yet generally resistant to hydroxyurea and formamide (**Figure 2-9D**). The allele-specific Mn^{2+} response amplified our previous observation that the GOF allele E1103G was highly sensitive to Mn^{2+} while the LOF allele H1085Y was resistant to, or even slightly suppressed by, Mn^{2+} (while the Mn^{2+} effects on both mutants were suppressed by Mg^{2+} supplementation)[52]. The TL phenotypic landscape showed that this Mn^{2+} response was general and class-specific for GOF and LOF mutants (**Figure 2-9D**). To validate this observation, we individually analyzed seven additional variants (two LOF and five GOF) for Mn^{2+} sensitivity in the presence or absence of Mg^{2+} supplementation. Notably, all tested LOF mutants conferred Mn^{2+} resistance while all tested GOF mutants conferred Mn^{2+} hypersensitivity (**Figure 2-9E**). Allele-specific Mn^{2+} responses could be suppressed by Mg^{2+} supplementation (**Figure 2-9E**). Mn^{2+} has been shown to stimulate transcriptional activity while compromising fidelity *in vitro*[248,249]. Our observations suggested that Mn^{2+} may suppress LOF mutants by stimulating transcriptional activity yet exacerbate GOF mutants by further decreasing their already compromised transcriptional fidelity *in vivo*[28,29]. Increased Pol II catalytic activity correlates strongly with upstream transcription start site (TSS) shifts *in vivo*[19,31]; therefore we assayed for TSS alterations upon Mn^{2+} treatment. Primer extension analysis at *ADH1* revealed that Mn^{2+} treatment shifted the TSS distribution upstream, and further exacerbated the upstream shift conferred by E1103G (**Figure 2-9F**). Deletion of *PMR1*, the golgi Mn^{2+}

export channel, causes accumulation cytosolic Mn^{2+} [252,253], and can be used to alter Mn^{2+} levels apart from supplementation of the medium. Our prior high throughput genetic interaction analyses of Pol II mutants showed that *pmr1*Δ strongly interacts with Pol II mutants in a highly allele-specific fashion[19], suggesting an intimate relationship between increased cellular Mn^{2+} levels and altered transcription activity. Here we find that *pmr1*Δ also shifted *ADH1* TSSs upstream (**Figure 2-9F**). While Mn^{2+} may have other indirect effects on Pol II mutants, these observations support direct effects of Mn^{2+} on Pol II transcription activity *in vivo*, raising the possibility that other allele-specific stress conditions (*e.g.* formamide) may also directly alter transcription *in vivo*.

Functional contributions of the TL tip region

The TL tip region (Rpb1 1090-1096) is a random-coil region that forms an α-helical structure when the TL is closed, and helical formation has been proposed to assist TL closing[25,41,234]. Mejia *et al.* characterized two *E.coli* RNAP TL tip mutants I1134V and G1136S (Equivalent to *Sce* Pol II V1094 and S1096) with decreased or increased transcription activity, respectively[234]. These results were interpreted as I1134V and G1136S substitutions decreasing or increasing helical propensity and thus disfavoring or favoring TL closing[234]. *Sce* Pol II contains each of these variants as the WT residue, therefore individual substitutions to the *E. coli* variants (V1094I and S1096G) would be predicted to confer opposite phenotypes under the helical propensity model. However, V1094I and S1096G did not confer phenotypes clearly consistent with either GOF or LOF (**Figure 2-9B**), failing to support the helical propensity model. We asked if the

proposed correlation from *E.coli* RNAP studies was a general property for TL substitutions in this region, if extended to more than two substitutions. Our data, calculated from 122 variants, fail to support a general correlation between helical propensity and predicted catalytic activity for Pol II substitutions in this region (**Figure 2-11A**). As discussed above, V1094 may be involved in interaction with BH residue K830, and LOF in most V1094 variants may result from disrupted BH/TL coordination. Therefore, we repeated the analyses excluding V1094 variants, yet still failed to observe a correlation (**Figure 2-11A**). We cannot rule out contributions of helical propensity in this region to TL function; however, we did not find compelling or widespread evidence for it.

A number of recent studies have suggested potential functions of the TL tip region in regulating TL dynamics[25,42,66]. In a simulated TL closing process, positively charged K1092 and K1093 were predicted to interact with several TL-proximal residues, and some of the predicted interactions were validated by subsequent Pol II crystal structures with alternative open TL states (**Figure 2-10A**). These interactions were proposed to stabilize the open, inactive TL state, and thus alanine (K1092A, K1093A) or charge reversing substitutions (K1092D/E, K1093D/E) were predicted to disrupt the inactive TL open state and result in GOF[25]. Contrary to this prediction, none of the above substitutions conferred GOF (**Figure 2-9B**). Networks of residue-residue interactions near the TL tip were observed[25,42], some of which may be functionally overlapping or redundant, adding complexity to simple models. Our previous point mutant epistatic

miniarray profile (p-EMAP) studies predicted two TL-proximal mutants (S713P and I1327V) to be GOF, which we confirm here (**Figure 2-10B**), suggesting that perturbation near the TL may interfere with native interactions, or create new ones, to destabilize the open TL. The tested variants here also extend the correlation between genetically predicted GOF and increased activity *in vitro* (**Figure 2-10B**). Additionally, several TL tip variants with bulky side chains (K1092W, K1093Y, K1093M) conferred GOF phenotypes (**Figure 2-9B**). Given the complexity and observation of both GOF/LOF phenotypes, we wished to further assess the functions of these residue-residue interactions.

Functional interactions among residues can be explored by the similarity between single substitution variants and the phenotypes of double mutants. We first sought evidence that variants in potential TL interaction partners could confer similar GOF or LOF phenotypes. In the simulation, K1092 switched interaction partners between two funnel helix residues D716 and E712[23], and other charged residues were either observed or simulated to interact with S1091, K1092 or K1093 (**Figure 2-10A**). Therefore, we constructed a panel of mutants in the residues D716, E712, R1281, E1307, and D1309 for phenotypic analyses. Notably, we observed GOF phenotypes (Mn^S and MPA^S) in E1307K but not E1307A, suggesting that E1307K gained an interfering interaction to destabilize the open TL state. In contrast, mutations in the residues E712 and D716 did not confer transcription-related phenotypes.

To further dissect functional relationships, we phenotyped double mutants from potential interaction partners, and observed a number of genetic interactions (**Figure 2-10D**, **Figure 2-12**). First, mutations in D716 (D716A/K) suppressed the S1091C and K1093M GOF phenotypes and conferred LOF phenotype (Gal^R) when combined with other S1091 alleles (D716A/S1091A, D716A/S1091E, D716K/S1091A, D716K/S1091E), although the D716A/K or S1091A/E single substitutions did not confer strong transcription-related phenotypes (**Figure 2-10D**, **Figure 2-11F**, **Figure 2-12A**). These allele-specific genetic interactions suggested that D716A/K and S1091A/E lost putatively redundant interactions that together conferred LOF phenotypes, and loss of D716 interaction(s) might also suppress the putatively gain of interactions in GOF mutants S1091C and K1093M. Second, mutations in E712 (E712A/E712R) did not confer transcription-related phenotypes, but suppressed the K1093M GOF phenotypes (**Figure 2-11F, Figure 2-12B**). Similarly, K1092A/D single substitutions did not confer transcription-related phenotypes, but were able to suppress the E1307K GOF phenotypes (**Figure 2-11F, Figure 2-12E**). This observed epistasis suggested that loss of potential interacting residues (E712 and K1092, respectively) relieved putative gain of interactions in the GOF mutants (K1093M and E1307K, respectively) (discussed above). Taken together, the observed allele-specific and epistatic interactions between TL tip and proximal residues suggest a highly complex genetic network of residues controlling TL dynamics, and illustrate how individual residues might constrain or allow diversification of the TL through evolution.

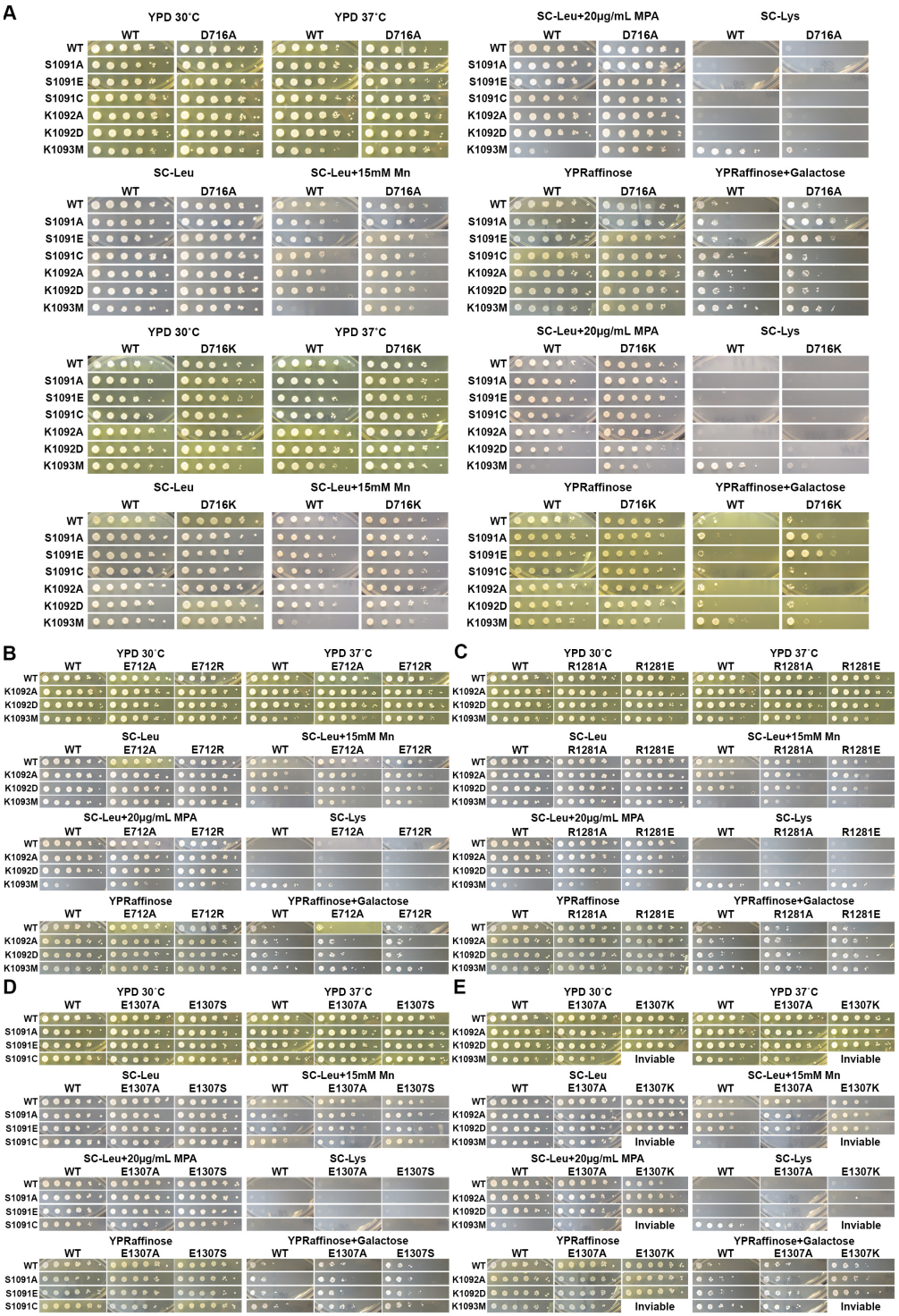


Figure 2-12. Genetic interactions of the TL tip and nearby charged residue variants on transcription-related phenotypes.

Genetic interactions between tip variants and nearby charged residues D716 (A), E712 (B), R1281 (C) and E1307 (D, E) variants detected by alterations in transcription-related phenotypes.

Functional interplay of the TL and Bridge helix (BH) domains

The BH is a strikingly conserved structural domain of multi-subunit RNA polymerases spanning the wide central cleft between polymerase "jaws", adjacent to the active site and proximal to the TL[167,168,254]. Although the BH is a straight helix in most published structures[20,21,24,166,167,228], some *Thermus thermophilus* RNAP structures revealed a bent BH conformation proposed to support translocation[168]. This BH bending mechanism was supported by a number of simulation studies but has never been directly tested[40,59,167,168,254]. In the archaeal *Mja* RNAP, proline substitutions at two hinge-proximal residues M808 and S824 (equivalent to *Sce* Rpb1 M818 and T834) resulted in GOF, suggesting kinking by the proline substitution results in increased translocation or catalysis[59,60]. Furthermore, *Mja* GOF TL and BH mutants were not additive when combined, suggesting mutual dependence on BH and TL functions[59].

To explore the functional consequence of BH kinking in *Sce* Pol II, we constructed and phenotyped BH mutants analogous to the characterized GOF and LOF variants in *Mja* RNAP. Notably, *Sce* T834 and other BH C-terminal hinge substitutions conferred *in vivo* phenotypes consistent with the altered transcriptional activities in *Mja* RNAP (**Figure 2-13F**), and we directly confirmed the altered activity of T834 variants *in vitro* (**Figure 2-14A**). In contrast, substitutions in M818, a predicted BH N-terminal hinge, showed defects deviating from expected conservation of function. M818P caused lethality, and could not be suppressed by any tested TL variants, precluding us from classifying it (**Figure 2-13A**). Furthermore, M818S and M818Y, although viable, did not confer any

clear phenotypes (**Figure 2-13F**). Therefore, we further assessed the functional interplay between BH and TL by double mutant analyses, including BH variants (M818S/Y, T834A/P) and TL substitutions covering a range of altered transcriptional activities (**Figure 2-14B-E**). Notably, the GOF BH variant T834P, along with M818S and M818Y, were mutually suppressive with biochemically strong LOF TL variants (**Figure 2-14B,C,E**), revealing both additive behavior between BH and TL for some combinations, and cryptic phenotypes for M818S/Y in others. The LOF BH variant T834A also suppressed GOF TL variants (**Figure 2-14D**). However, the additive interactions (exacerbation, synthetic lethality) we observed for GOF BH and TL double mutants were in contrast to the epistasis for *Mja* RNAP[59].

Multiple lines of evidence suggested additional, specific defects exist in BH mutants, beyond simple cooperation with the TL. First, M818P lethality could not be suppressed by any tested TL variants (**Figure 2-13**), which cover a wide range of transcriptional activities. Second, suppression between BH and TL mutants of different biochemical classes (GOF/LOF) was partial and not as strong as the previously observed intra-TL suppression. Third, GOF M818S, M818Y and T834P variants appeared to exhibit activity-dependent genetic interactions with TL variants. BH GOF variants suppressed strong LOF TL variants Q1078S and H1085Y but failed to suppress, or even exacerbated slightly LOF TL variants H1085Q and F1086S (**Figure 2-14B, C, E**), consistent with conditional epistasis, where GOF activity of BH variants can suppress either specific TL variants or otherwise exert their effects in specific contexts. Finally,

recent modeling studies predicted that the BH residue Y836 assists Pol II forward translocation[255] by interacting with the DNA:RNA hybrid. Y836A/H conferred Gal^R phenotypes, consistent with LOF and compromised translocation (**Figure 2-13F**). Notably, GOF T834P was suppressed by Y836A/H (**Figure 2-14E, Figure 2-15B**), consistent with T834P conferring a TL-independent fast translocation defect, suppressible by Y836A/H.

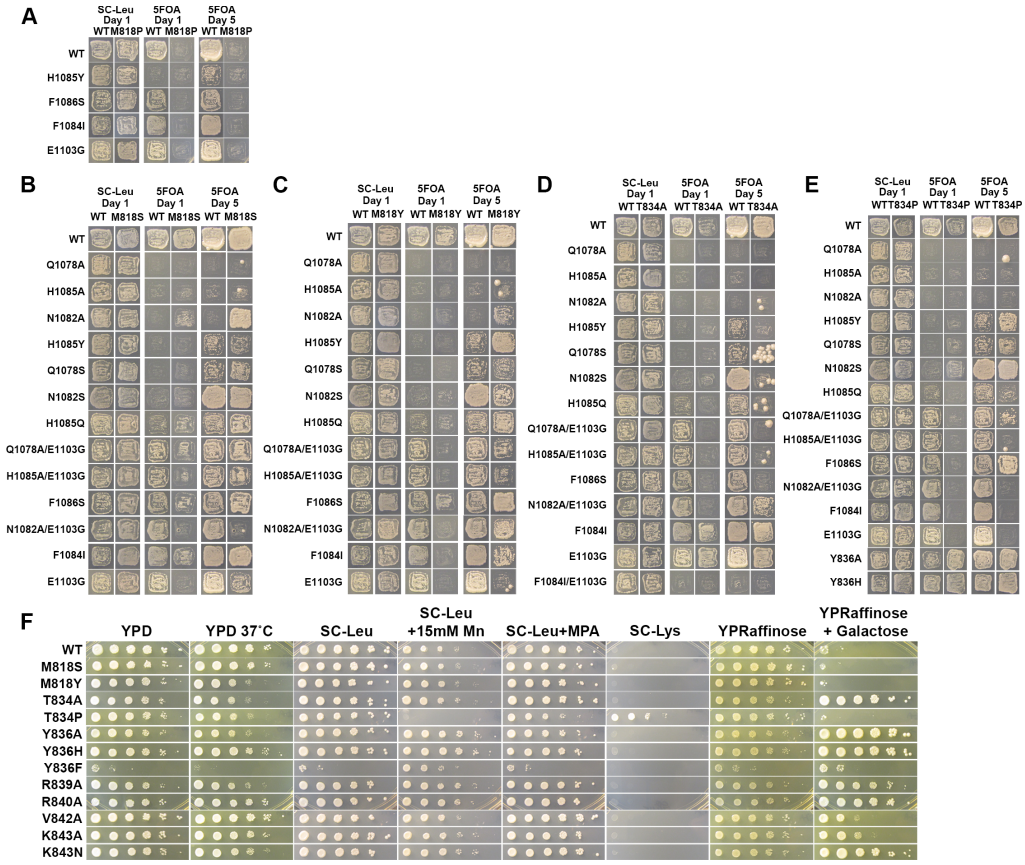


Figure 2-13. Construction and transcription-related phenotypes of TL and BH variants.
(A-E) Complementation ability of the indicated TL variants, BH single variants M818P (A), M818S (B), M818Y (C), T834A (D), T834P (E) and the corresponding double/triple mutants were determined by plasmid shuffling assays.
(F) Transcription-linked phenotypes of BH single-substituted mutants. M818S and M818Y are substitutions in a predicted BH N-terminal hinge; others are substitutions in predicted BH C-terminal hinge positions or additional C-terminal substitutions.

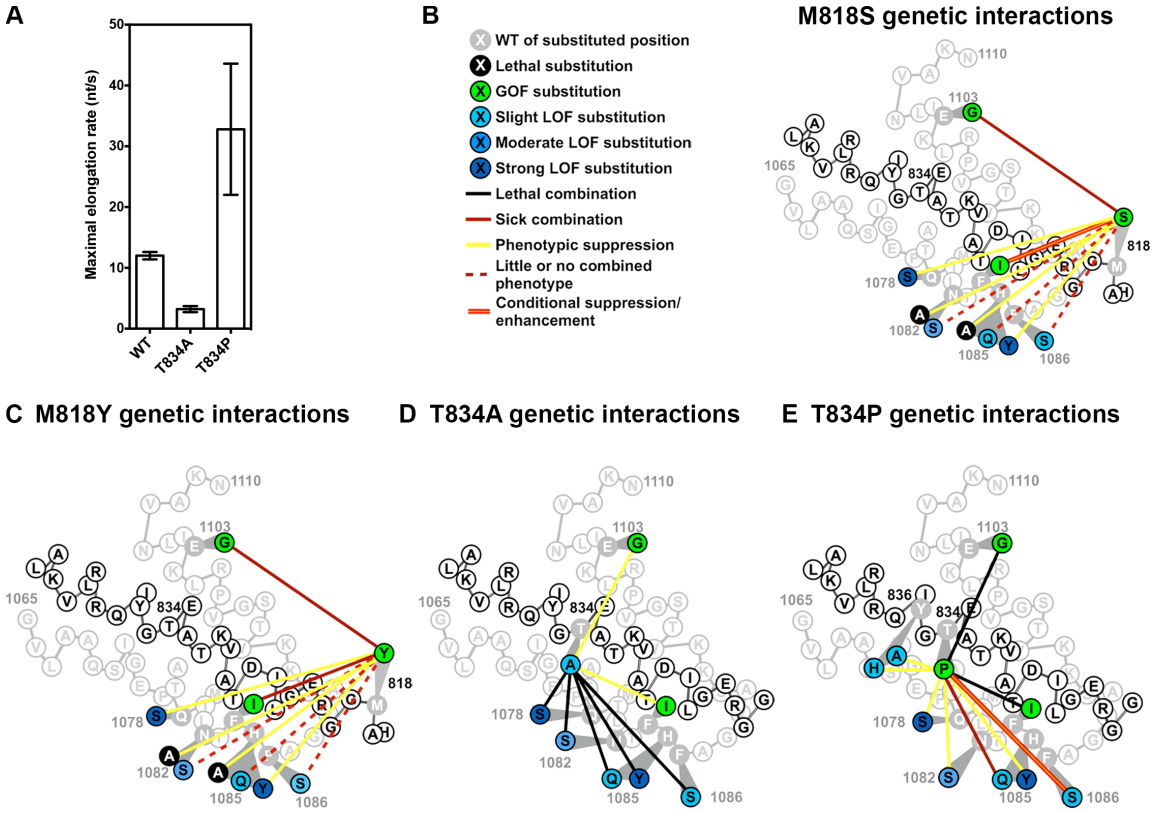


Figure 2-14. Functional interplay between the TL and Bridge Helix (BH).
(A) Maximal *in vitro* elongation rates (nucleotides/second) of BH variants T834A and T834P.
(B) Genetic interactions between BH M818S and TL substitutions. M818S suppressed (yellow lines) the strong LOF TL variants (dark blue) but not the slight and moderate LOF TL variants (light blue), and showed synthetic sickness (red lines) with the GOF TL variants (green).
(C) Genetic interactions between BH M818Y and TL substitutions. Similar to M818S genetic interactions with TL variants (**Figure 2-14B**), M818Y suppressed (yellow lines) the strong LOF TL variants (dark blue) but not the slight and moderate LOF TL variants (light blue), and showed synthetic sickness (red lines) with GOF TL variants (green).
(D) Genetic interactions between BH T834A and TL substitutions. T834A suppressed (yellow lines) the GOF TL variants and was synthetic lethal with all the tested LOF TL variants (blue).
(E) Genetic interactions between BH T834P and TL or BH. Similar to M818 variants (**Figure 2-14B, C**), T834P suppressed the strong and moderate LOF TL variants (dark blue) but was synthetic sick with weak LOF TL variants (light blue), while synthetically lethal with GOF TL variants (green). T834P was also suppressed (yellow line) by two LOF BH mutants Y836A/H.

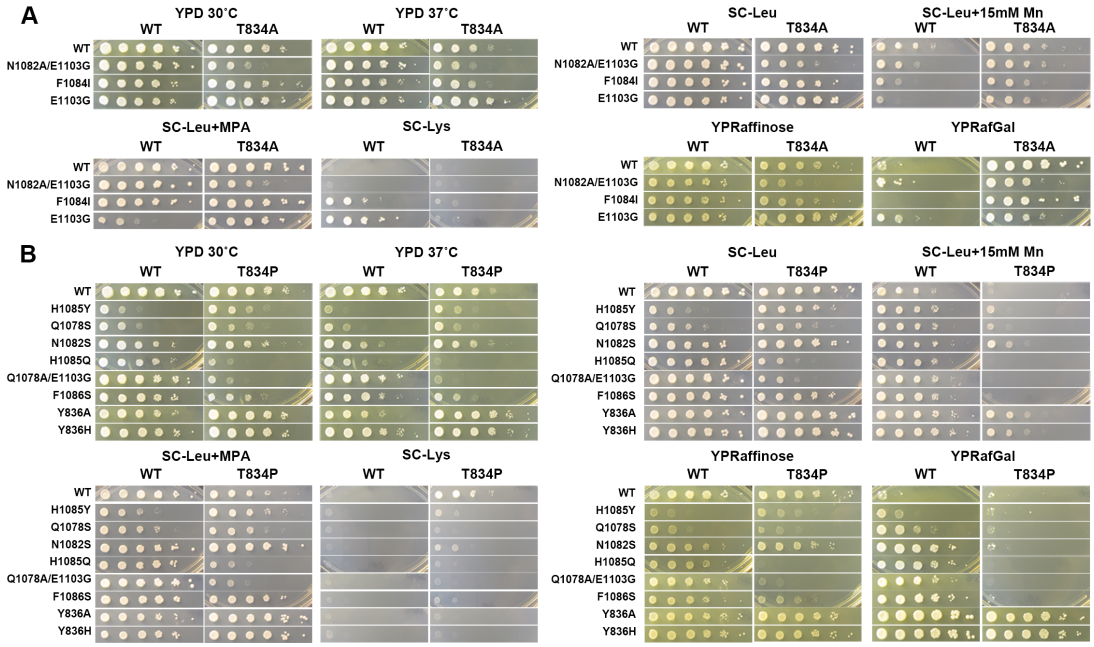


Figure 2-15. Genetic interactions of BH T834 and TL mutants on transcription-related phenotypes. Genetic interactions between TL variants and the BH variants T834A (A), T834P (B) were assessed by standard plate phenotyping of transcription-related phenotypes. Additional genetic interactions between T834P (GOF) and two LOF BH mutants (Y836A) and Y836H are included in (B).

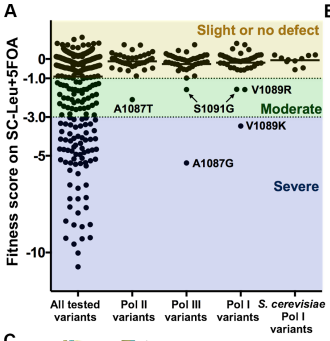
Context dependence of TL function

We previously observed that E1103G, a GOF allele in *Sce* Pol II, caused LOF in Pol I, highlighting divergent contributions of active site residues in different enzymatic contexts[56]. We also observed that the Pol I TL[56] and L1081M (this study) were functionally impaired in the Pol II context. We next sought to determine the functional compatibility of other evolutionary TL variants in the *Sce* Pol II context, using our fitness and phenotypic landscape (**Figure 2-16**). Most tested evolutionary TL variants did not confer fitness defects, with several exceptions (**Figure 2-16A**). Furthermore, some variants, although compatible for general growth, conferred transcription-related phenotypes and could be further classified by our phenotypic landscape (**Figure 2-16B**). These observations further suggest that the evolution of TL function is shaped by likely epistasis between the TL and proximal domains.

We next asked what substitutions might underlie the large difference in compatibility of the *Sce* Pol I TL (versus the *Sce* Pol III TL) within Pol II[56]. From our phenotypic landscape, although many individual *Sce* Pol I and Pol III TL substitutions appeared to be compatible, functionally impairing variants were identified (**Figure 2-16B**). The yeast Pol III TL contains Pol II GOF (A1076G) and LOF (N1082K) variants, both of which hypothetically could be mutually suppressive, resulting in close to WT activity in the Pol II context[56]. The Pol I TL contains three Pol II LOF substitutions (V1089H, A1090G and S1091A). The net incompatibility of Pol I TL is consistent with additive defects of the

three LOF variations, given that most TL LOF combinations show additive effects[31]. Since three evolutionarily observed variants with LOF phenotypes were all localized in the TL tip, we examined the difference between Pol I and Pol II structures for the TL tip proximal domains[42,256]. The Pol I funnel helix appears to impose less constraint than the Pol II funnel helix (**Figure 2-16C**), suggesting that Pol I controls its TL with a distinct network of interactions. In all, our mutational data, together with the recent Pol I crystal structure, reveal enzyme-specific mechanisms to control a highly conserved domain at the heart of eukaryotic transcription.

A



B

	Pol I evolutionary variants	Pol II evolutionary variants	Pol III evolutionary variants
Class 1 (LOF)	V1089H, V1089R, A1090G, S1091G, S1091A, V1094A, V1094N, V1094L	A1090G, S1091A, V1094T, K1102T	N1082K, N1082Q, V1089T, A1090G, S1091G, S1091A, S1091N, V1094L
Class 2	V1089F, V1098L	V1089I, V1098L	V1089I, K1102E
Class 3 (GOF)	N.A.	A1076I, S1096E	A1076G, T1080V
Slight or no defect	A1076C, A1076S, M1079L, F1086L, V1089L, A1090I, A1090L, A1090S, S1091D, K1092A, K1092M, K1093N, K1092S, K1092V, T1095K, T1095N, S1096L, S1096M, S1096Q, V1098I, V1098M, L1101I, L1101M, K1102R, I1104L, I1104M, L1105I, L1105V, N1106I, N1106L, N1106M	A1076S, A1076T, F1086H, F1086L, A1087T, A1090S, A1090T, S1091D, S1091T, K1092A, K1092H, K1092R, K1093D, K1093N, K1093S, K1093T, V1094I, S1096G, S1096K, S1096L, S1096T, L1101F, L1101V, K1102N, K1102Q, I1104L, L1105I, L1105M, N1106D	A1076S, T1083S, F1086H, V1089L, A1090S, S1091D, K1092L, K1092M, K1092Q, K1093A, K1093N, K1093S, V1094I, S1096A, S1096L, S1096Q, V1098I, L1101I, L1101M, K1102N, K1102Q, K1102V, I1104V, L1105I, L1105M, L1105V, N1106D, N1106G, N1106S

C

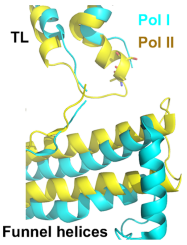


Figure 2-16. Phenotypic analyses of evolutionary variants suggest context-dependent functions for many TL residues.
(A) General growth fitness defects of the TL single-substituted variants observed in the TL across Pol I, II, III evolution including 38 Pol II, 42 Pol I and 42 Pol III amino acid variants relative to *Sce* Pol II.
(B) Evolutionary TL variants in three mutant classes from the TL phenotypic landscape (**Figure 2-9A, B**). Existing variants from *Sce* Pol I are colored in blue, and existing variants from *Sce* Pol III are colored in red. *Sce* Pol I has three substitutions (V1089H, A1090G and S1091A) that cause LOF in the Pol II context; *Sce* Pol III has one substitution (A1076G) classified as GOF and one substitution (N1082K) classified as LOF.
(C) Difference in positioning of funnel helices (relative to TL) in Pol I and Pol II. Cartoon representation of TL/funnel helices from Pol I and Pol II are shown in cyan and yellow, respectively (PDB: 5C4J and 2VUM).

Discussion

The ability of the TL to fold into multiple conformations and the dynamic conversion between these states are critical for its functions. Previous studies from us and others demonstrate that TL function is delicately balanced, such that perturbations result in either increased or decreased catalytic activity and altered translocation dynamics. Distinct consequences for transcriptional activity manifest *in vivo* as what we term LOF and GOF phenotypes. In this study, we have advanced our genetic framework with which to dissect Pol II mechanisms. From our phenotypic landscape, we assessed the functional contributions of almost all TL residues to fitness in *S. cerevisiae* under multiple conditions. Our data indicate that both intra-TL interactions and TL interactions with nearby domains (*e.g.* BH and funnel helices) are critical for TL function. This conclusion is also supported by recent work on Rpb9 organizing the TL indirectly through an Rpb1 TL-adjacent α-helix 21 (one of the funnel helices, discussed below)[66], interactions between the TL and F-loop regions in bacteria[49], and predictions of TL-proximal variants as GOF from our previous pEMAP analysis[19] (validated in this study). Our system allows efficient analysis of a large number of variants to evaluate accumulating computational[25,30,39,40] and structural[21,24,42,44] predictions for interactions within the TL and from without.

The major function of the TL is to link substrate recognition to catalysis, while it is also proposed to gate translocation such that translocation probability is linked to phosphodiester bond formation. Critical to this recognition is that a substrate be

positioned correctly by base-pairing to the DNA template, and that the 2'-OH allows NTPs to be selected over 2'-dNTPs by the TL residue Q1078[24,45,229]. We have proposed that the Q1078-substrate interaction releases the adjacent M1079 from its intra-TL hydrophobic pocket to trigger TL closing[42]. In this study, we find a great number of variants within the pocket residues A1076, M1079, G1097, L1101 to cause GOF phenotypes, providing evidence that disruption of the hydrophobic pocket destabilizes the open, inactive TL state. Additionally, while the TL shows incredibly high evolutionary conservation for a number of residues, prior work indicated alteration of ultra-conserved residues (*e.g.* E1103 in Pol II, E1224 in Pol I) in different RNA polymerases could have distinct effects, suggesting the importance of the evolved context within each enzyme[28,31,56]. Here, we evaluate many evolutionarily observed eukaryotic TL variants in the *Sce* Pol II system, and discover a number of functionally impaired TL variants. Our results highlight that TL proximal domains may impose constraint and also allow functional diversification in the molecular evolution of the highly conserved TL by epistatic interactions.

One example of a proximal region, the so-called "funnel helices" (Rpb1 α-20 and α-21) or "rim helices" in the bacterial RNAP literature, shows both evolutionary conservation and functional diversification. Funnel helices are both surface exposed and proximal to the TL[42,257]. Multiple pieces of evidence from three mutations in α-21 suggest roles for funnel helices in controlling TL function. One, the C4 allele of *Drosophila melanogaster*, corresponding to R726H in *Sce* Rpb1, confers a slow elongation rate in both Drosophila

(*in* vitro) and human Pol II enzymes (in cells)[258,259]. The molecular mechanism of this allele is not currently known, but based on another α-21 substitution (G730D) identified in yeast, we would speculate C4 enzymes have altered TL dynamics. G730D was identified in yeast twice, in independent genetic screens[260-262]. *rpb1*-G730D is catalytically slow[263], confers a severe growth defect but can be suppressed by a GOF mutant, *rpb9*Δ[66,260]. In fact, *rpb1-G730D* behaves as if it is incompatible with Rpb9[66]. Recent work from the Peterson lab strongly supports a model where Rpb9 normally coordinates a loop of Rpb1 – the "anchor loop" – to appropriately interact with the TL[66]. When Rpb9 is removed, anchor loop-TL interactions are disrupted, and the open conformation of the TL is destabilized. In *rpb1-G730D,* structural perturbations are proposed to alter Rpb9-Rpb1 interactions such that they interfere with the TL, therefore *rpb1-G730D* is incompatible with Rpb9. Removal of Rpb9 or alteration of specific Rpb9 residues that organize the Rpb1 anchor loop relieve the incompatibility between *rpb1-G730D* and the TL. Third, we previously identified *rpb1-S713P*, a substitution just proximal to the anchor loop (between α-20 and α-21), as conferring gene expression, genetic interaction, and initiation phenotypes indistinguishable from GOF TL mutants[19]. Here we show that *rpb1-S713P* also confers increased biochemical activity, similar to both TL GOF alleles and anchor loop GOF alleles. We propose that *rpb1-S713P*, through constraints of the proline on structure, alters the anchor loop and therefore TL dynamics. It is conceivable, given that the secondary channel and funnel helices are accessible to factors, factor binding might also be communicated to the TL from distal sites. In addition to the three previously identified mutants, we utilized a new set of TL

mutants to assess genetic interactions between the TL and the funnel helix α-21, and discover epistasis between K1093M (TL) and phenotypically inert E712A/R (funnel helix), along with multiple allele specific genetic interactions (**Figure 2-10D**). We have suggested a more relaxed control mechanism in the Pol I compared to Pol II (**Figure 2-16C**). Taken together, funnel helices may serve as a regulatory hotspot for direct or allosteric control of the Pol II active site through the TL. While structurally conserved, evolutionary diversification of sequence may allow distinct interactions with the TL in different msRNAPs.

The characterization of the unexpectedly healthy H1085L variant clouds the issue of how H1085 functions in substrate selection and catalysis. H1085 interacts with the substrate NTP through salt bridge and hydrogen bond[24], and previous simulations with limiting H1085 variants predicted the hydrogen bonding to be critical for maintaining substrate interaction[30]. The discovery of H1085L argues that productive substrate interactions may be supported by entirely different chemistry, although we cannot rule out the possibility that H1085L redirects substrate interactions to an alternative residue. Furthermore, H1085 variants may have multiple defects in NAC, such as substrate selection[28], catalysis[28,52], intrinsic cleavage[52] and PPi release[36,37], and whether or not H1085 or analogous residues act as a general acid remains controversial in different RNAPs[24,46,51,52,229]. Function of H1085L in all of these steps remains to be determined, but the H1085L phenotype suggests that function of H1085 as a general acid may be entirely bypassed.

The established TL phenotypic landscape can be further explored to study intra- and inter-TL epistasis. First, whether individual TL residues work collaboratively or independently to ensure balanced TL dynamics and proper function is an open question. Some TL residues can be functionally overlapping and act at similar steps, or functionally discrete, acting at distinct steps. For example, combination of LOF mutations in Q1078, N1082 and a TL-proximal residue N479 resulted in non-additive genetic interaction, suggesting functionally overlapping roles for these residues. In contrast, combination of variants from Q1078 (or N1082) and H1085 resulted in exacerbation or synthetic lethality, suggesting independent functions[31]. Coupled with structures of partially folded TL states, these genetic studies support the functional distinction between NIR residues and a multi-step TL folding model for the promotion of catalysis[31]. Here, we have identified many more predicted GOF and LOF TL variants (**Figure 2-9B**), some of which are predicted to confer epistatic interactions (*e.g.* F1086 and V1089). We expect the phenotypic landscape of a multiply-substituted TL library to be extremely informative for understanding functional relationship between TL residues.

Second, the TL phenotypic landscape is an extremely sensitive readout for assessing active site re-arrangement. Transcription is under control by many factors, some of which may alter the Pol II active site conformations, though few studies directly address these possibilities. Initiation factors and Pol II TL mutants confer similar alterations in transcription start site selection, consistent with initiation factors functioning through the

Pol II active site and altering the efficiency of Pol II catalysis during initiation[19,52,264]. Furthermore, TL may communicate with other Pol II sites, such as the RNA exit channel or clamp domain[63], or in direct competition with external factors, such as TFIIS[67]. Perturbations of this communication may alter TL dynamics and cause allele-specific genetic interactions (**Figure 2-10 and 2-14**). Specifically, an external perturbation by a relevant factor or Pol II TL distant domain may show epistasis or synergy only with specific TL alleles of a class (either LOF or GOF), whereas a non-interacting factor may not. Finally, similar perturbation of the TL phenotypic landscape by different factors would suggest functional similarity between them, thus clustering of phenotypic landscape changes upon different perturbations is expected to provide valuable insight.

The TL phenotypic landscape, along with our previous work[31], illustrates a strategy of utilizing *in vivo* genetic reporters or stress response profiles to distinguish a large number of mutants with distinct *in vivo* defects. As discussed above, the phenotypic landscape sheds light on functional contribution of TL residues to its dynamics, to the mechanism of catalysis and to the evolutionary constraints of the TL sequence and function. The phenotypic landscape strategy expands the current scope of existing deep mutational scanning studies[235-238], and can be generalized to study most, if not all, of the yeast proteins.

Materials and methods

Yeast strains, media and plasmids

All yeast strains are derived from a *GAL2*$^{+}$ derivative of S288C[265]. Genotypes and phenotypes of the yeast strains are listed in **Table 2-2** and as reported[266]. Standard yeast media and the media for assessing established transcription-related phenotypes are as described previously[31]. For studies with 15 mM caffeine (Sigma), 150 mM hydroxyurea (Sigma), 5 mM and 15 mM $MnCl_2$ (Sigma), 0.5 M NaCl (EMD), 3% formamide (JT Baker), 6% ethanol (KOPTEC), 0.07 µg/mL cycloheximide (Sigma), 10 mM HCl (EMD), 10 mM NaOH (EMD), 10 µg/mL benomyl (Sigma), each compound was added to the minimal SC-Leucine (SC-Leu) medium at the indicated concentration from concentrated stock solutions.

Table 2-2. Yeast strains used in chapter II

Strain Number	Genotype	Mating Type
CKY283	***MATa ura3-52 his3Δ200 leu2Δ1 or Δ0 trp1Δ63 met15Δ0 lys2-128∂ gal10Δ56 rpb1Δ::CLONATMX RPB3::TAP::KlacTRP1 [pRP112 RPB1 URA3 CEN]***	**a**
CKY1343	***ura3-52 his3Δ200 leu2Δ1 or Δ0 trp1Δ63 met15Δ0 lys2-128∂ gal10Δ56 rpb1Δ::CLONATMX RPB3::TAP::KlacTRP1 pmr1Δ::hphmx [pRP112 RPB1 CEN URA3]***	**a**

Detailed description of plasmids is in **Table 2-3**, and complete sequences of plasmids are available upon request. For studies involving individual analyses of Pol II mutants, site-directed mutagenesis was performed via the Quickchange strategy from Stratagene. All mutagenized regions have been verified by sequencing before sub-cloning into pRS315-derived plasmids, as previously described[31].

Genetic and biochemical analyses of individual Pol II mutants

Phenotypic analyses of individual Pol II mutants were performed by plasmid shuffling assays, with viable mutants further subjected to standard plate phenotyping. Each mutant in a pRS315-derived plasmid (*CEN LEU2*) was transformed into CKY283 (*rpb1*Δ*::CLONATMX*, pRP112 *RPB1 CEN URA3*). Transformants (Leu^+) were patched on SC-Leu plates and subsequently replica plated to SC-Leu+5FOA (1mg/mL) to assay complementation ability upon loss of the *RPB1 CEN URA3* plasmid. Experimental details are as previously described[28,31]. Saturated cultures from single colonies of viable and shuffled Pol II mutants were subject to 10-fold serial dilution and spotting on indicated phenotyping media, as described in various previous reports[28,31].

Table 2-3. Plasmids used in chapter II

Mutant	Plasmid number	Substituting Domain	Viability	Derivation of the mutants
WT (URA3)	pCK250 (pRP112)	N.A.	Viable	N.A.
WT	pCK859	N.A.	Viable	N.A.
TLΔRpb1	pCK892	TL	Inviable	N.A.
A1076T	pCK968	TL	Viable	Sub-cloned from previous screening
Q1078A	pCK935	TL	Inviable	Site-directed mutagenesis
Q1078S	pCK863	TL	Viable	Site-directed mutagenesis
M1079R	pCK872	TL	Viable	Site-directed mutagenesis
N1082A	pCK885	TL	Inviable	Site-directed mutagenesis
N1082S	pCK886	TL	Viable	Site-directed mutagenesis
F1084I	pCK955	TL	Viable	Sub-cloned from previous screening
H1085A	pCK861	TL	Inviable	Site-directed mutagenesis
H1085Q	pCK887	TL	Viable	Site-directed mutagenesis
H1085Y	pCK870	TL	Viable	Site-directed mutagenesis
F1086S	pCK871	TL	Viable	Sub-cloned from previous screening
G1097D (with silent mutation)	pCK867	TL	Viable	Sub-cloned from previous screening
L1101S	pCK864	TL	Viable	Sub-cloned from previous screening
E1103G	pCK960	TL	Viable	Sub-cloned from previous screening
Q1078A/E1103G	pCK947	TL	Viable	Site-directed mutagenesis
N1082A/E1103G	pCK897	TL	Viable	Site-directed mutagenesis
F1084I/E1103G	pCK952	TL	Inviable	Site-directed mutagenesis
H1085A/E1103G	pCK899	TL	Viable	Site-directed mutagenesis

Table 2-3. Continued

Mutant	Plasmid number	Substituting Domain	Viability	Derivation of the mutants
S713P	pCK866	Funnel helix	Viable	Sub-cloned from previous screening
I1327V	pCK610	Other Rpb1 domains	Viable	Sub-cloned from previous screening
A1087G	pCK1730	TL	Inviable	Site-directed mutagenesis
G1088A	pCK1731	TL	Inviable	Site-directed mutagenesis
A1087G/G1088A	pCK1732	TL	Inviable	Site-directed mutagenesis
A1087V	pCK1733	TL	Viable	Site-directed mutagenesis
G1088V	pCK1734	TL	Inviable	Site-directed mutagenesis
L1081M	pCK1735	TL	Viable	Site-directed mutagenesis
L1081P	pCK1736	TL	Inviable	Site-directed mutagenesis
H1085L	pCK1748	TL	Viable	Site-directed mutagenesis
M818P	pCK1089	BH	Inviable	Site-directed mutagenesis
M818S	pCK1165	BH	Viable	Site-directed mutagenesis
M818Y	pCK1164	BH	Viable	Site-directed mutagenesis
T834A	pCK910	BH	Viable	Site-directed mutagenesis
T834P	pCK1087	BH	Viable	Site-directed mutagenesis
Y836A	pCK1899	BH	Viable	Site-directed mutagenesis
Y836H	pCK1901	BH	Viable	Site-directed mutagenesis
Y836F	pCK1903	BH	Viable	Site-directed mutagenesis
R839A	pCK1904	BH	Viable	Site-directed mutagenesis
R840A	pCK1905	BH	Viable	Site-directed mutagenesis
V842A	pCK1906	BH	Viable	Site-directed mutagenesis

Table 2-3. Continued

Mutant	Plasmid number	Substituting Domain	Viability	Derivation of the mutants
K843A	pCK1907	BH	Viable	Site-directed mutagenesis
K843N	pCK1908	BH	Viable	Site-directed mutagenesis
M818P/H1085Y	pCK1117	BH/TL	Inviable	Site-directed mutagenesis
M818P/F1086S	pCK1118	BH/TL	Inviable	Site-directed mutagenesis
M818P/F1084I	pCK1119	BH/TL	Inviable	Site-directed mutagenesis
M818P/E1103G	pCK1120	BH/TL	Inviable	Site-directed mutagenesis
M818S/Q1078A	pCK1868	BH/TL	Inviable	Site-directed mutagenesis
M818S/H1085A	pCK1869	BH/TL	Viable	Site-directed mutagenesis
M818S/N1082A	pCK1870	BH/TL	Viable	Site-directed mutagenesis
M818S/H1085Y	pCK1871	BH/TL	Viable	Site-directed mutagenesis
M818S/Q1078S	pCK1872	BH/TL	Viable	Site-directed mutagenesis
M818S/N1082S	pCK1873	BH/TL	Viable	Site-directed mutagenesis
M818S/H1085Q	pCK1874	BH/TL	Viable	Site-directed mutagenesis
M818S/Q1078A/E1103G	pCK1875	BH/TL	Viable	Site-directed mutagenesis
M818S/H1085A/E1103G	pCK1876	BH/TL	Viable	Site-directed mutagenesis
M818S/F1086S	pCK1877	BH/TL	Viable	Site-directed mutagenesis
M818S/N1082A/E1103G	pCK1878	BH/TL	Inviable	Site-directed mutagenesis
M818S/F1084I	pCK1879	BH/TL	Viable	Site-directed mutagenesis
M818S/E1103G	pCK1880	BH/TL	Viable	Site-directed mutagenesis
M818Y/Q1078A	pCK1881	BH/TL	Inviable	Site-directed mutagenesis
M818Y/H1085A	pCK1882	BH/TL	Viable	Site-directed mutagenesis

Table 2-3. Continued

Mutant	Plasmid number	Substituting Domain	Viability	Derivation of the mutants
M818Y/N1082A	pCK1883	BH/TL	Viable	Site-directed mutagenesis
M818Y/H1085Y	pCK1166	BH/TL	Viable	Site-directed mutagenesis
M818Y/Q1078S	pCK1884	BH/TL	Viable	Site-directed mutagenesis
M818Y/N1082S	pCK1885	BH/TL	Viable	Site-directed mutagenesis
M818Y/H1085Q	pCK1886	BH/TL	Viable	Site-directed mutagenesis
M818Y/Q1078A/E1103G	pCK1887	BH/TL	Viable	Site-directed mutagenesis
M818Y/H1085A/E1103G	pCK1888	BH/TL	Viable	Site-directed mutagenesis
M818Y/F1086S	pCK1167	BH/TL	Viable	Site-directed mutagenesis
M818Y/N1082A/E1103G	pCK1889	BH/TL	Viable	Site-directed mutagenesis
M818Y/F1084I	pCK1169	BH/TL	Viable	Site-directed mutagenesis
M818Y/E1103G	pCK1168	BH/TL	Viable	Site-directed mutagenesis
T834A/Q1078A	pCK1403	BH/TL	Inviable	Site-directed mutagenesis
T834A/H1085A	pCK1399	BH/TL	Inviable	Site-directed mutagenesis
T834A/N1082A	pCK1411	BH/TL	Inviable	Site-directed mutagenesis
T834A/H1085Y	pCK1408	BH/TL	Inviable	Site-directed mutagenesis
T834A/Q1078S	pCK1400	BH/TL	Inviable	Site-directed mutagenesis
T834A/N1082S	pCK1401	BH/TL	Inviable	Site-directed mutagenesis
T834A/H1085Q	pCK1402	BH/TL	Inviable	Site-directed mutagenesis
T834A/Q1078A/E1103G	pCK1406	BH/TL	Inviable	Site-directed mutagenesis
T834A/H1085A/E1103G	pCK1405	BH/TL	Inviable	Site-directed mutagenesis
T834A/F1086S	pCK1407	BH/TL	Inviable	Site-directed mutagenesis

Table 2-3. Continued

Mutant	Plasmid number	Substituting Domain	Viability	Derivation of the mutants
T834A/N1082A/E1103G	pCK1404	BH/TL	Viable	Site-directed mutagenesis
T834A/F1084I	pCK1409	BH/TL	Viable	Site-directed mutagenesis
T834A/E1103G	pCK1410	BH/TL	Viable	Site-directed mutagenesis
T834A/F1084I/E1103G	pCK1112	BH/TL	Inviable	Site-directed mutagenesis
T834P/Q1078A	pCK1890	BH/TL	Inviable	Site-directed mutagenesis
T834P/H1085A	pCK1891	BH/TL	Inviable	Site-directed mutagenesis
T834P/N1082A	pCK1892	BH/TL	Inviable	Site-directed mutagenesis
T834P/H1085Y	pCK1113	BH/TL	Viable	Site-directed mutagenesis
T834P/Q1078S	pCK1893	BH/TL	Viable	Site-directed mutagenesis
T834P/N1082S	pCK1894	BH/TL	Viable	Site-directed mutagenesis
T834P/H1085Q	pCK1895	BH/TL	Viable	Site-directed mutagenesis
T834P/Q1078A/E1103G	pCK1896	BH/TL	Viable	Site-directed mutagenesis
T834P/H1085A/E1103G	pCK1897	BH/TL	Inviable	Site-directed mutagenesis
T834P/F1086S	pCK1114	BH/TL	Viable	Site-directed mutagenesis
T834P/N1082A/E1103G	pCK1898	BH/TL	Inviable	Site-directed mutagenesis
T834P/F1084I	pCK1115	BH/TL	Inviable	Site-directed mutagenesis
T834P/E1103G	pCK1116	BH/TL	Inviable	Site-directed mutagenesis
T834P/Y836A	pCK1900	BH	Viable	Site-directed mutagenesis
T834P/Y836H	pCK1902	BH	Viable	Site-directed mutagenesis
G1097E	pCK1737	TL	Viable	Site-directed mutagenesis
G1097D	pCK1738	TL	Viable	Site-directed mutagenesis

Table 2-3. Continued

Mutant	Plasmid number	Substituting Domain	Viability	Derivation of the mutants
S1091A	pCK1749	TL	Viable	Site-directed mutagenesis
S1091E	pCK1756	TL	Viable	Site-directed mutagenesis
S1091C	pCK1750	TL	Viable	Site-directed mutagenesis
K1092A	pCK1751	TL	Viable	Site-directed mutagenesis
K1092D	pCK1752	TL	Viable	Site-directed mutagenesis
K1093M	pCK1755	TL	Viable	Site-directed mutagenesis
D716A	pCK1747	Funnel helix	Viable	Site-directed mutagenesis
D716K	pCK2221	Funnel helix	Viable	Site-directed mutagenesis
E712A	pCK1740	Funnel helix	Viable	Site-directed mutagenesis
E712R	pCK1741	Funnel helix	Viable	Site-directed mutagenesis
E1307A	pCK1742	TL tip proximal	Viable	Site-directed mutagenesis
E1307S	pCK1743	TL tip proximal	Viable	Site-directed mutagenesis
E1307K	pCK1744	TL tip proximal	Viable	Site-directed mutagenesis
R1281A	pCK1745	TL tip proximal	Viable	Site-directed mutagenesis
R1281E	pCK1746	TL tip proximal	Viable	Site-directed mutagenesis
D716A/S1091A	pCK1757	Funnel helix/TL	Viable	Site-directed mutagenesis
D716A/S1091E	pCK1759	Funnel helix/TL	Viable	Site-directed mutagenesis
D716A/S1091C	pCK1761	Funnel helix/TL	Viable	Site-directed mutagenesis
D716A/K1092A	pCK1769	Funnel helix/TL	Viable	Site-directed mutagenesis
D716A/K1092D	pCK1777	Funnel helix/TL	Viable	Site-directed mutagenesis
D716A/K1093M	pCK1785	Funnel helix/TL	Viable	Site-directed mutagenesis

Table 2-3. Continued

Mutant	Plasmid number	Substituting Domain	Viability	Derivation of the mutants
D716K/S1091A	pCK2224	Funnel helix/TL	Viable	Site-directed mutagenesis
D716K/S1091E	pCK2239	Funnel helix/TL	Viable	Site-directed mutagenesis
D716K/S1091C	pCK2225	Funnel helix/TL	Viable	Site-directed mutagenesis
D716K/K1092A	pCK2226	Funnel helix/TL	Viable	Site-directed mutagenesis
D716K/K1092D	pCK2227	Funnel helix/TL	Viable	Site-directed mutagenesis
D716K/K1093M	pCK2228	Funnel helix/TL	Viable	Site-directed mutagenesis
E712A/K1092A	pCK1771	Funnel helix/TL	Viable	Site-directed mutagenesis
E712A/K1092D	pCK1779	Funnel helix/TL	Viable	Site-directed mutagenesis
E712A/K1093M	pCK1787	Funnel helix/TL	Viable	Site-directed mutagenesis
E712R/K1092A	pCK1772	Funnel helix/TL	Viable	Site-directed mutagenesis
E712R/K1092D	pCK1780	Funnel helix/TL	Viable	Site-directed mutagenesis
E712R/K1093M	pCK1788	Funnel helix/TL	Viable	Site-directed mutagenesis
E1307A/K1092A	pCK1775	TL tip proximal/TL	Viable	Site-directed mutagenesis
E1307A/K1092D	pCK1783	TL tip proximal/TL	Viable	Site-directed mutagenesis
E1307A/K1093M	pCK1791	TL tip proximal/TL	Viable	Site-directed mutagenesis
E1307K/K1092A	pCK1776	TL tip proximal/TL	Viable	Site-directed mutagenesis
E1307K/K1092D	pCK1784	TL tip proximal/TL	Viable	Site-directed mutagenesis
E1307K/K1093M	pCK1792	TL tip proximal/TL	Inviable	Site-directed mutagenesis
R1281A/K1092A	pCK1773	TL tip proximal/TL	Viable	Site-directed mutagenesis
R1281A/K1092D	pCK1781	TL tip proximal/TL	Viable	Site-directed mutagenesis

Table 2-3. Continued

Mutant	Plasmid number	Substituting Domain	Viability	Derivation of the mutants
R1281A/K1093M	pCK1789	TL tip proximal/TL	Viable	Site-directed mutagenesis
R1281E/K1092A	pCK1774	TL tip proximal/TL	Viable	Site-directed mutagenesis
R1281E/K1092D	pCK1782	TL tip proximal/TL	Viable	Site-directed mutagenesis
R1281E/K1093M	pCK1790	TL tip proximal/TL	Viable	Site-directed mutagenesis

Pol II enzymes were purified via a tandem-affinity tag (TAP) protocol derived from [267] with modifications described in [28]. Transcription elongation reactions were performed with Pol II elongation complexes assembled on a nucleic acid scaffold, in a procedure described in [28] with slight modifications in the amount of Pol II and nucleic acids as described in [42]. For each enzyme, elongation assays were performed with 25 μM, 125 μM, 500 μM and 750 μM NTPs (each of ATP, GTP, CTP, UTP), and maximal elongation rates were extracted exactly as previously described[28].

ADH1 transcription start site selection was analyzed by primer extension. In brief, indicated strains were grown in YPD until mid-log phase (~1×10^7 cells/mL), and diluted with YPD with 10mM $MnCl_2$ or equal volume of H_2O. Total RNA was extracted as described[268], and 30 μg of total RNA was subject to primer extension analysis, following a protocol derived from [269] with modifications described in [31].

High-throughput phenotypic analyses of the TL variants library

The TL variant library was synthesized by Sloning Biotechnology (now MorphoSys) with well-characterized TL variants excluded (specified in **Figure 2-1B**) using a building block approach[239,240]. The TL variant library was transformed into CKY283 via a gap-repair strategy as previously described[19]. In brief, the amplified TL variant library with flanking sequence was transformed into CKY283 together with a linearized pRS315-derived plasmid (*CEN LEU2*) containing *rpb1* deleted for the TL (*TL*Δ) and linearized at the deletion junction, allowing *in vivo* homologous recombination.

Homologous recombination produces a library of complete *rpb1* genes containing TL variants. The gap-repaired TL variants (Leu^+) were titered and plated at 200-300 colonies per plate to reduce inter-colony growth competition, and Leu^+ colonies were first replica-plated to SC-Leu+5FOA (1mg/mL), and subsequently to additional selective and control media. Three independent biological replicate screens were performed. In each replicate, we pooled 6000 to 12000 colonies. Each cell pool was subjected to genomic DNA extraction and TL amplification by emulsion PCR. Amplification of the TL region was performed using Micellula DNA Emulsion & Purification (ePCR) Kit (Chimerx) per manufacturer's instructions. To minimize amplification bias, each sample was amplified in a 15-cycle ePCR reaction, purified and subject to additional 13-15 cycle scale-up ePCR reactions. The two-step ePCR amplification protocol ensured sufficient yield of DNA for NGS sequencing while minimizing perturbation of the allele distribution in the DNA pool. The amplified samples were subject to Illumina HiSeq 2500 sequencing, and on average over 2 million reads were obtained from each replicate of a sample, with high reproducibility and minimal perturbation of the mutant distribution within the TL variant library (**Figure 2-2D**).

Allele frequency was subsequently measured by deep sequencing of the TL amplicons. All the sequencing data (FASTQ format) for the reported analyses are deposited and available under the NCBI bioproject PRJNA340979. To identify the mutations that were present for each set of paired-end reads, a codon-based alignment algorithm was developed to align each paired-end read set in which the overlapping substrings from

both flanking regions agreed perfectly to the WT sequence. The purpose of our approach was to identify real variants using an expected set of mutant codons used in the programmed library synthesis from sequencing errors. A dynamic programming algorithm was applied so that an exact match of three letters was assigned a positive score, a mismatch of at least one letter in a codon was assigned a negative score, and the insertion or deletion of either one, two or three letters was assigned a constant negative score. The allele frequency was subsequently calculated from the mapped reads, and the phenotypic score of each TL variant was calculated by allele frequency change (normalized to WT) under each condition, as below:

$$f = \log\left[\frac{f^{mut,sele}}{f^{mut,unsele}}\right] - \log\left[\frac{f^{wt,sele}}{f^{wt,unsele}}\right]$$

Mutants with less than 200 reads in the transformed pool (SC-Leu) and allele frequency changes assessed from less than 50 reads from both conditions were excluded from further analyses. Median values from three independent biological replicates were used for fitness and phenotype scoring. Fitness score cutoff for lethality was estimated based on fitness scores (on SC-Leu and 5FOA) of 163 known viable TL and 16 known lethal mutants. Hierarchical clustering for generating phenotypic landscape was performed by Gene Cluster 3.0 using centered correlation[270]. Figures displaying structural information were generated using Pymol (https://www.pymol.org/).

Evolutionary analyses

Eukaryotic RNA polymerase large subunit sequences were obtained from BLAST using *Sce* Rpb1 (Pol II), *Sce* Rpa190 (Pol I), and *Sce* Rpo31 (Pol III) sequences as queries. Sequences were assigned to Pol I, II, or III based on highest similarity when compared to each of the three query sequences, with prokaryotic sequences further filtered out. Multiple sequence alignments (MSAs) were generated by first applying CD-HIT[271] to cluster sequences so that the identity between sequences in different clusters was less than 90%, then applying MUSCLE[272] to obtain an alignment that contains one representative sequence from each cluster. The TL conservation score was generated using Jalview 2.8 version 14.0[247] and plotted as a heatmap using Gene-E (http://www.broadinstitute.org/cancer/software/GENE-E/index.html).

CHAPTER III

MODE OF ACTION OF THIOLUTIN

Overview

Thiolutin is a well-known and routinely used transcription inhibitor with an unresolved mode of action. Recent studies have identified Zn^{2+}-chelation activity for both thiolutin and the related dithiolopyrrolone holomycin, and have appeared to rule out direct inhibition of RNA polymerases as a mode of action. However, negative results for direct transcription inhibition of RNA polymerases in this recent work contradicts previously observed thiolutin inhibition of yeast RNA polymerases I, II, and III *in vitro*. Here, we present chemical genetic and biochemical approaches to investigate the mode of action of thiolutin. We identify diverse classes of mutants that are resistant or sensitive to thiolutin. We functionally dissect the multidrug resistance and thioredoxin pathways controlling thiolutin sensitivity. We provide evidence that thiolutin causes oxidation of thioredoxins *in vivo*, and suggest that thiolutin induces oxidative stress and alters Zn^{2+} and Cu^{2+} homeostasis *in vivo*, recapitulating thiolutin interactions with Zn^{2+}. Finally, our results resolve contradictory biochemical results for thiolutin inhibition of transcription by direct demonstration of thiolutin inhibition of RNA polymerase II (Pol II) *in vitro*. Inhibitory activity requires both appropriate reduction of thiolutin and the presence of Mn^{2+}. Thio/Mn^{2+} inhibition is abrogated when template DNA is pre-bound to Pol II or when excess DTT is present, and renders Pol II pause prone in elongation if initiation is bypassed using a pre-melted template and RNA primer. Together, we propose that

thiolutin directly inhibits Pol II transcription through a novel mechanism distinct from known transcription inhibitors.

Introduction

Thiolutin is a commonly used transcription inhibitor with possibly multiple modes of action. Thiolutin and holomycin are structurally highly similar dithiolopyrrolone compounds, which feature an intra-molecular and redox-sensitive disulfide bond[210,212]. Thiolutin has been shown to inhibit bacterial and yeast transcription *in vivo*, and has been used to study mRNA stability in multiple species[108-111]. However, the mode of action of thiolutin for transcription inhibition remains unclear and complicated (discussed below). In addition, thiolutin appears to affect multiple cellular pathways, including glucose metabolism, mRNA degradation, oxidative stress response, and proteasome activity[214,223,227,273].

Recent progress has suggested that thiolutin and holomycin appear to be a redox sensitive Zn^{2+} chelator, thus explaining diverse cellular effects[212-214]. It has been demonstrated *in vitro* that thiolutin and holomycin can be reduced by strong reductants DTT or TCEP[212,214], and reduced thiolutin and holomycin can chelate Zn^{2+} [213,214]. The Zn^{2+} chelating activity explains multiple thiolutin induced phenotypes, including alteration of glucose metabolism and proteasome inhibition through Zn^{2+} specific metalloproteins[214,227]. In addition, two observations suggest a possible mechanism for thiolutin-induced expression of oxidative stress response genes[273]. First, thiolutin and

holomycin appear to be only active when they are reduced, and reduction of thiolutin could in turn oxidize the cellular thiolutin reductant or reducing proteins[212,213]. Second, reduced holomycin can be spontaneously oxidized by molecular oxygen, suggesting that holomycin may be a redox cycling compound[225,226] and cause accumulation of ROS. Whether thiolutin can act as a redox cycler has not yet been tested, and the cellular reductant for thiolutin or holomycin remain unclear, because glutathione does not reduce holomycin *in vitro*[212]. Together, the Zn^{2+} chelating and redox activities are consistent with the previously observed diverse thiolutin-induced phenotypes. However, both thiolutin and holomycin are inert against RNAPs under conditions where either should be able to chelate Zn^{2+}. In light of these negative results, transcription inhibition by dithiolopyrrolones was concluded to be a secondary outcome of a distinct target, such as the proteasome, which was inhibited by thiolutin through removal of zinc from Rpn11, an essential deubiquitinase[213,214].

It seems clear that thiolutin inhibits transcription *in vivo* in multiple species[217,221,222], though from *in vitro* studies it has been concluded that it inhibits prokaryotic and eukaryotic RNAPs differently. Thiolutin inhibited all three partially purified yeast RNAPs *in vitro*[216], but failed to inhibit all tested prokaryotic RNAPs[220-222]. Similar RNAP inhibition *in vivo* with lack of inhibition *in vitro* was also observed for holomycin[164,213]. In addition, the recently reported lack of thiolutin inhibition for all three fully purified yeast RNAPs adds further confusion[214]. Together, thiolutin and holomycin do not appear to inhibit prokaryotic or fully purified eukaryotic RNAPs,

suggesting the *in vivo* transcription inhibition could be indirect result of other thiolutin or holomycin activities. However, without comprehensive assessment of the experimental parameters for different transcription assays, it seems difficult to rule out direct mechanisms of action for thiolutin on eukaryotic RNAPs, as recent work does not satisfactorily explain why thiolutin inhibited all three partially purified yeast RNA polymerases *in vitro*[216].

The thiolutin inhibition of partially purified yeast RNA polymerases *in vitro* appears to be specific and is not satisfactorily explained by Zn^{2+} chelation[216]. Thiolutin inhibited RNAPs prior to template addition, but did not inhibit the template DNA bound RNAPs[216]. This critical order-of-addition requirement for treatment prior to DNA addition is reminiscent of RNAP switch region inhibitors' behavior (reviewed in Chapter I)[174,187] and is consistent with inhibition of a specific and early step in transcription.

We imagine a number of possibilities to reconcile the existing data. First, thiolutin may target a secondary protein, which is present *in vivo* and in partially purified RNA polymerase fractions[216] but not others[214,220,221]. Second, thiolutin may act as a prodrug that is activated *in vivo* and in the early report[216] but not in assays used by others[214,220,221]. To further investigate the mode of action of thiolutin, we have undertaken multiple chemical genetics and genomics approaches to screen for thiolutin resistant and sensitive mutants, as similar studies have contributed to understanding of modes of actions of many compounds[274-276], including holomycin[213]. We identify drug efflux pumps

functioning in thiolutin resistance, and demonstrate that thiolutin treatment induces oxidative stress. We propose that thiolutin likely induces the oxidative stress via oxidizing thioredoxins and possibly also redox-cycling. In addition, we confirm that reduced thiolutin directly interacts with Zn^{2+} *in vitro*, consistent with two very recent reports on holomycin and thiolutin using different approaches[213,214]. We also suggest that thiolutin alters cellular Zn^{2+} homeostasis in yeast, similar to holomycin's effects in *E.coli*[213].

Finally, we have also revisited *in vitro* biochemical studies with thiolutin and Pol II. We find that thiolutin indeed inhibits Pol II but critically requires the presence of small amounts of Mn^{2+}, providing an immediate explanation for disparate results across studies. First, DNA bound Pol II is resistant to thiolutin inhibition, recapitulating Tipper et al.'s early studies[216]. Second, though thiolutin appears to inhibit a very early step in Pol II transcription, when initiation is bypassed by use of specific nucleic acid templates, thiolutin-treated Pol II exhibits pause-prone, slow elongation. Third, high DTT reverses the thiolutin inhibition, suggesting involvement of disulfide bond or redox chemistry in inhibition. We find that thiolutin appears to block initiation *in vivo* immediately after treatment, consistent with Pol II inhibition *in vivo* being direct. We propose that thiolutin inhibits Pol II transcription through a novel mechanism distinct from most of known transcription inhibitors.

Material and methods

Yeast strains and reagents

Yeast strains used are listed in **Table 3-1**. *YAP1* C-terminal tagging with EGFP and gene deletion strains were constructed as described[277,278]. Phenotyping of mutants using plate assays were performed as previously described[31,266]. Chemicals were commercially obtained from the following: Cayman Chemical (Thiolutin), Gold Biotechnology (DTT, TCEP, 5FOA), Sigma ($MnCl_2$), JT Baker($ZnCl_2$), BDH Chemicals ($MgCl_2$).

Variomics screens

Two separate pools of diploid Variomics libraries were gifts from Dr. Xuewen Pan: one for non-essential and the other for essential genes[279]. The pooled diploid Variomics libraries were grown, sporulated and selected for haploid Variomics libraries as described[279]. Haploid Variomics libraries were subsequently used for the genetic screens. To convert the Variomics libraries into plasmid-free deletion libraries, pooled Variomics libraries (diploid for essential genes and haploid for non-essential genes) were grown in liquid SC media and plated on SC+5FOA plates to select against the mutants with the *URA3* plasmids containing the Variomics mutations. The diversity of the libraries were confirmed by deep sequencing of barcodes ("uptags" and "downtags") flanking the deletion cassettes.

Table 3-1. Yeast strains used in chapter III

Strain number	Genotype	Mating Type	Comments
CKY457	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂*	a	Wild type strain
CKY767	*his3Δ1 leu2Δ0 met15Δ0 ura3Δ0*	a	BY4741
CKY769	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0*	a/alpha	BY4743
CKY1132	*ura3-52 his3Δ200 leu2Δ1 or Δ0 trp1Δ63 met15Δ0 lys2-128∂ rpb1Δ::CLONATMX RPB3::3XFLAG::kanMX kanMX::TEFp::YLR454w*	alpha	*YLR454W* promoter was replaced with *TEF1* promoter. *RPB3* was 3XFLAG tagged. For *in vivo* transcription shutoff and Pol II chromatin IP
CKY1327	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ yap1Δ::hphNT1*	a	CKY457 with *yap1Δ*, for testing thiolutin sensitivity.
CKY1328	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ yap1Δ::hphNT1*	a	CKY457 with *yap1Δ*, for testing thiolutin sensitivity.
CKY1329	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ yrr1Δ::kanMX*	a	CKY457 with *yrr1Δ*, for testing thiolutin sensitivity.
CKY1330	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ yrr1Δ::kanMX*	a	CKY457 with *yrr1Δ*, for testing thiolutin sensitivity.
CKY1342	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂*	alpha	CKY457 with mating type switched to alpha via *GAL-HO* plasmid (plasmid removed). For bulk-segregant analysis.
CKY1419	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ pdr1Δ::natNT2*	a	CKY457 with *pdr1Δ*, for testing thiolutin sensitivity.
CKY1420	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ pdr1Δ::natNT2*	a	CKY457 with *pdr1Δ*, for testing thiolutin sensitivity.
CKY1604	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ snq2Δ::KlacTRP1*	a	CKY457 with *snq2*Δ, for testing thiolutin sensitivity.
CKY1605	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ snq2Δ::KlacTRP1*	a	CKY457 with *snq2*Δ, for testing thiolutin sensitivity.
CKY1608	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ trr2Δ::kanMX*	a	CKY457 with *trr2*Δ, for testing thiolutin sensitivity.
CKY1609	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ trr2Δ::kanMX*	a	CKY457 with *trr2*Δ, for testing thiolutin sensitivity.

Table 3-1. Continued

Strain number	Genotype	Mating Type	Comments
CKY1610	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ trx1Δ::natNT2*	a	CKY457 with *trx1*Δ, for testing thiolutin sensitivity.
CKY1611	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ trx1Δ::natNT2*	a	CKY457 with *trx1*Δ, for testing thiolutin sensitivity.
CKY1612	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ trx2Δ::kanMX*	a	CKY457 with *trx2*Δ, for testing thiolutin sensitivity.
CKY1613	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ trx2Δ::kanMX*	a	CKY457 with *trx2*Δ, for testing thiolutin sensitivity.
CKY1650	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ trr1Δ::hphNT1*	a	CKY457 with *trr1*Δ, for testing thiolutin sensitivity.
CKY1651	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ trr1Δ::hphNT1*	a	CKY457 with *trr1*Δ, for testing thiolutin sensitivity.
CKY1932	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ tsa1Δ::kanMX*	a	CKY457 with *tsa1*Δ, for testing thiolutin sensitivity.
CKY1933	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ tsa1Δ::kanMX*	a	CKY457 with *tsa1*Δ, for testing thiolutin sensitivity.
CKY1934	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ zap1Δ::natNT2*	a	CKY457 with *zap1*Δ, for testing thiolutin sensitivity.
CKY1935	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ zap1Δ::natNT2*	a	CKY457 with *zap1*Δ, for testing thiolutin sensitivity.
CKY1984	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ trx2Δ::kanMX trx1Δ::natNT2*	a	CKY457 with *trx1Δtrx2*Δ, for testing thiolutin sensitivity.
CKY1985	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ trx2Δ::kanMX trx1Δ::natNT2*	a	CKY457 with *trx1Δtrx2*Δ, for testing thiolutin sensitivity.
CKY1986	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ sod1Δ::hphNT1*	a	CKY457 with *sod1*Δ, for testing thiolutin sensitivity.
CKY1987	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ sod1Δ::hphNT1*	a	CKY457 with *sod1*Δ, for testing thiolutin sensitivity.
CKY2038	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ yap1::EGFP::kanMX*	a	CKY457 with Yap1 tagged with EGFP, for visualizing Yap1 cellular localization.
CKY2040	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ pdr5Δ::natNT2*	a	CKY457 with *pdr5*Δ, for testing thiolutin sensitivity.
CKY2041	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ pdr5Δ::natNT2*	a	CKY457 with *pdr5*Δ, for testing thiolutin sensitivity.

Table 3-1. Continued

Strain number	Genotype	Mating Type	Comments
CKY2046	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ trx2Δ::kanMX trx1Δ::natNT2 trr1Δ::hphNT1*	a	CKY457 with *trr1Δtrx1Δtrx2Δ*, for testing thiolutin sensitivity.
CKY2047	*leu2Δ1 ura3-52 trp1Δ63 his3Δ200 lys2-128∂ trx2Δ::kanMX trx1Δ::natNT2 trr1Δ::hphNT1*	a	CKY457 with *trr1Δtrx1Δtrx2Δ*, for testing thiolutin sensitivity.
CKY2132	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 erv1Δ::hphNT1/ERV1*	a/alpha	BY4743 with *erv1Δ*, for validating bar-seq results
CKY2133	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 erv1Δ::hphNT1/ERV1*	a/alpha	BY4743 with *erv1Δ*, for validating bar-seq results
CKY2134	*his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 gcn5Δ::hphNT1*	a	BY4741 with *gcn5Δ*, for validating bar-seq results
CKY2135	*his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 gcn5Δ::hphNT1*	a	BY4741 with *gcn5Δ*, for validating bar-seq results
CKY2136	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 mia40Δ::hphNT1/MIA40*	a/alpha	BY4743 with *mia40Δ*, for validating bar-seq results
CKY2137	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 mia40Δ::hphNT1/MIA40*	a/alpha	BY4743 with *mia40Δ*, for validating bar-seq results
CKY2138	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 pre10Δ::hphNT1/PRE10*	a/alpha	BY4743 with *pre10Δ*, for validating bar-seq results
CKY2139	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 pre10Δ::hphNT1/PRE10*	a/alpha	BY4743 with *pre10Δ*, for validating bar-seq results
CKY2140	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 pup1Δ::hphNT1/PUP1*	a/alpha	BY4743 with *pup1Δ*, for validating bar-seq results

Table 3-1. Continued

Strain number	Genotype	Mating Type	Comments
CKY2141	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 pup1Δ::hphNT1/PUP1*	a/alpha	BY4743 with *pup1*Δ, for validating bar-seq results
CKY2142	*his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ssn3Δ::natNT2*	a	BY4741 with *ssn3*Δ, for validating bar-seq results
CKY2143	*his3Δ1 leu2Δ0 met15Δ0 ura3Δ0 ssn3Δ::natNT2*	a	BY4741 with *ssn3*Δ, for validating bar-seq results
CKY2144	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 rpn5Δ::hphNT1/RPN5*	a/alpha	BY4743 with *rpn5*Δ, for validating bar-seq results
CKY2145	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 rpn5Δ::hphNT1/RPN5*	a/alpha	BY4743 with *rpn5*Δ, for validating bar-seq results
CKY2146	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 rrp4Δ::hphNT1/RRP4*	a/alpha	BY4743 with *rrp4*Δ, for validating bar-seq results
CKY2147	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 rrp4Δ::hphNT1/RRP4*	a/alpha	BY4743 with *rrp4*Δ, for validating bar-seq results
CKY2148	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 rrp46Δ::hphNT1/RRP46*	a/alpha	BY4743 with *rrp46*Δ, for validating bar-seq results
CKY2149	*his3Δ1/his3Δ1 leu2Δ0/leu2Δ0 LYS2/lys2Δ0 met15Δ0/MET15 ura3Δ0/ura3Δ0 rrp46Δ::hphNT1/RRP46*	a/alpha	BY4743 with *rrp46*Δ, for validating bar-seq results

For manual screening of variomics libraries, about 1-3×10^7 cells were plated on SC+10μg/mL thiolutin plates, and the potential thiolutin resistant candidates were restruck out on SC+10μg/mL thiolutin plates to validate the resistance. The plasmids from the validated resistant candidates were subsequently recovered for Sanger sequencing to identify putative resistance-conferring variants. We subsequently tested dominance for the repeatedly isolated candidates. The recovered plasmids from all reproducibly isolated candidates were transformed into the wild-type (CKY457) strain, and thiolutin sensitivity for transformants was phenotyped and compared to the empty vector control.

For high-throughput screens, Variomics libraries were screened on plates, while deletion libraries were screened in liquid, as previously described[274,275,279], and three independent screens (biological replicates) were performed. For Variomics screens, each library was screened on SC-URA+DMSO, SC-URA+8μg/mL thiolutin or SC-URA+10μg/mL plates. Each biological replicate was screened on 9 plates (6×10^7 cells/plate) for 3 days, and cells were scraped to screen for an additional set of 9 plates for 3 days (6×10^7 cells/plate). Deletion libraries were grown in liquid SC media to 3×10^7 cells/mL and diluted in SC+3μg/mL thiolutin (for haploid non-essential gene deletion library) or SC+4μg/mL thiolutin (for diploid essential gene deletion library) to grow for 20 generations. During the selection, the cells were diluted to 1×10^6 cells/mL every five generations to maintain cultures in log phase. For both screens, yeast cells were pooled, and genomic DNA was prepared using a YeaStar Genomic DNA kit (Zymo Research) for subsequent PCR

amplification of barcode regions. Amplicons were sequenced by illumina Hiseq2500 in rapid mode.

Bar-seq data processing

Sequencing reads are mapped to the re-annotated barcode sequences[274] using Bowtie2 (version 2.2.4) with the -N flag set to 0 and the --no-unal flag to suppress unaligned reads for further analyses. Bowtie2 outputs were written into SAM format and further extracted using Samtools (version 1.3.1). Barcode sequences shorter than 15nts or were mapped to multiple reference barcodes were discarded. On average, 98% of sequencing reads were uniquely mapped.

As with Robinson et al.[280], we filtered out barcodes with lower than 100 total reads across all samples, since low count barcodes across all conditions were likely from sequencing errors and did not exist. In addition, mutants with less than 20 reads in both treated and the corresponding untreated controls are further filtered out, to exclude large changes from a small amount of reads with low confidence.

Following Robinson et al.[280], we added 1 pseudo-count to each barcode count to avoid division by 0, and performed TMM normalization using the edgeR package (version 3.12.1). For differential abundance analyses, we used the edgeR to compute mutant specific dispersions, p-value and log transformed abundance changes using exactTest function (default setting). The transformed data were subsequently subjected to

differential abundance analyses and clustering analyses. Differential abundance analyses were performed independently for the uptag and downtag sequences, and reproducibility between uptag and downtag data was assessed with Pearson correlation in R. Hierarchical clustering of the thiolutin phenotypic profile to an existing drug response profile for yeast mutants, consisting of 3356 compounds[275] was performed in Cluster (3.0), using centered correlation and complete linkage.

Mutants with significantly altered abundance ($p<0.01$) were subjected to gene-ontology (GO) analyses as previously reported[275]. In brief, all three GO terms, including biological processes (BP), Molecular Function (MF) and Cellular components (CC), were included in our analyses. Following the previous report[275], BP and MF GO terms that are too specific (present in less than 5 genes) or too generic (present in greater than 300 genes) were excluded, although smaller GO term groups were allowed for protein complexes (more than 2 genes) and larger groups (greater than 300 genes) were allowed for cellular components. We computed the GO term enrichment analyses for the significantly resistant or sensitive mutants ($p<0.01$) using hypergeometric test in R (version 3.2.2, physer function, default setting). Raw counts and p-value for each GO term were reported.

Visualization of Yap1 localization under fluorescence microscopy

Microscopy experiments were performed as previously described[80]. In brief, CKY2038 was grown to mid-log ($1\text{-}2\times10^7$ cells/mL) in SC liquid media and added to the ConA

treated perfusion chamber gasket (ThermoFisher, 4 chamber: 19 × 6mm) for treatment at indicated conditions (SC+1% DMSO, SC+10µg/mL Thiolutin, SC+10µg/mL Holomycin or SC+0.4mM H2O2) under the fluorescence microscope. The procedure to prepare the chamber and the microscope setting were the same as previously described.

Glutathione quantitation assays

WT yeast (CKY457) cultures were grown (in YPD at 30˚C) until mid-log phase (~2×10^7 cells/mL), washed and resuspended in the indicated conditions (YPD+1% DMSO, YPD+10µg/mL Thiolutin or YPD+10µg/mL Holomycin). The suspended culture were grown at 30˚C for one hour, and 5×10^7 cells were subject to glutathione (GSH) quantitation. The active and total GSH levels were quantified using the GSH detection kit from Arbor Assays (K006-F5) following the manual. The glutathione disulfide (GSSG) level was calculated by (Total GSH - Active GSH)/2.

UV-Vis assays

Thiolutin reduction and Zn^{2+} chelation reactions were performed in 250µL reaction with 50µM Thiolutin, 50µM $ZnCl_2$ and 100mM potassium phosphate buffer (pH=6.5), as previously described[212]. Mn^{2+} chelation reaction and relevant controls were performed under almost identical conditions except for the 100mM Tris buffer (pH=8) to keep consistent pH with the *in vitro* transcription assays. Each reaction had 250µL final volume and was quantified in Nanodrop 2000c spectrophotometer using UV-transparent cuvettes.

Growth curve, viability and canavanine resistance assays

Yeast growth curve assays with Tecan Infinite F200 plate readers were performed as previously described[80]. For viability assays and canavanine resistance assays, WT yeast strain CKY457 was grown in YPD until mid-log (1-2×10^7 cells/mL), washed and treated with the indicated conditions (YPD, YPD+3µg/mL thiolutin, YPD+5µg/mL thiolutin and YPD+10µg/mL thiolutin) for an hour. The viability was quantified by staining the yeast cells in 0.1% Trypan blue staining (GE Healthcare Life Sciences, Catalog Number SV30084.01) for 3 mins, and the viability was quantified by counting the fraction of unstained yeast cells under the microscope. At least 300 yeast cells were counted in each repeat. Canavanine resistance frequency was quantified by plating the culture onto YPD and SC-Arg+60µg/mL canavanine plates and dividing the number of canavanine resistant colonies by the number of viable colonies on YPD plates. Three experimental replicates were performed.

Pol II transcription activity assays

Pol II was purified using a strain expressing Rpb3 tagged with tandem-affinity tag (TAP), following a procedure as previously described[28]. The DTT was removed from the Pol II buffer using Zeba spin desalting column (Thermo Scientific).

Standard ssDNA transcription assays were performed in the *in vitro* transcription buffer (20mM Tris-HCl pH8, 40mM KCl, 5mM $MgCl_2$). Each reaction had 2µM Pol II, 2µg

denatured sheared salmon sperm DNA as the template and 200μM ATP, CTP, GTP, 9μM UTP, 1μM radioactive α-^{32}P UTP as the substrates. The concentration of thiolutin, reductants (DTT or TCEP) and $MnCl_2$ varies among the experiments and were indicated within each experimental schematic. The addition of order prior to the start of the transcription reaction was also shown in the schematic. Transcription reactions were incubated at room temperature for 30 minutes and stopped with stop buffer (10M Urea, 5mM EDTA in TBE). Synthesized transcripts were separated from unincorporated α-^{32}P UTPs in 10% acrylamide/7M urea gels, and visualized with a Bio-Rad Pharos Phosphorimager. Transcription elongation assays were performed as previously described[28].

Results

Thiolutin or reduced thiolutin alone failed to inhibit fully purified Pol II *in vitro*

Recent publications suggest that thiolutin does not directly inhibit Pol II, in contrast to classic *in vitro* transcription experiments demonstrating thiolutin inhibition of partially purified yeast extract fractions containing Pol I, II, or III[216]. We set out to understand this discrepancy by first examining thiolutin inhibition of purified Pol II. Given that thiolutin and holomycin activities could be sensitive to reduction of the disulfide bond by DTT[212-214] (**Figure 3-1A**), we removed DTT from the Pol II storage buffer by buffer exchange and performed the thiolutin treatment with or without equivalent molar DTT (**Figure 3-1B**). Under buffer conditions employed here we failed to observe thiolutin inhibition of Pol II, with DTT having no effect (**Figure 3-1B**). While these results are

consistent with assertions that thiolutin does not directly target Pol II, they might similarly be explained by the absence of a cofactor or subtleties in experimental conditions between *in vitro* assays. These biochemical assays are revisited later.

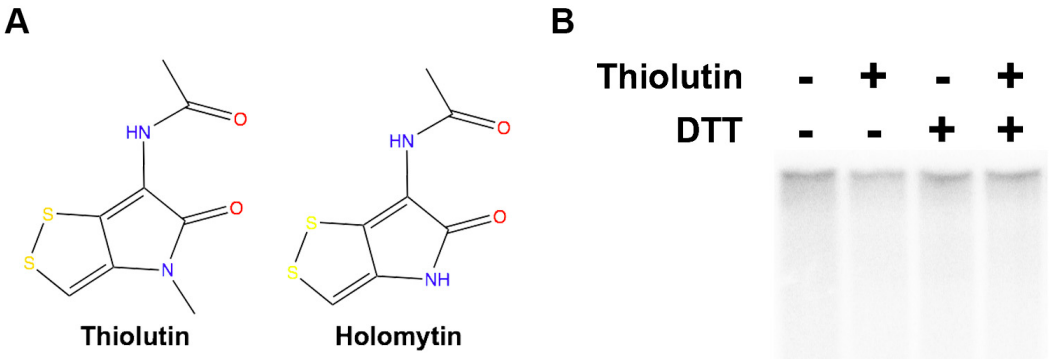


Figure 3-1. Thiolutin or reduced thiolutin alone fail to inhibit purified yeast Pol II *in vitro*
(A) Thiolutin and holomycin structure
(B) Thiolutin, or thiolutin treated with equimolar DTT, show no inhibition of Pol II transcription *in vitro*. Transcription activity assay was performed with ssDNA as the template and NTPs, including ^{32}P labeled UTP. The transcribed ^{32}P containing RNA was separated from free ^{32}P UTPs on 10% polyacrylamide gels for visualization. Experiments were performed for at least three times, and a representative replicate is shown.

Three independent genetic screens for thiolutin resistant and sensitive mutants

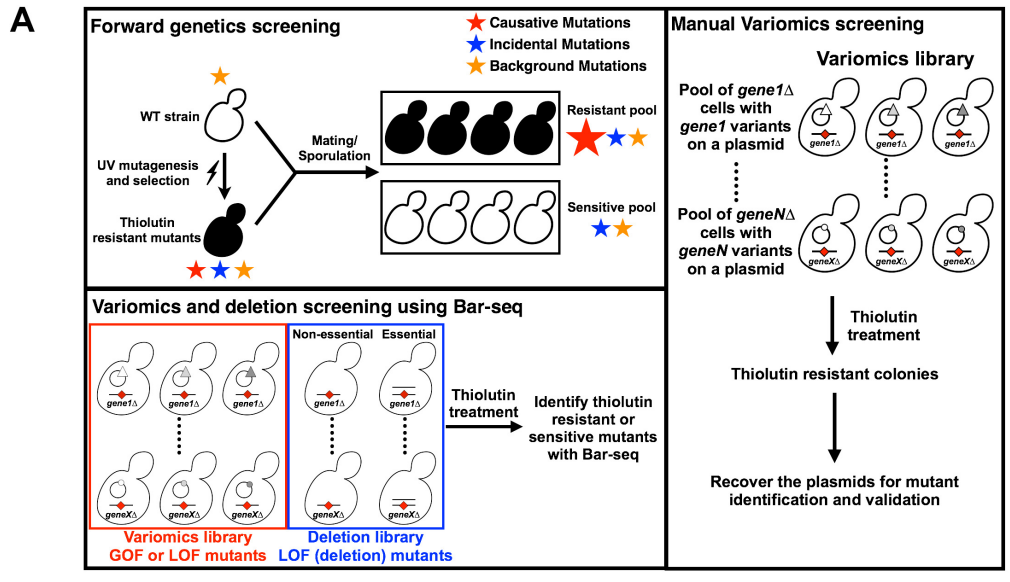
To gain insights into factors controlling the cellular response to thiolutin, we performed three independent genetic screens for modifiers of thiolutin sensitivity (**Figure 3-2A**). First, we performed a conventional forward genetic screen for thiolutin resistant mutants through UV mutagenesis followed by thiolutin resistance selection. The causal mutations for thiolutin resistance were identified by bulk segregant analyses and whole genome sequencing (**Figure 3-2A**)[281]. Second, we screened a recently constructed yeast Variomics library for individual thiolutin resistant candidates[279]. Variomics library has possibly Gain-of-function (GOF) and Loss-of-function (LOF) mutants cloned in a *CEN* based plasmid and expressed in strains with the same endogenous gene deleted. The *CEN* based plasmid is known to replicate 1-2 copies in cells and could result in increased dosage of the gene. Mutants from the forward genetics screen or reproducibly isolated from the manual Variomics screens were re-tested for resistance (**Figure 3-2B**, **Table 3-2**).

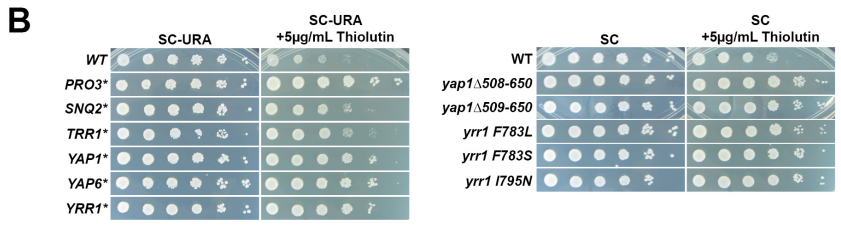
Third, we performed a set of high-throughput screens using both pooled Variomics and deletion libraries (**Figure 3-2A,** material and methods). The thiolutin resistant and sensitive mutants were identified by quantifying changes of mutant-linked DNA barcode frequency under the thiolutin treated versus control conditions, using a well-established barcode sequencing (Bar-seq) approach[274,275]. A series of quality controls were performed on the reproducibility among biological replicates, consistency between the two barcodes on each gene deletion (uptag and downtag) for the same strain, common

and distinct properties between Variomics and deletion libraries following a well-established statistical analysis pipeline (**Figures 3-3**)[275,280]. A subset of the statistically significant resistant or sensitive mutants were further validated by re-constructing the deletion strain for individual phenotyping (**Figure 3-4**). Overall, our data exhibited excellent reproducibility, and revealed valuable common and distinct features between Variomics (GOF and LOF mutants) and deletion (LOF mutants) libraries. Different classes of genes controlling thiolutin sensitivity emerging from these studies are discussed below.

Functional dissection of the multidrug resistance (MDR) and oxidative stress response (OSR) pathways in response to thiolutin

We first identified a series of thiolutin resistant mutants that harbored mutations in MDR and OSR genes from all three genetic screens. These results are perhaps not surprising given that yeast has well characterized MDR pathways that sense and pump out the toxic small molecule compounds as a self-protection mechanism, or given the recent reports that thiolutin and holomycin activities require reduction *in vitro*[213,214]. Reduction of thiolutin may itself oxidize cellular proteins and cause oxidative stress. In addition, it is also reported that reduced holomycin may act as a redox cycler, which could possibly generate reactive oxygen species (ROS) in the cell. In fact, it was suggested in one report that thiolutin appeared to induce expression of several OSR genes[273]. However, whether thiolutin indeed caused oxidative stress and the specific MDR factors for thiolutin resistance have not been rigorously tested. In light of our isolated thiolutin resistant





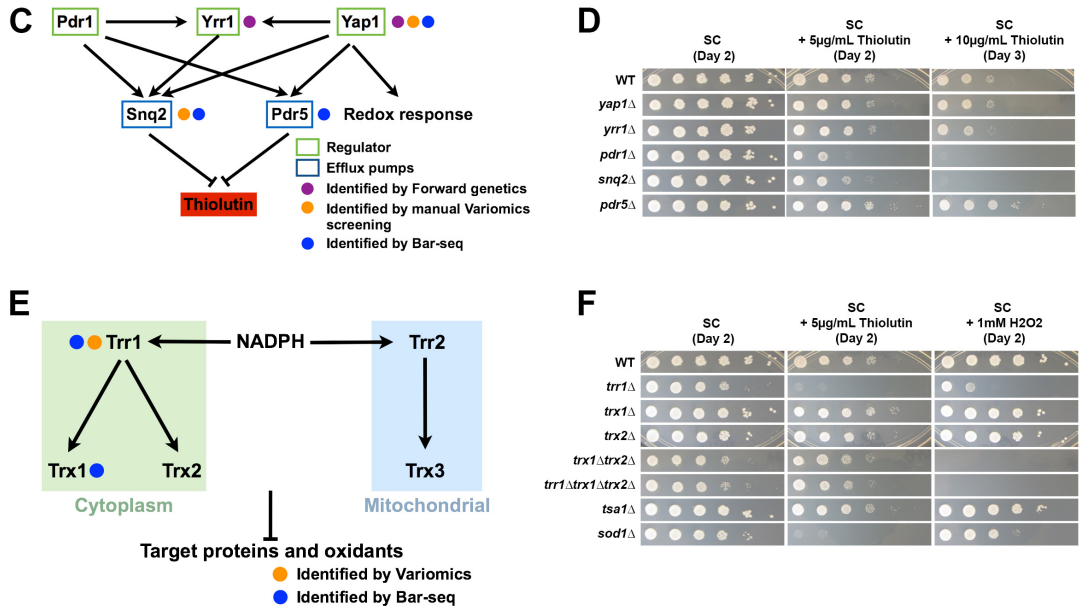


Figure 3-2. Mutants in multiple drug resistance (MDR) and oxidative stress responses (OSR) pathways confer distinct resistance or sensitivity to thiolutin

(A) Schematic diagram of the three independent genetic screens for thiolutin resistant mutants. Top left: For the forward genetic screening, WT yeast strain (CKY457) was UV mutagenized and screened for thiolutin resistant mutants, and mutations were identified by whole-genome sequencing. Causative mutations were enriched by bulk segregant analyses. Top right: The recently constructed gene-specific Variomics libraries consists of many mutants for a given gene in a yeast cells pool, and Variomics libraries > 4000 genes are further pooled together for thiolutin resistance screening. Plasmids from the validated thiolutin resistant mutants are recovered for Sanger sequencing to identify both the gene and the mutations. Bottom left: Four libraries (Variomics libraries for non-essential or essential genes, Deletion libraries for non-essential or essential genes) are used for screening thiolutin resistant or sensitive mutants. Changes in abundance in the library was detected by deep sequencing of the PCR amplicon of the barcode region. Variomics libraries consist of point mutation variants and can be used to screen for both Gain-of-Function (GOF) and Loss-of-Function (LOF) alleles, whereas "Deletion" libraries consist of single-gene-deleted mutants that can be used to screen for complete LOF mutants.

B: Reproducibly isolated thiolutin resistant mutants from the forward genetics and manual Variomics screening.

C: Schematic diagram summarizing a partial regulatory network of multiple drug resistance, with isolated resistant mutants indicated. The reported or hypothesized functional interactions are indicated by pointed (activation) or blunt-end (inhibition) arrows.

D: Distinct thiolutin sensitivity of MDR deficient mutants. *pdr1*Δ and *snq2*Δ are sensitive to thiolutin, whereas other tested MDR deficient strains retain resistance.

E: Schematic diagram summarizing the regulatory network of thioredoxin system in yeast, with isolated resistant mutants indicated.

F: Distinct thiolutin sensitivity of some OSR deficient mutants.

Table 3-2. Reproducibly isolated resistant candidates from the manual Variomics screens

Candidate Number	Class	Plasmid dependence	Dominance	Gene	Mutation
15	Essential	N.A.	No	*PRO3*	H75Q, I268T
16	Essential	N.A.	No	*PRO3*	WT
42	Essential	N.A.	No	*PRO3*	WT
220	Non-essential	Yes	Yes	*SNQ2*	R307 silent (AGA->AGG), P1358 silent (CCT->CCC)
221	Non-essential	Yes	Yes	*SNQ2*	R307 silent (AGA->AGG), P1358 silent (CCT->CCC)
7	Essential	N.A.	No	*TRR1*	T76A
9	Essential	N.A.	No	*TRR1*	S242G
186	Non-essential	Yes	Yes	*YAP1*	D382G, E516G, S567G, N581D
211	Non-essential	Yes	Yes	*YAP1*	V260I, T312A, S528P
24	Non-essential	Yes	No	*YAP6*	H93Y, C-terminal truncation *yap6Δ368-383*
27	Non-essential	Yes	No	*YAP6*	L6 silent mutation (TTG->CTG), Q152R, C-terminal truncation *yap6Δ368-383*
28	Non-essential	Yes	No	*YAP6*	G80 silent mutation (GGT->GGC), S218 silent mutation (TCA->TCG), C-terminal truncation *yap6Δ368-383*
31	Non-essential	Yes	No	*YAP6*	Upstream mutation (T->C), A179T, C-terminal truncation *yap6Δ380-383*
33	Non-essential	Yes	No	*YAP6*	WT

Table 3-2. Continued

Candidate Number	Class	Plasmid dependence	Dominance	Gene	Mutation
44	Non-essential	Yes	Yes	*YRR1*	V499 silent mutation (GTA->GTG), T610 silent mutation (ACA->ACG), N703D
46	Non-essential	Yes	Yes	*YRR1*	Y134C, Y165C, K493E, T696 silent mutation (ACC->ACA), N703D, T802 silent mutation (ACT->ACC)

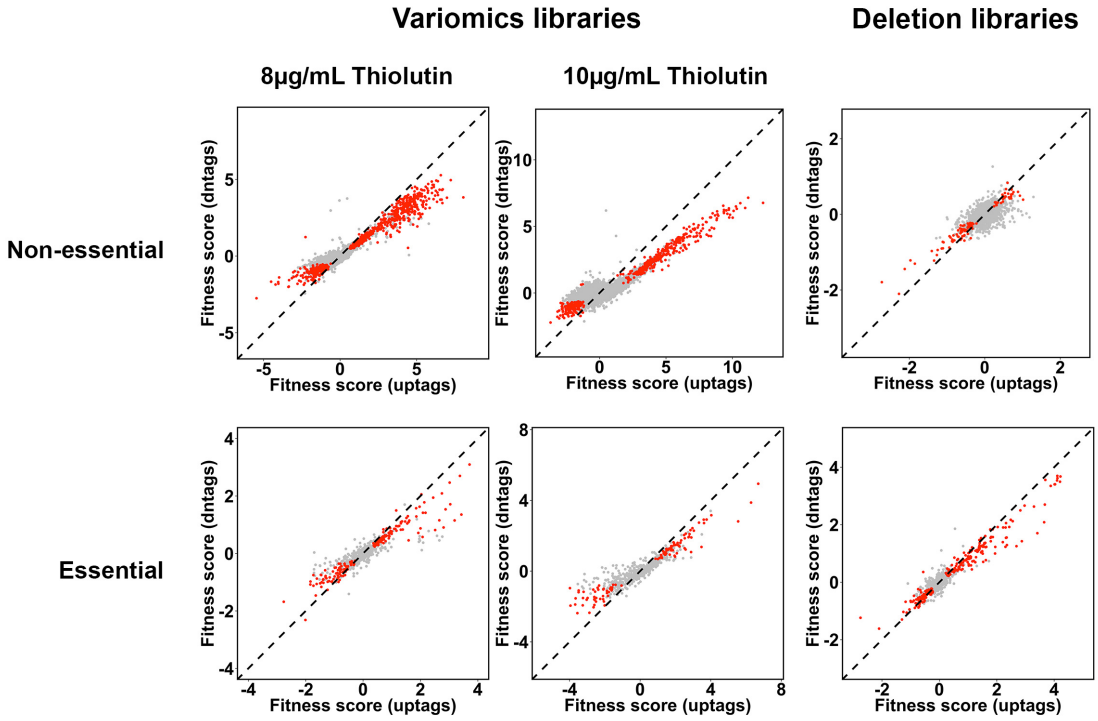


Figure 3-3. Correlation between uptag and downtags results
Xyplots showing the fitness score (log2 of the fold change in relative abundance) computed from both uptag and downtags. Mutants are declared significantly resistant or sensitive in both tags are colored in red, whereas others are colored in light grey. y=x is shown as a dashed line.

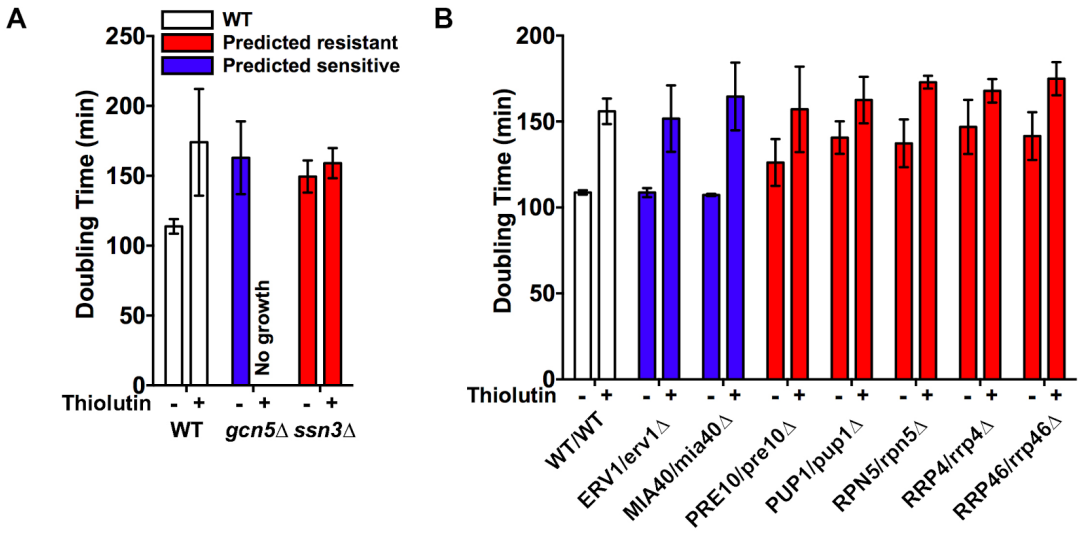


Figure 3-4. Validation of several statistically significantly resistant or sensitive mutants. Mutants were reconstructed in haploid BY4741 or diploid BY4743 backgrounds originally used to construct the deletion libraries, and mid-log growth at 30˚C was assessed using a Tecan plate reader. Three independent repeats were performed and the error bars represent standard deviation of the mean. We failed to validate the observed slight sensitivity in *ERV1/erv1*Δ and *MIA40/mia40*Δ, two strains that were statistically significantly sensitive but the sensitivity was very slight.

mutants, we functionally dissected several MDR and OSR pathways for thiolutin sensitivity, as discussed below.

From both manual Variomics and forward genetic screens, we isolated multiple mutants for Yap1 and Yrr1, two transcription factors that have been implicated in MDR pathways (**Figure 3-2B**). Though implicated in MDR, Yap1 is also known as a master regulator for oxidative stress responses and will be discussed again in a later section. In addition, we also isolated mutants in a well-known MDR efflux pump Snq2 from the manual Variomics screen. Finally, from high-throughput Bar-seq screens, we have observed distinct thiolutin resistance and sensitivity in many more barcodes (**Figure 3-6A**) linked to MDR transcription factors or efflux pumps.

Several lines of evidence suggested that the isolated thiolutin resistant MDR mutants were GOF mutations. First, isolated *yap1* mutants showed clustering of mutations in regions encoding the C-terminus, where a variety of mutants are known to cause GOF[282-286]. Second, isolated MDR mutants from the manual Variomics screen were dominant or possibly functioned through increased dosage (**Figure 3-5**). Third, in high-throughput Bar-seq screens, we found Variomics mutants in several MDR genes (*YAP1*, *YRR1*, *SNQ2*) conferred thiolutin resistance whereas deletion of the same genes either conferred sensitivity or had no strong effect (**Figure 3-6A**), suggesting that these MDR Variomics mutants behave differently from the gene deletion mutants (complete LOF).

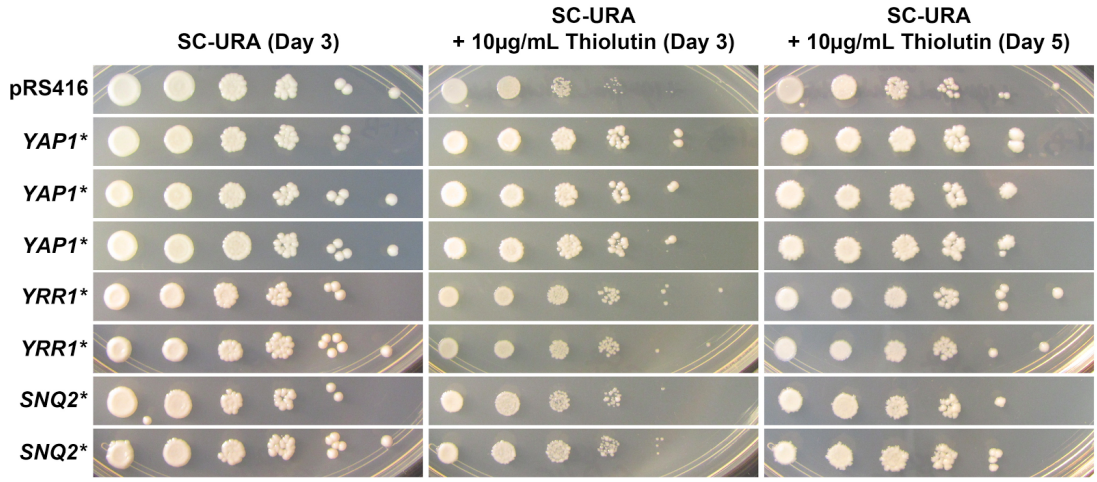


Figure 3-5. Isolated thiolutin resistant MDR Variomics candidates are dominant or dosage dependent

Plasmids recovered from the Variomics candidates are transformed into WT yeast strain along with an empty vector (pRS416) control, thiolutin resistance are assessed for at least two independent transformants, and the dominant mutants are shown in this figure.

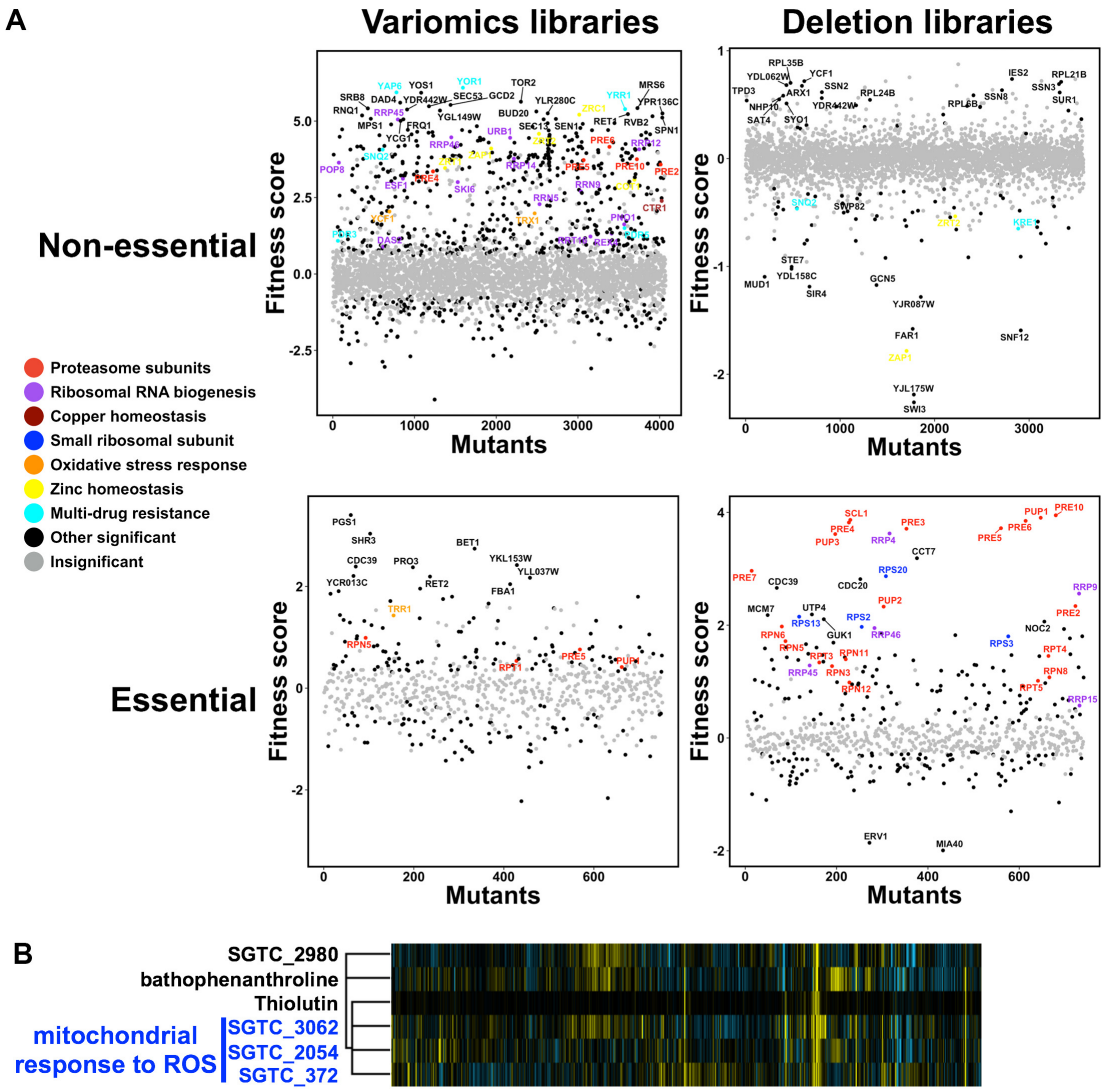


Figure 3-6. Mutations in multiple distinct cellular pathways confer resistance/sensitivity to thiolutin
(A) Different thiolutin resistant or sensitive mutant classes are revealed in the Bar-seq based screenings of pooled Variomics and Deletion libraries.
(B) Thiolutin induced phenotypic profile in the pooled deletion libraries is co-clustered with bathophenanthroline. The heatmap shows 4683 deletion mutants' responses (in columns) to different compounds (in rows). Resistant strains are colored in yellow, and sensitive strains are colored in blue. All 3356 compounds were used for hierarchical clustering, but only the thiolutin closely correlated compounds are shown in this figure for clarity.

In order the better understand the nature of identified MDR mutants from variomics and forward genetics, we constructed a series of MDR gene deletion mutants and tested their sensitivity to thiolutin (**Figure 3-2D**). We found that none of the tested deletion mutants was thiolutin resistant (**Figure 3-2D**). The lack of resistance for all the tested MDR deletion strains is consistent with the isolated thiolutin resistant MDR Variomics and forward genetics mutants increasing MDR activity. In addition, *pdr1*Δ and *snq2*Δ conferred hypersensitivity to thiolutin (**Figure 3-2D**), consistent with the Pdr1 and Snq2 functions in promoting thiolutin resistance for WT strains. Together, we conclude that the MDR transcription factor Pdr1 is the dominant factor promoting expression of efflux pumps to reduce thiolutin concentration in the cell, whereas neither Yrr1 nor Yap1 appear to promote baseline resistance, given that *yap1*Δ and *yrr1*Δ are not hypersensitive to thiolutin. However, increased Yap1 or Yrr1 activities through mutations or dosage may increase efflux pump expression to promote thiolutin resistance, due to the known linkages to many multidrug efflux pumps[287-289]. We also suggest that Snq2 is a thiolutin efflux pump whereas Pdr5 is not.

Holomycin did not inhibit yeast cell growth at low concentration[164], and we recapitulated these observations (**Figure 3-7A**). Given that structurally similar compounds may have different cell permeability and efflux efficiency, we asked whether the same set of MDR deletion strains conferred sensitivity to holomycin, but failed to observe any (**Figure 3-7A**), although we cannot rule out the possibility that thiolutin and holomycin may be transported by different efflux pumps.

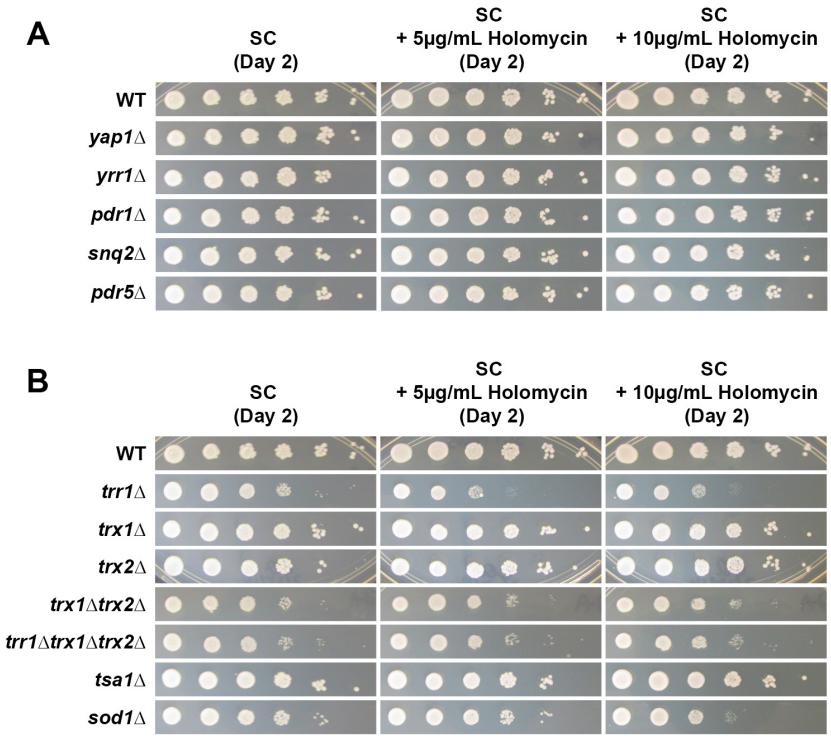


Figure 3-7. Tested MDR and OSR deficient mutants do not confer same hypersensitivity to holomycin as to thiolutin.

(A) Tested MDR mutants do not confer hypersensitivity to holomycin

(B) Most tested OSR mutants do not confer same hypersensitivity to holomycin as to thiolutin. *trr1*Δ and *sod1*Δ are slightly sensitive to holomycin, while others are not.

In addition to MDR pathways, we also identified thiolutin modulating mutations in OSR-related genes, including *YAP1* (discussed above) and genes in the thioredoxin pathway. Thioredoxins are a series of small anti-oxidant proteins that primarily function in reducing specific cysteines in client proteins (reviewed in [290] and references therein), and in turn are themselves reduced by thioredoxin reductases (**Figure 3-2E**). In yeast, there are two thioredoxin reductases (cytoplasmic Trr1 and mitochondrial Trr2) and three thioredoxins (cytoplasmic Trx1, Trx2; mitochondrial Trx3) (**Figure 3-2E**). Mutants in *TRR1* were reproducibly isolated from our manual Variomics screen (**Figure 3-2B,E; Table 3-2**). In addition, *trr1* and *trx1* related barcodes were linked to thiolutin resistance in high-throughput screening of Variomics but not deletion libraries (**Figures 3-2E, 3-6A**), suggesting that the isolated *trr1* mutants could be GOF due to specific alleles or increased dosage of the plasmid.

TRR1 was found to be essential in large scale gene deletion analysis[291] but nonessential in classical genetic experiments[292-296]. To understand the nature of *TRR1* function in thiolutin resistance/sensitivity, we constructed a *trr1*Δ strain to directly test its thiolutin sensitivity (**Figure 3-2F**). Interestingly, *trr1*Δ conferred hypersensitivity to thiolutin (**Figure 3-2F**), consistent with Trr1 activity antagonizing thiolutin and specific *trr1* mutants or increased Trr1 dosage enhancing this antagonism. The Trr1 function in antagonizing thiolutin also suggests that our isolated *yap1* alleles may confer thiolutin resistance through promoting *TRR1* expression. In fact, several *YAP1* GOF alleles have

been shown to increase expression of both MDR and OSR pathways through Yap1 nuclear accumulation[282-286].

We conceive of *trr1*Δ hypersensitivity to thiolutin in two ways. It could be due to loss of Trr1 function in directly counteracting thiolutin activity or the accumulation of the oxidized thioredoxins, or both. The observed *trx1*-linked thiolutin resistance in high-throughput Variomics screens suggested critical functions of thioredoxin(s) in thiolutin resistance downstream of *TRR1*. Therefore, we examined thioredoxin deletions directly for thiolutin resistance. We found that the *trx1*Δ, *trx2*Δ or *trx1*Δ*trx2*Δ mutants did not confer thiolutin sensitivity (**Figure 3-2F**). In contrast, *trx1*Δ*trx2*Δ conferred thiolutin resistance, and further, suppressed thiolutin hypersensitivity in *trr1*Δ (**Figure 3-2F**), consistent with the hypothesis that *trr1*Δ hypersensitivity to thiolutin is due to accumulation of oxidized thioredoxins. The observed *trx1*Δ*trx2*Δ thiolutin resistance reveals critical requirement of Trx1 and Trx2 for thiolutin activity, and suggests that Trx1 and Trx2 may function either directly or indirectly in thiolutin reduction. However, further experiments are needed to demonstrate the direct thiolutin reduction by Trx1 and Trx2.

The distinct thiolutin resistance conferred by *yap1* alleles and thioredoxin mutants suggests possible connections to OSR. Therefore, we set to test the roles of two additional OSR genes, *TSA1*, encoding a thioredoxin peroxidase, and *SOD1*, encoding a cytosolic copper-zinc superoxide dismutase, in thiolutin resistance. Interestingly, *sod1*Δ,

but not *tsa1*Δ, conferred hypersensitivity to thiolutin (**Figure 3-2F**). Given that holomycin is a known redox cycler, we asked if simple addition of H_2O_2 could mimic the thiolutin effect in tested thioredoxin and OSR gene deletion mutants (**Figure 3-2F**). Thiolutin sensitivities in tested strains were distinct from H_2O_2 sensitivities, suggesting distinct effects on cellular function. Finally, we tested if yeast could be sensitized to holomycin by the same set of mutants (**Figure 3-7B**). Interestingly, *trr1*Δ and *sod1*Δ were slightly sensitive to 10μg/mL holomycin (**Figure 3-7B**), suggesting that the induction of oxidative stress is a general property of dithiolopyrrolone compounds and that yeast are at least somewhat permeable to holomycin.

Thiolutin induces apparent oxidative stress

Multiple lines of evidence presented above indicated that thiolutin may induce oxidative stress and increased Yap1 function may promote thiolutin resistance. Yap1 functions by sensing the cellular oxidants through its C-terminal cysteines, causing it to translocate into the nucleus and up-regulating a series of OSR genes to counteract the cellular oxidants upon oxidative stress[282-286]. Therefore, nuclear localization of Yap1 is in general a sign of increased Yap1 function and OSR[282,285,286]. To test whether thiolutin treatment caused nuclear translocation of Yap1, we fused *YAP1* gene DNA to sequence encoding EGFP on the Yap1 C-terminus, and monitored Yap1::EGFP localization after thiolutin treatment. Thiolutin indeed induced Yap1::EGFP nuclear localization, consistent with induction of oxidative stress, although the thiolutin induced Yap1::EGFP nuclear translocation appeared to be slower and weaker than that induced by H_2O_2

(**Figure 3-8A**). In addition, thiolutin depleted total glutathione within an hour of treatment (**Figure 3-8B**). However, thiolutin depleted both active GSH and GSSG, a behavior similar to glutathione synthesis defective mutants but distinct from several well-known oxidants such as H_2O_2 and menadione[297-299]. In contrast, holomycin did not have a clear effect on Yap1 localization or cellular glutathione level (**Figure 3-8A,B**), consistent with the lack of growth inhibition by holomycin and the distinct activities between the two structurally similar analogs.

Previous studies suggested that the reduced holomycin was spontaneously re-oxidized when exposed to air[212], consistent with the behavior of redox-cycling compounds[225,226]. To test if thiolutin exhibited similar behavior, we monitored thiolutin reduction and re-oxidation by its unique UV absorbance due to the conjugated ene-dithiol groups (**Figure 3-8C**), as with the recent reports[212-214]. Thiolutin could be reduced by DTT or TCEP, but not glutathione (**Figure 3-8C**), consistent with the behavior of holomycin[212]. Reduced thiolutin could be also re-oxidized (**Figure 3-8D**), suggesting that thiolutin can act as a redox-cycler, similar to holomycin. Together, thiolutin induced the Yap1 nuclear accumulation and hypersensitivity of *sod1*Δ and *trr1*Δ mutant *in vivo* and appeared to be a redox cycler *in vitro*, consistent with the induction of oxidative stress. However, we did not find direct evidence of glutathione oxidation, suggesting that additional experiments are necessary to validate the thiolutin-mediated apparent OSR.

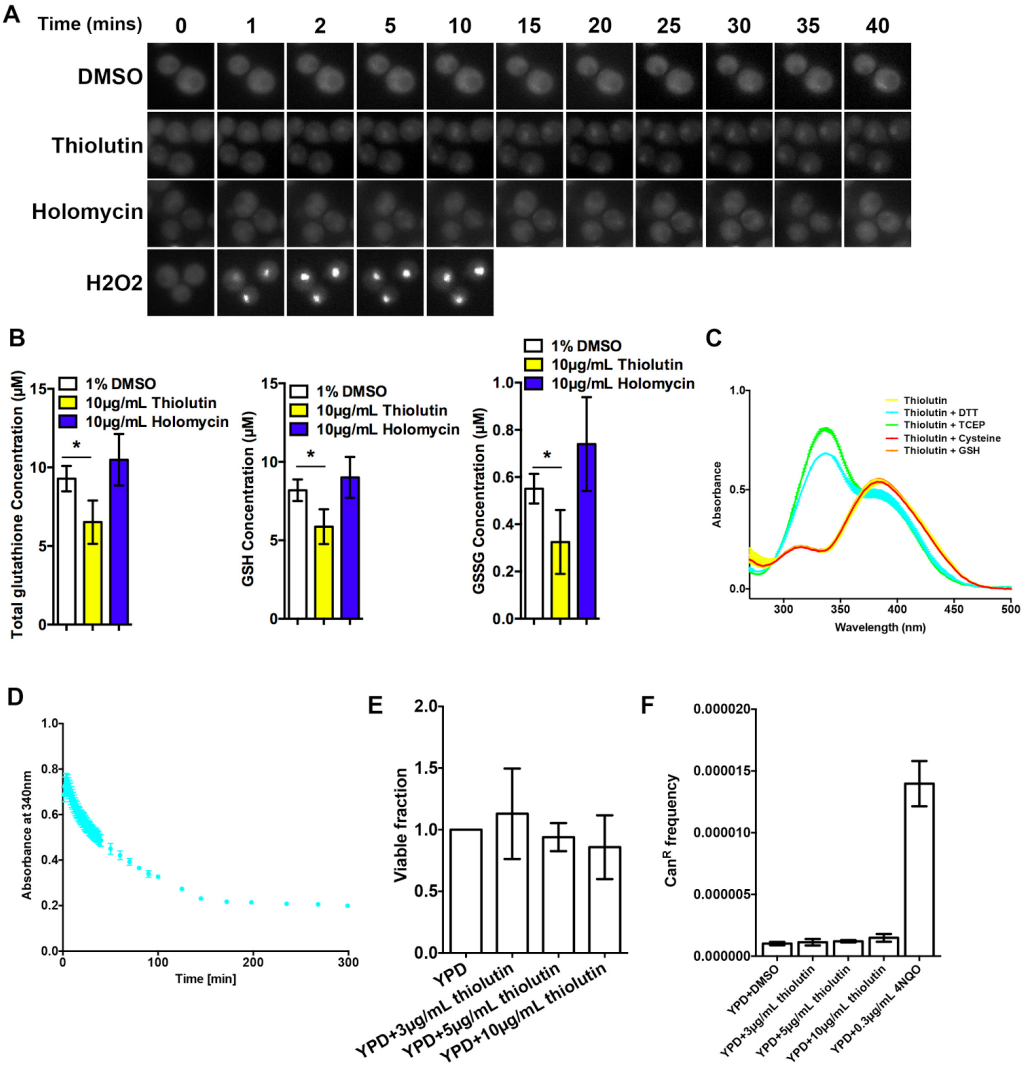


Figure 3-8. Thiolutin appears to induce oxidative stress partially through redox cycling
(A) Thiolutin induces nuclear localization of the redox-sensitive Yap1 transcription activator. EGFP-tagged Yap1 strain treated with 1% DMSO (control), 10µg/mL thiolutin or 0.4mM H_2O_2. GFP fluorescence was monitored after treatment at the indicated time points.
(B) Thiolutin depletes total glutathione *in vivo*. Growing WT yeast cultures are washed and treated with indicated conditions at 30˚C for an hour. Reduced glutathione (GSH) and total glutathione are measured in three independent replicates, and the oxidized glutathione (GSSG) are calculated. The error bars represent standard deviation of the mean. *p=0.01 (Two-tailed paired t-test).
(C) Thiolutin can be reduced by DTT and TCEP, but not glutathione and cysteine *in vitro*. The reactions were performed in 100mM phosphate buffer (pH=6.5) with equivalent molar reductant with thiolutin. UV spectra were measured 1min after the reactions. Three independent replicates were performed and the error bars represent standard deviation of the mean.
(D) Reduced thiolutin is spontaneously re-oxidized when exposed to air. The reactions were performed as described in Figure 3-9C, and UV absorbance at 340nm plotted over a time course. Three independent replicates were performed and the error bars represent standard deviation of the mean.
(E) Thiolutin does not affect yeast viability.
(F) Thiolutin does not cause DNA damage, as measured by mutation frequency at *CAN1* gene.

DNA damage can be a consequence of redox-cycling and can lead to transcription inhibition and cytotoxicity[300,301]. We asked if thiolutin treatment could cause DNA damage or decrease viability. We did not observe increase in mutation rate as measured by resistance to canavanine, consistent with little or no effect on DNA damage rates, nor did we observe a viability defect (**Figure 3-8E, F**). The lack of viability defect is consistent with an early report showing that thiolutin inhibition was reversible after several hours' treatment[215]. Together, we suggest that thiolutin is a redox cycler *in vitro* and possibly also *in vivo*, but does not appear to inhibit transcription through widespread DNA damage.

Thiolutin alters Zn^{2+} homeostasis

The recent detection of direct Zn^{2+} chelation activity of reduced thiolutin and holomycin explains multiple thiolutin-induced phenotypes, such as inhibition of proteasome and glucose utilization[213,214,227]. From the high-throughput screens, we observed distinct thiolutin resistance or sensitivity linked to diverse Zn^{2+} trafficking genes, including the master transcription regulator Zap1 and various Zn^{2+} transporters Zrt1, Zrt2, Zrc1 and Cot1 (**Figure 3-6A**). Variomics strains linked to *ZAP1*, *ZRT1*, *ZRT2*, *ZRC1* and *COT1* conferred thiolutin resistance whereas deletion mutants were either not resistant (*zrt1*Δ*, zrc1*Δ*, cot1*Δ) or hypersensitive (*zap1*Δ *and zrt2*Δ) (**Figure 3-6A**), suggesting that these Variomics mutants could be allele-specific or dosage modifiers of thiolutin. We further tested Zap1 function in thiolutin resistance by constructing a *zap1*Δ strain and testing its sensitivity to thiolutin (**Figure 3-9**). *zap1*Δ confers hypersensitivity to thiolutin,

consistent with Zap1 functions in cellular resistance to thiolutin and thiolutin alteration of Zn^{2+} homeostasis *in vivo*.

Surprisingly, *zap1*Δ also conferred hypersensitivity to holomycin (**Figure 3-9**). This observation suggests that altering zinc homeostasis is a conserved mechanism among dithiolopyrrolones and indicates yeast cells are permeable to holomycin. This result also argues that holomycin can access essential targets in yeast under the right conditions. *zap1*Δ is known to be sensitive to oxidative stress[302-305], meaning it is possible that *zap1*Δ hypersensitivity to thiolutin and holomycin could be due to zinc deficiency-induced oxidative stress. However, *zap1*Δ hyper-sensitivity to thiolutin and holomycin was much stronger than that to H_2O_2, suggesting Zn^{2+} homeostasis defects in addition to the oxidative stress exacerbation.

Given our observation that a Cu^{2+} transporting gene CTR1 may be also linked to thiolutin resistance (**Figure 3-6A**), we tested whether reduced thiolutin chelates Cu^{2+} and also confirmed Zn^{2+} chelation. Consistent with the two recent reports, Zn^{2+} caused a distinct UV shift of reduced thiolutin from ~340nm to ~370nm but did not change the UV spectra of non-reduced thiolutin (**Figure 3-9B**). In addition, we found that Cu^{2+} can also bind to reduced but not non-reduced thiolutin, in a manner likely similar to Zn^{2+} binding (**Figure 3-9B**). In contrast, Mg^{2+} does not bind to either reduced or non-reduced thiolutin, distinct from the other two tested metals (**Figure 3-9B**). Finally, Zn^{2+} supplementation partially suppressed the thiolutin inhibition (**Figure 3-9C**). The lack of

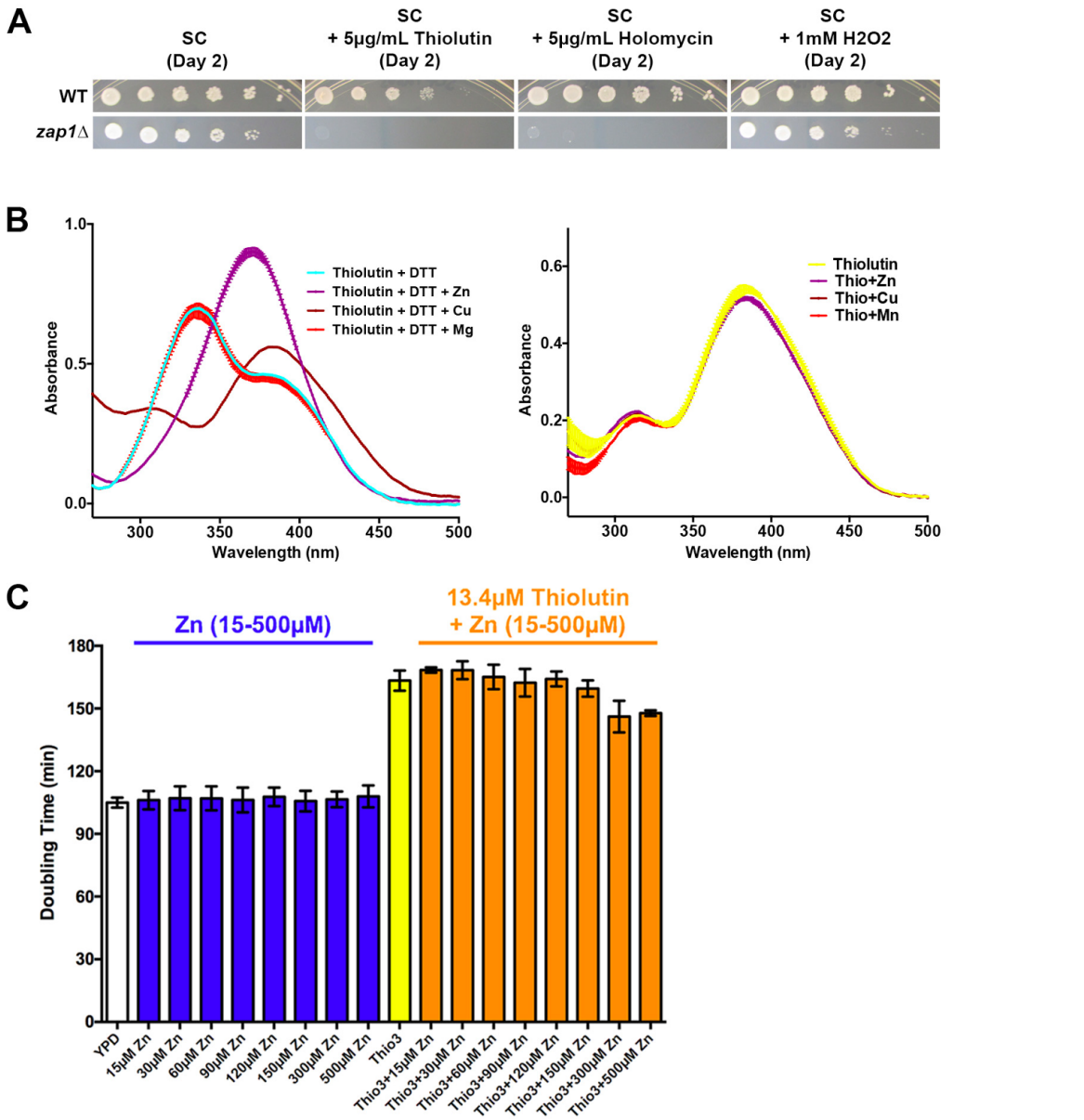


Figure 3-9. Reduced thiolutin chelates Zn^{2+} and Cu^{2+} *in vitro* and alters Zn^{2+} homeostasis *in vivo*
(A) *zap1*Δ is hypersensitive to thiolutin and holomycin.
(B) Zn^{2+} and Cu^{2+}, but not Mg^{2+}, alter the UV spectra of reduced thiolutin. The reactions are performed as described in 4C. Three independent repeats were performed and the error bars represent standard deviation of the mean.
(C) Zn^{2+} supplementation does not fully suppress thiolutin sensitivity. Doubling time were derived from growth curve measured by Tecan plate reader. Three independent repeats were performed and the error bars represent standard deviation of the mean.

full suppression may suggest that there are other thiolutin-mediated defects in addition to Zn^{2+} chelation, although we cannot rule out that inefficient Zn^{2+} trafficking limits effective Zn^{2+} supplementation in the cell.

Additional cellular pathways are involved in thiolutin resistance

The high-throughput screens of pooled yeast Variomics and Deletion libraries, compared to conventional forward genetics and manual Variomics screens, revealed several additional pathways in thiolutin resistance (**Figure 3-6A**). We isolated mutants involved in the proteasome, ribosomal RNA biogenesis, and small ribosomal subunits (**Figure 3-6A**). Many observations from the Variomics screens were confirmed from two conventional screens and direct deletion analyses (**Figure 3-2, Table 3-2**). We further constructed 9 homozygous (for non-essential genes) or heterozygous (for essential genes) deletion mutants to validate the observation from the bar-seq screens of the deletion libraries. We validated the observations in 7 (out of 9) mutants, including one resistant and one sensitive mutants from the non-essential pool, and 5 resistant mutants from essential pool. Among them, we validated the observed thiolutin resistance conferred by three heterozygous deletion strains in proteasome subunits (*PRE10/pre10*Δ, *PUP1/pup1*Δ, *RPN5/rpn5*Δ) (**Figure 3-4**).

The reconstructed strains validated the surprising but clear observation that heterozygous deletion mutants in proteasome subunits universally conferred thiolutin resistance (**Figure 3-6A**). This is surprising because thiolutin was recently demonstrated to be a

proteasome inhibitor through Zn^{2+} chelation[214], and heterozygous proteasome subunit deletion mutants presumably may lead to reduced proteasome activity and would be expected to confer sensitivity to proteasome inhibitors. Interestingly, two recent reports observed similar paradoxical resistance to proteasome inhibitors in yeast strains and human cell lines with decreased level of the regulatory 19S proteasome subunits[306,307]. In yeast and human, the fully assembled 26S proteasome consists of a 20S catalytic core and a 19S regulatory complex. It was demonstrated that decrease or transient inhibition of 19S proteasome subunits may in turn induce the level and activity of partially assembled 20S proteasome instead of the fully assembled 26S proteasome[307], resulting in a net increase of proteasome function. It was also recently suggested that specific reduced expression of 19S proteasome subunits induced an altered cellular state and altered the global transcriptome in response to proteasome inhibitors[308]. However, our observation is distinct, because we observed decreases in both 19S and 20S proteasome subunits conferred thiolutin resistance. Whether the proteasome has increased level and activity in these strains has yet to be further tested.

Thiolutin, when activated by DTT and Mn^{2+}, directly inhibits Pol II *in vitro*

Despite the progress in understanding thiolutin induced diverse cellular responses, whether and how thiolutin directly inhibits yeast Pol II *in vitro* remained unresolved. Multiple lines of evidence suggest possible involvement of divalent metals in this process. First, the co-clustering of thiolutin and bathophenanthroline (**Figure 3-10**) induced phenotypic profiles suggests similar modes of action (**Figure 3-6B**).

Bathophenanthroline is highly similar to a well-known transcription inhibitor 1,10-phenanthroline (**Figure 3-10**), which appears to require Cu^{2+} for direct transcription inhibition *in vitro*[309,310], though these compounds may have similar activities because they are also well-known Zn chelators[311]. Second, the direct interaction between reduced thiolutin and Zn^{2+} (or Cu^{2+}) suggests the possibility that thiolutin may function with metal co-factors.

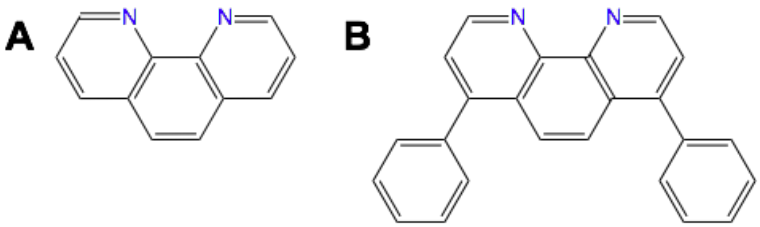


Figure 3-10. Structures of 1,10-phenanthroline and bathophenanthroline
(A) 1,10-phenanthroline
(B) bathophenanthroline

Many divalent metals inhibit RNAP activity (*e.g.* Pb^{2+}, Zn^{2+}, Cu^{2+}, Be^{2+}, Cd^{2+}, Ca^{2+})[248], and additionally are tightly controlled in cells due to toxicity (*e.g.* Fe^{2+}, Co^{2+}), so are unlikely to participate in possible Pol II inhibition by thiolutin. Mn^{2+} does not inhibit but instead increases the activity of multiple RNA polymerases (including Pol II)[248]. In addition, Mn^{2+} is readily available in yeast cells (0.04-2mM depending on the types of measurement)[312-316] at a relevant range to the reported thiolutin inhibitory concentration (~20μM in [216]). We note that in the original Tipper observations, their transcription buffer contained 1.6mM Mn^{2+}, whereas most other reports did not report Mn^{2+} in buffers[164,213,220,222] (One report appears to have an unusually high concentration of 10mM Mn^{2+} and possibly also reductant[221]). Therefore, we set out to test whether Mn^{2+} might participate in thiolutin-mediated inhibition of Pol II. Remarkably, Mn^{2+} addition to reduced thiolutin potently inhibited purified Pol II activity (**Figure 3-11A)**, in sharp contrast to the activation of Pol II activity caused by Mn^{2+} by itself in the absence of reduced thiolutin. Our observation is consistent with the known activation of RNAPs by Mn^{2+} [65,248], but also suggests a highly potent Pol II inhibition by Mn^{2+} activated reduced thiolutin (hereafter termed as Thio/Mn^{2+}) (**Figure 3-11A**).

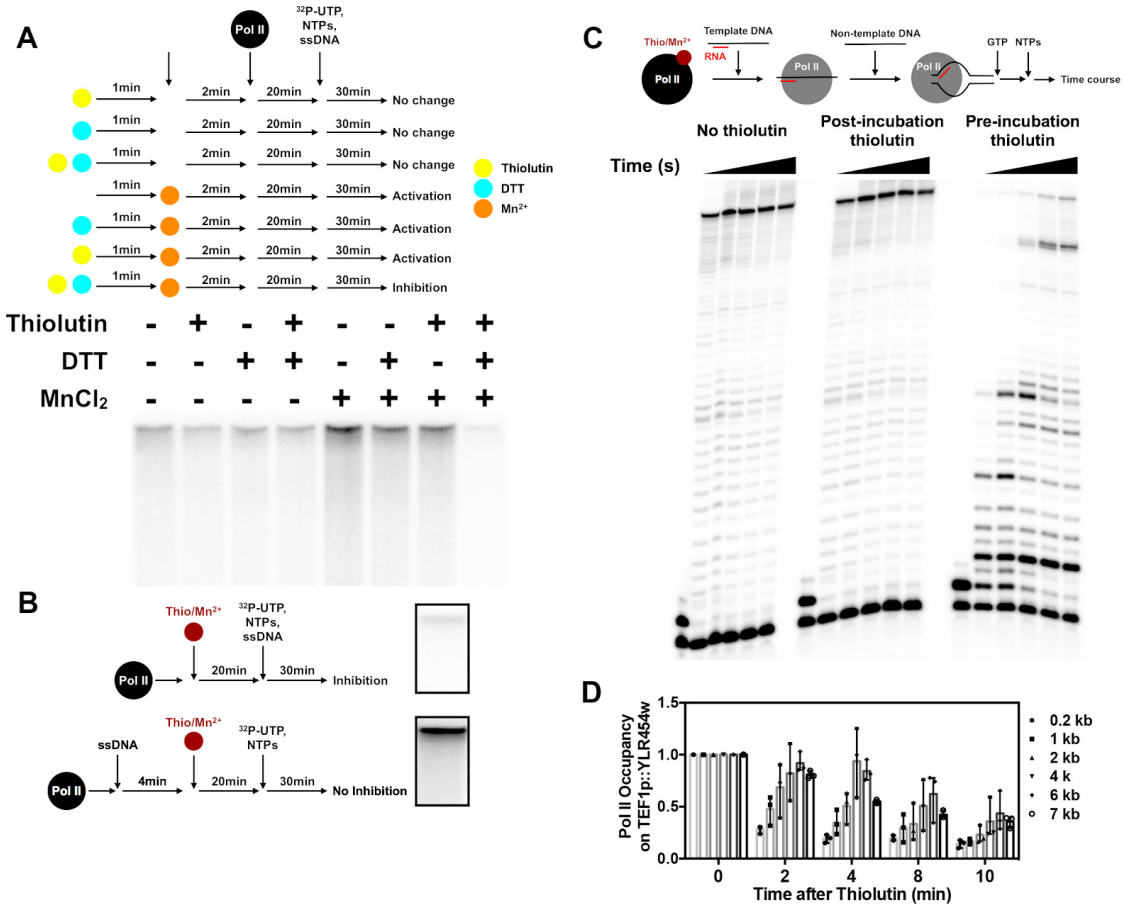


Figure 3-11. Thiolutin/Mn^{2+} complex inhibits Pol II transcription *in vitro* and *in vivo*
(A) Thiolutin requires both DTT (reductant) and Mn^{2+} to inhibit Pol II transcription *in vitro*. At least three experimental replicates were performed, and a representative replicate is shown.
(B) Pre-binding to DNA renders Pol II resistant to Thiolutin/Mn^{2+} complex. Three experimental replicates were performed, and a representative replicate is shown.
(C) Thio/Mn^{2+} inhibited Pol II can be manually assembled into a distinct and slow elongating complex. Two experimental replicates were performed, and a representative replicate is shown.
(D) Thio/DTT/Mn^{2+} inhibits Pol II transcription *in vivo*. Three experimental replicates were performed, and the error bars represent standard deviation of the mean.

We next designed a series of experiments to investigate properties of the inhibited Pol II and nature of the inhibitory species. We first tested if order-of-addition among Pol II, DNA and Thio/Mn^{2+} is critical for the inhibition, since order-of-addition is often informative on the mode of transcription inhibition (reviewed in chapter I), especially given the observation of Tipper et al. that order-of-addition was critical for thiolutin inhibition. Consistent with the early observation, we found that the template ssDNA binding protects Pol II from the Thio/Mn^{2+} inhibition (**Figure 3-11B**). Template ssDNA protection of Pol II generally suggests Thio/Mn^{2+} inhibits Pol II DNA interaction or similarly early step in transcription, consistent with the behavior of transcription initiation inhibitor. Therefore, we tested if Thio/Mn^{2+} affected Pol II elongation on a transcription bubble template with RNA primer, where transcription initiation is bypassed. Surprisingly, Thio/Mn^{2+} altered, but did not block Pol II transcription elongation, inducing a highly specific pause prone Pol II elongation mode (**Figure 3-11C, right panel**). This is unexpected because most specific initiation inhibitors do not cause elongation defects, and DNA binding does not protect RNAPs from elongation inhibitors. In addition, pre-assembled transcription elongation complexes were also resistant to Thio/Mn^{2+} (**Figure 3-11C, middle panel**), validating the critical order-of-addition from another assay (**Figure 3-11B**). Thio/Mn^{2+} inhibited Pol II was pause-prone and appeared to irreversibly arrest at specific template positions (**Figure 3-11C**). These results together suggested that thiolutin entirely blocks initiation, yet if initiation is bypassed, elongation is allowed though pause-prone.

Whether thiolutin inhibits transcription initiation or elongation *in vivo* is unclear[217,220]. In light of our *in vitro* results, we next tested thiolutin mediated Pol II transcription inhibition *in vivo*, by monitoring Pol II occupancy on a long gene *YLR454w*. We observed a specific decrease of Pol II occupancy on the 5' end of the gene within the first 2 mins, consistent with immediate transcription initiation inhibition after thiolutin treatment (**Figure 3-11D**). After 2 mins, we also observed a relatively slower loss of Pol II from the template compared to previous experiments where transcription was inhibited by other means[80]. Additionally, this decrease in Pol II occupancy at later time points does not appear polar as it would if elongation were proceeding normally. These results suggest that thiolutin may have additional non-immediate inhibitory effects on Pol II elongation *in vivo* (**Figure 3-11D**). Together, we conclude that thiolutin can inhibit Pol II initiation both *in vitro* and *in vivo* (**Figure 3-11**).

To investigate the nature of the inhibitory species, we first asked if reduced thiolutin interacts with Mn^{2+}. Interestingly, we observed continuous and reproducible changes of the UV spectra after Mn^{2+} is added to the reduced thiolutin (**Figure 3-12A**), suggesting changing chemical species in the reaction. The UV peak was shifted from the 340nm (reduced thiolutin) to around 380nm in the first two minutes, and the 380nm peak gradually decreased as a third species at 300nm started to accumulate (**Figure 3-12A**). The reaction reached a relative stable equilibrium after 20 minutes. Our data revealed the dynamic change and two major species after Mn^{2+} addition to the reduced thiolutin, but could not distinguish if either of the two was the inhibitory species. To test this question,

we prepared Thio/Mn^{2+} and either freshly treated Pol II or incubated at room temperature for 20 minutes before treating Pol II (**Figure 3-12B**). Remarkably, we observed strong Pol II inhibition by freshly prepared Thio/Mn^{2+} but complete loss of Thio/Mn^{2+} inhibitory activity after 20 minutes (**Figure 3-12B**). This result is consistent with the 380nm species possibly contributing to Pol II inhibition yet unstable in solution. In contrast to the window for inhibitory activity of Thio/Mn^{2+}, immediately treated Pol II was inhibited for up to 50 minutes (20-minute incubation and 30-minute reaction time), suggesting that the Pol II inhibition is stable (**Figure 3-12B**).

We propose two models to reconcile the observation that Pol II stabilized the unstable 380nm inhibitory species. First, Pol II might stabilize the inhibitory species in a tight pocket that prevented further reaction into other inactive species. Second, Mn^{2+} might coordinate with reduced thiolutin to form a disulfide bond within Pol II or possibly a thiolutin adduct, as Cu^{2+} does to facilitate disulfide formation through cysteine sulfenylation in other proteins[317-319]. If sulfenylation and disulfide bond formation were involved, the Thio/Mn^{2+} inhibition is expected to be suppressible by a high concentration of DTT. Indeed, we found that 50X excess of DTT fully abolished the inhibition, consistent with the involvement reductant sensitive inhibition (**Figure 3-12C**).

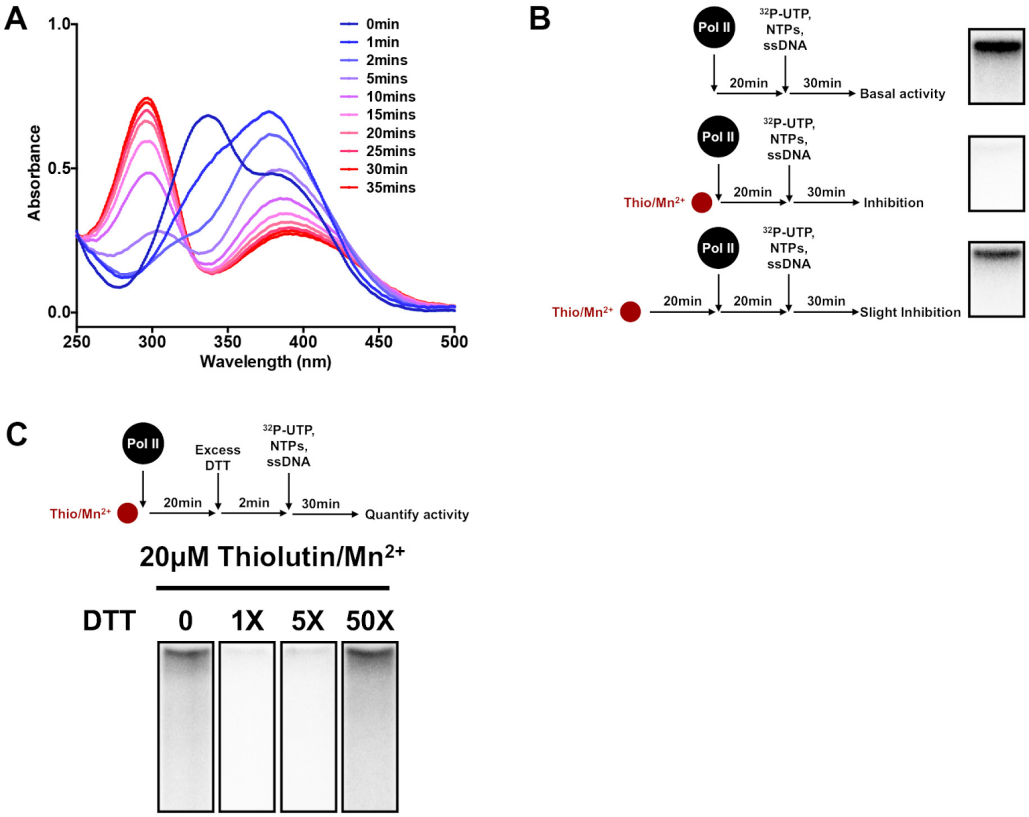


Figure 3-12. The apparent Thiolutin/Mn^{2+} complex is unstable in solution but stable in Pol II
(A) Time course of changing UV spectra of reduced thiolutin treated with Mn^{2+}. Reduced thiolutin (blue, 0 min) was treated with equivalent molar Mn^{2+}, and the UV spectra was acquired at different time points. Spectra over time was colored in a series of gradient from blue to red. Three experimental replicates were performed, and a representative replicate is shown.
(B) Thio/Mn^{2+} lost the inhibitory activity after 20 mins in solution but keeps Pol II inhibited. Two experimental replicates were performed and consistent. One replicate is shown.
(C) Thio/Mn^{2+} can be reversed by excess DTT. Excess of DTT was added after 20 minutes of Pol II inhibition by Thio/Mn^{2+}. Final DTT excess (relative to thiolutin) was indicated on the figure. At least three experimental replicates were performed, and a representative replicate is shown.

Discussion

Thiolutin is a routinely used transcription inhibitor for studying mRNA stability, but the exact mode of inhibition remained complicated and unresolved. It has been demonstrated that reduced thiolutin and holomycin chelated Zn^{2+} *in vitro*, and the thiolutin chelation of Zn^{2+} could specifically inhibit diverse classes of metalloproteins[213,214]. However, both reports failed to observe direct thiolutin inhibition of RNAPs, under conditions where Zn^{2+} chelation was permissible. It was interpreted as thiolutin and holomycin inhibiting transcription through secondary effects. Here we present three independent genetic screens for thiolutin resistant or sensitive mutants, providing a genetic basis for understanding thiolutin altered cellular responses and direct mode of action against Pol II *in vivo*. We show that alterations in diverse cellular pathways can modulate cellular sensitivity to thiolutin. In light of our genetic results, we discovered that both reductant DTT and Mn^{2+} together activate thiolutin direct inhibition of Pol II function, countering the recent narrative while upholding classic observations[214,216].

Thiolutin inhibits Pol II through a novel mode of action

We propose that Thio/Mn^{2+} inhibits Pol II through a novel mode of action. The critical order of treating Pol II with thiolutin prior to template DNA binding is a classical behavior for RNAP clamp inhibitors, which are distinct from inhibitors targeting the active site, NTP uptake or RNA exit channels (see Chapter I). In addition, we show that thiolutin inhibits Pol II elongation on an initiation-bypassing transcription bubble

template, distinct from all the three characterized inhibitors that lock the clamp in the closed state (Myx, Cor and Rip) and do not inhibit RNAP elongation[170,174]. It should be noted that Lpm, an inhibitor that appears to lock the clamp in the partially or fully open state[187], has never been tested on a bubble template. Therefore, thiolutin behaves differently from Myx, Cor and Rip, but further experiments are in need to test whether thiolutin behaves similarly to Lpm. Finally, the exact thiolutin binding site on Pol II remains unclear, and our data cannot rule out the possibility that thiolutin may inhibit Pol II regions other than clamp controlling switch regions.

Diverse cellular pathways modulate thiolutin multiple modes of action

Our genetic screens reveal contribution of diverse cellular pathways in thiolutin sensitivity. Many of these pathways can be reconciled in a unified model based on current understanding of thiolutin mode of action. In the cell, the thiolutin intra-molecular disulfide bond appears to be reduced, and we propose thioredoxins Trx1 and Trx2 directly or indirectly contribute to thiolutin reduction *in vivo*, while in turn being oxidized by this process. Reduced thiolutin may subsequently chelate Zn^{2+} to inhibit multiple metalloproteins (*e.g.* proteasome), affects Zn^{2+} homeostasis, interacts with Mn^{2+} to inhibit RNAP transcription, or get re-oxidized by molecular oxygen. The reduction or redox cycling of thiolutin may induce the observed Yap1 nuclear localization and oxidative stresses. In addition, evidence from us and others have shown that reduced thiolutin may also chelate other divalent metals such as Cu^{2+} or Ca^{2+} (but not Mg^{2+}), but

whether other thiolutin-metal complexes have additional activities remain to be further tested.

Thiolutin should not be used as a tool without caution

Our work underpins the caveats for using thiolutin to investigate other cellular process, as suggested by previous studies[213,214,223]. As a routinely used transcription inhibitor to study mRNA stability, it was shown that thiolutin itself inhibits mRNA degradation and complicates the quantitation of mRNA half-life at slightly higher dose[223]. In addition, studies from us and others reveal that thiolutin induces multiple cellular stress through inhibiting various targets, suggesting the lack of thiolutin specificity even at low dose. Unfortunately, there is currently no available specific and permeable Pol II inhibitors for yeast cells, and other approaches of Pol II inhibition appear to be complicated by other factors. For example, inhibiting Pol II transcription through Pol II temperature sensitive allele *rpb1-1* requires a temperature shift and a heat-shock response, and gene-specific transcription inhibition through glucose shutoff of a galactose inducible promoter appears to be complicated by the glucose repressing kinetics[80]. Before development of specific and permeable Pol II inhibitors, we suggest that multiple strategies for Pol II inhibition have to be taken to validate the observations in relevant studies.

Thiolutin mode of action may reveal insights into Pol II pausing

The thiolutin induced pause (and arrest) pattern appears to be highly specific and position-dependent, and it will be interesting to further investigate the pause sequence

preference and its possible connection to specific inhibited Pol II conformation. As discussed above, although remaining to be further tested, biochemical properties of thiolutin inhibited Pol II most closely resembles the clamp inhibition into a non-closed state, likely a partially or fully open state. It is tempting to hypothesize that the thiolutin induced pause prone Pol II is linked to a specific Pol II conformation, such as a clamp-opening state. This hypothesis is also consistent with the observations in *E.coli* RNAP that paused elongation complexes appear to correlate with open clamp and TL states, which can be reversed and possibly regulated by elongation factor RfaH[35,63]. It has been proposed that this RfaH (Spt5 in yeast) pause suppression through clamp may be a conserved regulatory mechanism for RNAP elongation in all domains of life[173]. Further characterization of thiolutin inhibited Pol II may reveal additional insights into this process in eukaryotes.

CHAPTER IV

SUMMARY AND FUTURE DIRECTIONS

Summary

The work described in this dissertation utilizes two entirely different approaches to investigate Pol II function. First, we developed a high-throughput phenotypic system to functionally dissect almost all TL single substitution mutants, providing a comprehensive view for the *in vivo* consequences of different Pol II TL perturbations (Chapter II). Some phenotypic consequences have already been connected to specific functional defects by our lab[31,52,80,244,266], and more are under investigation. These phenotypes allowed us to greatly expanded the membership of two previously identified mutant classes: the LOF mutants with decreased elongation rate and the GOF mutants with increased elongation rate. We have also identified a minor class of mutants with distinct phenotypes to be further investigated. In addition, we investigated the functional relationship between TL and several surrounding domains, and provide evidence consistent with complex functional interplay among these domains. Finally, our data suggest possible determinants for the incompatibility of Pol I TL in the Pol II context, likely through evolutionary divergence of the funnel helix surrounding TL.

Second, we took a chemical genetic and biochemical approach to investigate the mode of action of thiolutin, a natural product transcription inhibitor with unclear mode of transcription inhibition (Chapter III). Based on the observation that thiolutin inhibited

partially purified but not fully purified Pol II, we hypothesized that thiolutin inhibition of RNAPs required additional co-factors, and performed three independent genetic screens for yeast mutants that are resistant or sensitive to thiolutin. We characterized functional contribution of several cellular response pathways to thiolutin resistance. In addition, in light of our genetic data, we discovered surprising and stringent requirements for activation of thiolutin inhibition of Pol II: Mn^{2+} and appropriate levels of reductant DTT to reduce thiolutin while not antagonizing inhibition. We studied the nature of this Mn^{2+} and reductant-activated inhibitory species, and characterized the properties of the thiolutin inhibited Pol II. We showed that thiolutin mainly inhibited Pol II transcription initiation, and only inhibited free Pol II prior to template DNA binding. We further showed that thiolutin inhibited Pol II could elongate on an initiation bypassed transcription bubble template, but was prone to pausing and arrests. We suggest that thiolutin inhibits Pol II through a novel mechanism that is distinct from most other characterized transcription inhibitors (reviewed in Chapter I).

Each project advances our understanding of the Pol II function, and allows us to answer many critical open questions. The open questions and proposed future experiments are discussed below.

Future directions

Address functional relationship and evolutionary questions using the high-resolution phenotypic system

Our phenotypic system distinguishes TL mutants with similar general fitness defects on a large scale[19,31,266], thus providing the basis for studying several exciting and previously challenging questions.

First, the functional relationship among the residues within and surrounding the TL remains an open question. The TL residues can work collaboratively or independently to ensure proper TL function, as evidenced by TL mutations conferring diverse types of genetic interactions[31]. In addition, many TL interactions with the surrounding domains appear to be critical but largely untested[24,25,42,266]. Our phenotypic system is in place to explore these intra- and inter-TL genetic interactions.

To comprehensively evaluate the intra-TL genetic interactions, we propose to phenotype a carefully selected subset of double mutants. Based on the single mutant phenotypes, we will choose at least two substitutions for every residue and perform pairwise combinations for these substitutions. We expect most of the combinations to confer independent and activity-based additive effects, as suggested by previous studies[31,266]. We would also expect a subset of the combinations to confer intra-TL epistasis that suggest inter-dependent and non-additive functions for the residues. The non-additive intra-TL epistasis, as discussed above, has already contributed and will continue to

contribute to insights of collaborative TL residue functions[31]. Finally, we will also combine several well characterized single mutants (*e.g.* E1103G) with all possible single substitutions to specifically screen for intra-TL epistasis for these well-known mutants.

To study TL genetic interactions with the surrounding domains, we propose to phenotype the existing single substituted TL variants in the context of mutations in the surrounding domain. We expect the TL external mutations to alter the TL phenotypic landscape in a highly specific way, where both activity-based additive interactions and specific inter-domain epistasis can also be revealed. Together, our phenotypic system allows us to draw a semi-comprehensive epistasis map within and surrounding TL, and is expected to expand our current understanding of this complex functional network of residue-residue interactions.

Second, we propose to evaluate the complementation compatibilities of different evolutionary TLs in our Pol II system. The incompatibility of Pol I TL in the Pol II context revealed a critical and stringent requirement for the environments surrounding TL[56,266]. To understand the evolutionary divergence surrounding TL, we propose to comprehensively phenotype many more evolutionary TL variants in the Pol II context, and we propose to further dissect all the intermediate mutations for several compatible and incompatible evolutionary variants (we term as "evolutionary paths"). For incompatible evolutionary variants, this evolutionary path experiment allows us to pinpoint the causal mutations for incompatibility. For some compatible evolutionary

variants with potential detrimental mutations (*e.g.* Pol III TL discussed in chapter II), we expect to identify intra-TL epistasis within the evolutionary path that tolerate possible detrimental mutations. Together, we expect such experiments to be informative for our understanding of molecular evolution of the extremely highly conserved TL.

Third, we propose to test the potential roles of inter-residue co-evolution and functional coupling within the TL. Functional constraints among residues appear to drive co-evolution, which can be detected by an approach termed as statistical coupling analysis (SCA)[320-322]. Extensive work from Ranganathan lab have shown that the co-evolving residues form a physically continuous cluster of residues (termed as "sectors") and confer multiple functions in protein folding, stability, catalysis, substrate specificity, allostery and evolvability[236,320-326]. Despite the multiple functions, whether functional coupling contributes to the function of highly conserved domains remain difficult to test, partially due to the technical challenges to detect co-evolution within highly conserved domains and lack of sensitive functional assays to test them in a large scale.

With the accumulating sequences and our phenotypic system, we propose to test this previously difficult question using Pol II TL as a model system. The massive number of available sequences now allows us to assemble a diverse set of evolutionary TL variants (371 non-redundant TL sequences with 40-80% identity) and to detect co-evolution within the TL (**Figure 4-1, left panel**)[320,321]. We next performed two different previously reported scrambling approach to (1) specifically disrupt the co-evolution but preserve the

conservation (**Figure 4-1, middle panel**) or (2) preserve both conservation and coevolution (**Figure 4-1, right panel)** (See [323] for details). Each scrambling approach creates more than 350 artificial TL sequences for experimental validation. Our established phenotypic system distinguishes mutants with similar general growth defects and allows us to phenotype hundreds of mutants all at once. We propose to phenotype the natural TL sequences (as proposed above) along with the randomly scrambled (conservation preserved, co-evolution disrupted) and the Monte-Carlo-based scrambled (conservation and co-evolution preserved) libraries. We expect the phenotypic difference among these three libraries to reveal the potential role of co-evolution on the TL function.

Further explore the novel mode of action of thiolutin

Our observations in Chapter III suggest a novel mode of action for thiolutin dependent Pol II inhibition. However, several critical questions remain to be further clarified: 1. What is the structural basis of the thiolutin inhibition? 2. What is the nature of this inhibitory species when thiolutin is activated by DTT and Mn^{2+}? 3. Are there additional untested properties of the thiolutin inhibited Pol II? Each of these questions can be further divided into more specific questions, and we propose several critical experiments to address these questions.

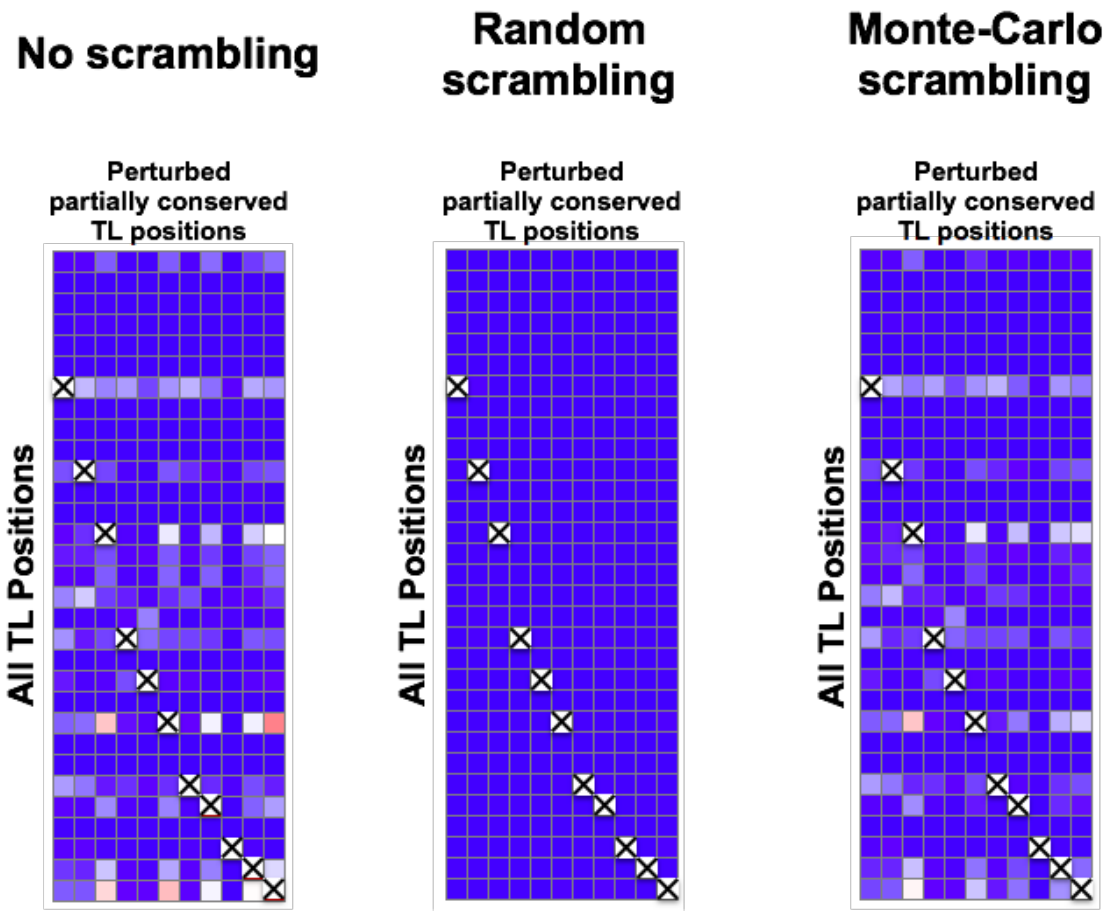


Figure 4-1. Intra-TL co-evolution detected by statistical coupling analysis (SCA).
A matrix showing the co-evolution scores between all 31 TL positions (rows) and 11 moderately conserved TL positions (columns). The co-evolution scores are computed by quantifying the changes of amino acid frequency at all TL position (rows) after sampling (perturbing) the moderately conserved TL positions (columns) for a specific amino acid, as previously reported[320]. The colors are coded as no co-evolution (blue), medium co-evolution (white) and maximal co-evolution (red). The self-correlation of positions to its own perturbation is crossed out from the matrix.

We propose to crystallize the thiolutin inhibited Pol II complex, which is expected to answer several major questions above, including the nature and the Pol II binding site of the inhibitory species. We also expect a thiolutin bound Pol II structure to reveal insights into the inhibition and the Pol II function. However, it should be noted that there are technical challenges in this experiment. First, the high sensitivity of thiolutin inhibition to excess reductants (DTT or TCEP) can pose a significant challenge for crystallography. Pol II crystallization is generally performed under conditions with high DTT concentration (5mM), which reverses thiolutin inhibition at relevant concentrations (Chapter III). In addition, the redox-cycling nature of thiolutin and presence of Mn^{2+} are known to facilitate DTT oxidation in the presence of molecular oxygen. Anaerobic crystallization conditions are likely needed for this experiment. In addition, the low solubility of thiolutin can be a limiting factor for crystallization. Thiolutin is only moderately soluble in DMSO (up to 4.4mM) but not in water or transcription buffer, thus limiting the concentration of thiolutin employed in the crystallization conditions, because high DMSO concentration is known to be detrimental to the Pol II crystal formation.

Given the potential challenges in crystallography, we proposed several parallel approaches to investigate the structural basis of thiolutin inhibition. First, as discussed in Chapter III, thiolutin may inhibit Pol II through forming covalent disulfide bond with the Pol II cysteines. A covalent inhibitor binding can be potentially detected with Pol II fragmentation followed by high resolution mass spectrometry. We are currently

collaborating with Amber Mosley's lab (at IUSM) to investigate this possibility. Second, as reviewed in Chapter III, thiolutin inhibition appears to most closely resemble the RNAP clamp inhibitors. The RNAP clamp is a highly mobile domain that can possibly move up to 20 angstroms[170], which is within the detection limit of cryo-EM experiments. We propose to test this potential inhibition of clamp motion using cryo-EM. Third, we propose to investigate the thiolutin inhibitory activity on RNA polymerases from other species, especially *E.coli*. Cross-species inhibitory activity is expected to reveal additional information on the conservation of thiolutin targeted region. Finally, if thiolutin similarly inhibits *E.coli* RNAP (to be tested), we could perform a genetic screen for *E.coli* RNAP mutants that confers resistance to thiolutin, as resistance pocket has been shown to be highly informative on the potential binding site of RNAP inhibitors (reviewed in Chapter I).

REFERENCES

1. Mason, P.B. & Struhl, K. Distinction and relationship between elongation rate and processivity of RNA polymerase II *in vivo*. *Molecular Cell* **17**, 831-840 (2005).

2. Thummel, C.S., Burtis, K.C. & Hogness, D.S. Spatial and temporal patterns of E74 transcription during *Drosophila* development. *Cell* **61**, 101-11 (1990).

3. Shermoen, A.W. & O'Farrell, P.H. Progression of the cell cycle through mitosis leads to abortion of nascent transcripts. *Cell* **67**, 303-310 (1991).

4. O'Brien, T. & Lis, J.T. Rapid changes in *Drosophila* transcription after an instantaneous heat shock. *Mol Cell Biol* **13**, 3456-63 (1993).

5. Yao, J., Ardehali, M.B., Fecko, C.J., Webb, W.W. & Lis, J.T. Intranuclear distribution and local dynamics of RNA polymerase II during transcription activation. *Mol Cell* **28**, 978-90 (2007).

6. Singh, J. & Padgett, R.A. Rates of in situ transcription and splicing in large human genes. *Nat Struct Mol Biol* **16**, 1128-33 (2009).

7. Darzacq, X. et al. In vivo dynamics of RNA polymerase II transcription. *Nat Struct Mol Biol* **14**, 796-806 (2007).

8. Danko, C.G. et al. Signaling pathways differentially affect RNA polymerase II initiation, pausing, and elongation rate in cells. *Mol Cell* **50**, 212-22 (2013).

9. Fuchs, G. et al. 4sUDRB-seq: measuring genomewide transcriptional elongation rates and initiation frequencies within cells. *Genome Biology* **15**, R69 (2014).

10. Jonkers, I., Kwak, H. & Lis, J.T. Genome-wide dynamics of Pol II elongation and its interplay with promoter proximal pausing, chromatin, and exons. *Elife* **3**, e02407 (2014).

11. Veloso, A. et al. Rate of elongation by RNA polymerase II is associated with specific gene features and epigenetic modifications. *Genome Research* **24**, 896-905 (2014).

12. Kwak, H. & Lis, J.T. Control of transcriptional elongation. *Annu Rev Genet* **47**, 483-508 (2013).

13. Jonkers, I. & Lis, J.T. Getting up to speed with transcription elongation by RNA polymerase II. *Nat Rev Mol Cell Biol* **16**, 167-77 (2015).

14. Imashimizu, M., Oshima, T., Lubkowska, L. & Kashlev, M. Direct assessment of transcription fidelity by high-resolution RNA sequencing. *Nucleic Acids Research* **41**, 9090-104 (2013).

15. Gout, J.F., Thomas, W.K., Smith, Z., Okamoto, K. & Lynch, M. Large-scale detection of in vivo transcription errors. *Proc Natl Acad Sci U S A* **110**, 18584-9 (2013).

16. Reid-Bayliss, K.S. & Loeb, L.A. Accurate RNA consensus sequencing for high-fidelity detection of transcriptional mutagenesis-induced epimutations. *Proc Natl Acad Sci U S A* **114**, 9415-9420 (2017).

17. Bentley, D.L. Coupling mRNA processing with transcription in time and space. *Nat Rev Genet* **15**, 163-75 (2014).

18. Larson, M.H. et al. Trigger loop dynamics mediate the balance between the transcriptional fidelity and speed of RNA polymerase II. *Proc Natl Acad Sci U S A* **109**, 6555-60 (2012).

19. Braberg, H. et al. From structure to systems: high-resolution, quantitative genetic analysis of RNA polymerase II. *Cell* **154**, 775-88 (2013).

20. Kaplan, C.D. Basic mechanisms of RNA polymerase II activity and alteration of gene expression in Saccharomyces cerevisiae. *Biochim Biophys Acta* **1829**, 39-54 (2013).

21. Liu, X., Bushnell, D.A. & Kornberg, R.D. RNA polymerase II transcription: structure and mechanism. *Biochim Biophys Acta* **1829**, 2-8 (2013).

22. Martinez-Rucobo, F.W. & Cramer, P. Structural basis of transcription elongation. *Biochim Biophys Acta* **1829**, 9-19 (2013).

23. Svetlov, V. & Nudler, E. Basic mechanism of transcription by RNA polymerase II. *Biochim Biophys Acta* **1829**, 20-8 (2013).

24. Wang, D., Bushnell, D.A., Westover, K.D., Kaplan, C.D. & Kornberg, R.D. Structural basis of transcription: role of the trigger loop in substrate specificity and catalysis. *Cell* **127**, 941-54 (2006).

25. Wang, B., Predeus, A.V., Burton, Z.F. & Feig, M. Energetic and structural details of the trigger-loop closing transition in RNA polymerase II. *Biophys J* **105**, 767-75 (2013).

26. Xu, L. et al. Dissecting the chemical interactions and substrate structural signatures governing RNA polymerase II trigger loop closure by synthetic nucleic acid analogues. *Nucleic Acids Research* **42**, 5863-70 (2014).

27. Vassylyev, D.G. et al. Structural basis for substrate loading in bacterial RNA polymerase. *Nature* **448**, 163-8 (2007).

28. Kaplan, C.D., Larsson, K.M. & Kornberg, R.D. The RNA polymerase II trigger loop functions in substrate selection and is directly targeted by alpha-amanitin. *Mol Cell* **30**, 547-56 (2008).

29. Kireeva, M.L. et al. Transient reversal of RNA polymerase II active site closing controls fidelity of transcription elongation. *Mol Cell* **30**, 557-66 (2008).

30. Huang, X. et al. RNA polymerase II trigger loop residues stabilize and position the incoming nucleotide triphosphate in transcription. *Proc Natl Acad Sci U S A* **107**, 15745-50 (2010).

31. Kaplan, C.D., Jin, H., Zhang, I.L. & Belyanin, A. Dissection of Pol II trigger loop function and Pol II activity-dependent control of start site selection in vivo. *PLoS Genet* **8**, e1002627 (2012).

32. Xu, L. et al. Strand-specific (asymmetric) contribution of phosphodiester linkages on RNA polymerase II transcriptional efficiency and fidelity. *Proc Natl Acad Sci U S A* **111**, E3269-76 (2014).

33. Hwang, C.S. et al. Functional interplay between NTP leaving group and base pair recognition during RNA polymerase II nucleotide incorporation revealed by methylene substitution. *Nucleic Acids Research* **44**, 3820-3828 (2016).

34. Malinen, A.M. et al. Active site opening and closure control translocation of multisubunit RNA polymerase. *Nucleic Acids Research* **40**, 7442-51 (2012).

35. Nayak, D., Voss, M., Windgassen, T., Mooney, R.A. & Landick, R. Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase. *Mol Cell* **50**, 882-93 (2013).

36. Da, L.T., Wang, D. & Huang, X. Dynamics of pyrophosphate ion release and its coupled trigger loop motion from closed to open state in RNA polymerase II. *J Am Chem Soc* **134**, 2399-406 (2012).

37. Liu, B., Zuo, Y. & Steitz, T.A. Structures of E. coli sigmaS-transcription initiation complexes provide new insights into polymerase mechanism. *Proc Natl Acad Sci U S A* **113**, 4051-6 (2016).

38. Toulokhonov, I., Zhang, J., Palangat, M. & Landick, R. A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol Cell* **27**, 406-19 (2007).

39. Seibold, S.A. et al. Conformational coupling, bridge helix dynamics and active site dehydration in catalysis by RNA polymerase. *Biochim Biophys Acta* **1799**, 575-87 (2010).

40. Silva, D.A. et al. Millisecond dynamics of RNA polymerase II translocation at atomic resolution. *Proc Natl Acad Sci U S A* **111**, 7665-70 (2014).

41. Windgassen, T.A. et al. Trigger-helix folding pathway and SI3 mediate catalysis and hairpin-stabilized pausing by *Escherichia coli* RNA polymerase. *Nucleic Acids Research* **42**, 12707-21 (2014).

42. Barnes, C.O. et al. Crystal Structure of a Transcribing RNA Polymerase II Complex Reveals a Complete Transcription Bubble. *Mol Cell* **59**, 258-69 (2015).

43. Weixlbaumer, A., Leon, K., Landick, R. & Darst, S.A. Structural basis of transcriptional pausing in bacteria. *Cell* **152**, 431-41 (2013).

44. Wang, D. et al. Structural basis of transcription: backtracked RNA polymerase II at 3.4 angstrom resolution. *Science* **324**, 1203-6 (2009).

45. Cheung, A.C. & Cramer, P. Structural basis of RNA polymerase II backtracking, arrest and reactivation. *Nature* **471**, 249-53 (2011).

46. Zhang, J., Palangat, M. & Landick, R. Role of the RNA polymerase trigger loop in catalysis and pausing. *Nat Struct Mol Biol* **17**, 99-104 (2010).

47. Yuzenkova, Y. & Zenkin, N. Central role of the RNA polymerase trigger loop in intrinsic RNA hydrolysis. *Proc Natl Acad Sci U S A* **107**, 10878-83 (2010).

48. Sosunova, E., Sosunov, V., Epshtein, V., Nikiforov, V. & Mustaev, A. Control of transcriptional fidelity by active center tuning as derived from RNA polymerase endonuclease reaction. *J Biol Chem* **288**, 6688-703 (2013).

49. Miropolskaya, N. et al. Interplay between the trigger loop and the F loop during RNA polymerase catalysis. *Nucleic Acids Research* **42**, 544-52 (2014).

50. Esyunina, D. et al. Lineage-specific variations in the trigger loop modulate RNA proofreading by bacterial RNA polymerases. *Nucleic Acids Research* **44**, 1298-308 (2016).

51. Castro, C. et al. Nucleic acid polymerases use a general acid for nucleotidyl transfer. *Nat Struct Mol Biol* **16**, 212-8 (2009).

52. Cabart, P., Jin, H., Li, L. & Kaplan, C.D. Activation and reactivation of the RNA polymerase II trigger loop for intrinsic RNA cleavage and catalysis. *Transcription* **5**, e28869 (2014).

53. Mishanina, T.V., Palo, M.Z., Nayak, D., Mooney, R.A. & Landick, R. Trigger loop of RNA polymerase is a positional, not acid–base, catalyst for both transcription and proofreading. *Proc Natl Acad Sci U S A*, 201702383 (2017).

54. Irvin, J.D. et al. A genetic assay for transcription errors reveals multilayer control of RNA polymerase II fidelity. *PLoS Genet* **10**, e1004532 (2014).

55. Dangkulwanich, M. et al. Complete dissection of transcription elongation reveals slow translocation of RNA polymerase II in a linear ratchet mechanism. *Elife* **2**, e00971 (2013).

56. Viktorovskaya, O.V. et al. Divergent contributions of conserved active site residues to transcription by eukaryotic RNA polymerases I and II. *Cell Rep* **4**, 974-84 (2013).

57. Malagon, F. et al. Mutations in the *Saccharomyces cerevisiae* RPB1 gene conferring hypersensitivity to 6-azauracil. *Genetics* **172**, 2201-9 (2006).

58. Trinh, V., Langelier, M.F., Archambault, J. & Coulombe, B. Structural perspective on mutations affecting the function of multisubunit RNA polymerases. *Microbiol Mol Biol Rev* **70**, 12-36 (2006).

59. Tan, L., Wiesler, S., Trzaska, D., Carney, H.C. & Weinzierl, R.O. Bridge helix and trigger loop perturbations generate superactive RNA polymerases. *J Biol* **7**, 40 (2008).

60. Weinzierl, R.O. The nucleotide addition cycle of RNA polymerase is controlled by two molecular hinges in the Bridge Helix domain. *BMC Biol* **8**, 134 (2010).

61. Miropolskaya, N., Artsimovitch, I., Klimasauskas, S., Nikiforov, V. & Kulbachinskiy, A. Allosteric control of catalysis by the F loop of RNA polymerase. *Proc Natl Acad Sci U S A* **106**, 18942-7 (2009).

62. Miropolskaya, N., Nikiforov, V., Klimasauskas, S., Artsimovitch, I. & Kulbachinskiy, A. Modulation of RNA polymerase activity through the trigger loop folding. *Transcription* **1**, 89-94 (2010).

63. Hein, P.P. et al. RNA polymerase pausing and nascent-RNA structure formation are linked through clamp-domain movement. *Nat Struct Mol Biol* **21**, 794-802 (2014).

64. Nesser, N.K., Peterson, D.O. & Hawley, D.K. RNA polymerase II subunit Rpb9 is important for transcriptional fidelity *in vivo*. *Proc Natl Acad Sci U S A* **103**, 3268-3273 (2006).

65. Walmacq, C. et al. Rpb9 subunit controls transcription fidelity by delaying NTP sequestration in RNA polymerase II. *J Biol Chem* **284**, 19601-12 (2009).

66. Kaster, B.C., Knippa, K.C., Kaplan, C.D. & Peterson, D.O. RNA Polymerase II Trigger Loop Mobility: Indirect effects of Rpb9. *J Biol Chem* **291**, 14883-95 (2016).

67. Kettenberger, H., Armache, K.J. & Cramer, P. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol Cell* **16**, 955-65 (2004).

68. Sobell, H.M., Jain, S.C., Sakore, T.D. & Nordman, C.E. Stereochemistry of actinomycin--DNA binding. *Nat New Biol* **231**, 200-5 (1971).

69. Kamitori, S. & Takusagawa, F. Crystal structure of the 2: 1 complex between d (GAAGCTTC) and the anticancer drug actinomycin D. *J Mol Biol* **225**, 445-456 (1992).

70. Chen, H., Liu, X. & Patel, D.J. DNA bending and unwinding associated with actinomycin D antibiotics bound to partially overlapping sites on DNA. *J Mol Biol* **258**, 457-79 (1996).

71. Kastan, M.B., Onyekwere, O., Sidransky, D., Vogelstein, B. & Craig, R.W. Participation of p53 protein in the cellular response to DNA damage. *Cancer Res* **51**, 6304-11 (1991).

72. Wadkins, R.M. & Jovin, T.M. Actinomycin D and 7-aminoactinomycin D binding to single-stranded DNA. *Biochemistry* **30**, 9469-78 (1991).

73. Nelson, W.G. & Kastan, M.B. DNA strand breaks: the DNA template alterations that trigger p53-dependent DNA damage response pathways. *Molecular and Cellular Biology* **14**, 1815-1823 (1994).

74. Zandomeni, R. & Weinmann, R. Inhibitory effect of 5,6-dichloro-1-beta-D-ribofuranosylbenzimidazole on a protein kinase. *J Biol Chem* **259**, 14804-11 (1984).

75. Zandomeni, R., Zandomeni, M.C., Shugar, D. & Weinmann, R. Casein kinase type II is involved in the inhibition by 5, 6-dichloro-1-beta-D-ribofuranosylbenzimidazole of specific RNA polymerase II transcription. *J Biol Chem* **261**, 3414-3419 (1986).

76. Exinger, F. & Lacroute, F. 6-Azauracil inhibition of GTP biosynthesis in *Saccharomyces cerevisiae*. *Curr Genet* **22**, 9-11 (1992).

77. Sintchak, M.D. et al. Structure and mechanism of inosine monophosphate dehydrogenase in complex with the immunosuppressant mycophenolic acid. *Cell* **85**, 921-30 (1996).

78. Titov, D.V. et al. XPB, a subunit of TFIIH, is a target of the natural product triptolide. *Nature Chemical Biology* **7**, 182-188 (2011).

79. Peterlin, B.M. & Price, D.H. Controlling the elongation phase of transcription with P-TEFb. *Mol Cell* **23**, 297-305 (2006).

80. Malik, I., Qiu, C., Snavely, T. & Kaplan, C.D. Wide-ranging and unexpected consequences of altered Pol II catalytic activity *in vivo*. *Nucleic Acids Research* **45**, 4431-4451 (2017).

81. Hyle, J.W., Shaw, R.J. & Reines, D. Functional distinctions between IMP dehydrogenase genes in providing mycophenolate resistance and guanine prototrophy to yeast. *J Biol Chem* **278**, 28470-8 (2003).

82. Jenks, M.H., O'Rourke, T.W. & Reines, D. Properties of an intergenic terminator and start site switch that regulate *IMD2* transcription in yeast. *Mol Cell Biol* **28**, 3883-93 (2008).

83. Kuehner, J.N. & Brow, D.A. Regulation of a eukaryotic gene by GTP-dependent start site selection and transcription attenuation. *Mol Cell* **31**, 201-11 (2008).

84. Darst, S.A. New inhibitors targeting bacterial RNA polymerase. *Trends Biochem Sci* **29**, 159-60 (2004).

85. Ma, C., Yang, X. & Lewis, P.J. Bacterial transcription as a target for antibacterial drug development. *Microbiology and Molecular Biology Reviews* **80**, 139-160 (2016).

86. Villain-Guillot, P., Bastide, L., Gualtieri, M. & Leonetti, J.P. Progress in targeting bacterial transcription. *Drug Discov Today* **12**, 200-8 (2007).

87. Gartel, A.L. Transcriptional inhibitors, p53 and apoptosis. *Biochimica et Biophysica Acta (BBA)-Reviews on Cancer* **1786**, 83-86 (2008).

88. Momand, J., Zambetti, G.P., Olson, D.C., George, D. & Levine, A.J. The mdm-2 oncogene product forms a complex with the p53 protein and inhibits p53-mediated transactivation. *Cell* **69**, 1237-45 (1992).

89. Juven-Gershon, T. & Oren, M. Mdm2: the ups and downs. *Mol Med* **5**, 71-83 (1999).

90. Momand, J., Wu, H.-H. & Dasgupta, G. MDM2—master regulator of the p53 tumor suppressor protein. *Gene* **242**, 15-29 (2000).

91. Michael, D. & Oren, M. The p53–Mdm2 module and the ubiquitin system. *Seminars in cancer biology* Vol. 13 49-58 (Elsevier, 2003).

92. Blagosklonny, M.V. Flavopiridol, an inhibitor of transcription: implications, problems and solutions. *Cell Cycle* **3**, 1537-42 (2004).

93. Demidenko, Z.N. & Blagosklonny, M.V. Flavopiridol induces p53 via initial inhibition of Mdm2 and p21 and, independently of p53, sensitizes apoptosis-reluctant cells to tumor necrosis factor. *Cancer Res* **64**, 3653-60 (2004).

94. Radhakrishnan, S.K. & Gartel, A.L. A novel transcriptional inhibitor induces apoptosis in tumor cells and exhibits antiangiogenic activity. *Cancer Res* **66**, 3264-70 (2006).

95. Lu, W., Chen, L., Peng, Y. & Chen, J. Activation of p53 by roscovitine-mediated suppression of MDM2 expression. *Oncogene* **20**, 3206-16 (2001).

96. O'hagan, H.M. & Ljungman, M. Nuclear accumulation of p53 following inhibition of transcription is not due to diminished levels of MDM2. *Oncogene* **23**, 5505 (2004).

97. Arima, Y. et al. Transcriptional blockade induces p53-dependent apoptosis associated with translocation of p53 to mitochondria. *J Biol Chem* **280**, 19166-76 (2005).

98. Drygin, D., Rice, W.G. & Grummt, I. The RNA polymerase I transcription machinery: an emerging target for the treatment of cancer. *Annu Rev Pharmacol Toxicol* **50**, 131-56 (2010).

99. Bywater, M.J., Pearson, R.B., McArthur, G.A. & Hannan, R.D. Dysregulation of the basal RNA polymerase transcription apparatus in cancer. *Nat Rev Cancer* **13**, 299-314 (2013).

100. Bywater, M.J. et al. Inhibition of RNA Polymerase I as a Therapeutic Strategy to Promote Cancer-Specific Activation of p53. *Cancer Cell* **22**, 51-65 (2012).

101. Andrews, W.J. et al. Old drug, new target: ellipticines selectively inhibit RNA polymerase I transcription. *J Biol Chem* **288**, 4567-82 (2013).

102. Fetherston, J., Werner, E. & Patterson, R. Processing of the external transcribed spacer of murine rRNA and site of action of actinomydn D. *Nucleic Acids Research* **12**, 7187-7198 (1984).

103. Rey, J.P., Scott, R. & Muller, H. Induction and removal of interstrand crosslinks in the ribosomal RNA genes of lymphoblastoid cell lines from patients with Fanconi anemia. *Mutat Res* **289**, 171-80 (1993).

104. Pondarre, C., Strumberg, D., Fujimori, A., Torres-Leon, R. & Pommier, Y. In vivo sequencing of camptothecin-induced topoisomerase I cleavage sites in human colon carcinoma cells. *Nucleic Acids Research* **25**, 4111-6 (1997).

105. Mahajan, P.B. Modulation of transcription of rRNA genes by rapamycin. *Int J Immunopharmacol* **16**, 711-21 (1994).

106. Ghoshal, K. & Jacob, S.T. An alternative molecular mechanism of action of 5-fluorouracil, a potent anticancer drug. *Biochem Pharmacol* **53**, 1569-75 (1997).

107. Jordan, P. & Carmo-Fonseca, M. Cisplatin inhibits synthesis of ribosomal RNA in vivo. *Nucleic Acids Research* **26**, 2831-6 (1998).

108. Rossini, C., Taylor, W., Fagan, T. & Hastings, J.W. Lifetimes of mRNAs for clock-regulated proteins in a dinoflagellate. *Chronobiol Int* **20**, 963-76 (2003).

109. Grigull, J., Mnaimneh, S., Pootoolal, J., Robinson, M.D. & Hughes, T.R. Genome-wide analysis of mRNA stability using transcription inhibitors and microarrays reveals posttranscriptional control of ribosome biogenesis factors. *Molecular and Cellular Biology* **24**, 5534-5547 (2004).

110. Kebaara, B.W., Nielsen, L.E., Nickerson, K.W. & Atkin, A.L. Determination of mRNA half-lives in *Candida albicans* using thiolutin as a transcription inhibitor. *Genome* **49**, 894-899 (2006).

111. Morey, J.S. & Van Dolah, F.M. Global analysis of mRNA half-lives and *de novo* transcription in a dinoflagellate, Karenia brevis. *PLoS One* **8**, e66347 (2013).

112. Archambault, J., Lacroute, F., Ruet, A. & Friesen, J. Genetic interaction between transcription elongation factor TFIIS and RNA polymerase II. *Molecular and Cellular Biology* **12**, 4142-4152 (1992).

113. Powell, W. & Reines, D. Mutations in the second largest subunit of RNA polymerase II cause 6-azauracil sensitivity in yeast and increased transcriptional arrest in vitro. *J Biol Chem* **271**, 6866-73 (1996).

114. Desmoucelles, C., Pinson, B., Saint-Marc, C. & Daignan-Fornier, B. Screening the yeast "disruptome" for mutants affecting resistance to the immunosuppressive drug, mycophenolic acid. *J Biol Chem* **277**, 27036-44 (2002).

115. Reines, D. Use of RNA yeast polymerase II mutants in studying transcription elongation. *Methods in Enzymology* **371**, 284-292 (2003).

116. Riles, L., Shaw, R.J., Johnston, M. & Reines, D. Large-scale screening of yeast mutants for sensitivity to the IMP dehydrogenase inhibitor 6-azauracil. *Yeast* **21**, 241-8 (2004).

117. Ezekiel, D.H. & Hutchins, J.E. Mutations affecting RNA polymerase associated with rifampicin resistance in *Escherichia coli*. *Nature* **220**, 276-7 (1968).

118. Khesin, R., Shemyakin, M., Gorlenko, Z.M., Mindlin, S. & Ilyina, T. Studies on the RNA polymerase in *Escherichia coli K12* using the mutation affecting its activity. *J Mol Biol* **42**, 401-411 (1969).

119. Rabussay, D. & Zillig, W. A rifampicin resistent RNA-polymerase from *E. coli* altered in the beta-subunit. *FEBS Lett* **5**, 104-106 (1969).

120. Austin, S. & Scaife, J. A new method for selecting RNA polymerase mutants. *J Mol Biol* **49**, 263-7 (1970).

121. McClure, W.R. & Cech, C.L. On the mechanism of rifampicin inhibition of RNA synthesis. *J Biol Chem* **253**, 8949-56 (1978).

122. Campbell, E.A. et al. Structural mechanism for rifampicin inhibition of bacterial RNA polymerase. *Cell* **104**, 901-12 (2001).

123. Jin, D.J. & Gross, C.A. Mapping and sequencing of mutations in the *Escherichia coli* rpoB gene that lead to rifampicin resistance. *J Mol Biol* **202**, 45-58 (1988).

124. Artsimovitch, I. et al. Allosteric modulation of the RNA polymerase catalytic reaction is an essential component of transcription control by rifamycins. *Cell* **122**, 351-363 (2005).

125. Artsimovitch, I. et al. Tagetitoxin inhibits RNA polymerase through trapping of the trigger loop. *J Biol Chem* **286**, 40395-40400 (2011).

126. Sevostyanova, A., Belogurov, G.A., Mooney, R.A., Landick, R. & Artsimovitch, I. The β subunit gate loop is required for RNA polymerase modification by RfaH and NusG. *Molecular cell* **43**, 253-262 (2011).

127. Klyuyev, S. & Vassylyev, D.G. The binding site and mechanism of the RNA polymerase inhibitor tagetitoxin: an issue open to debate. *Transcription* **3**, 46-50 (2012).

128. Temiakov, D. et al. Structural basis of transcription inhibition by antibiotic streptolydigin. *Mol Cell* **19**, 655-66 (2005).

129. Tuske, S. et al. Inhibition of bacterial RNA polymerase by streptolydigin: stabilization of a straight-bridge-helix active-center conformation. *Cell* **122**, 541-52 (2005).

130. Siddhikol, C., Erbstoeszer, J.W. & Weisblum, B. Mode of action of streptolydigin. *J Bacteriol* **99**, 151-5 (1969).

131. Cassani, G., Burgess, R.R., Goodman, H.M. & Gold, L. Inhibition of RNA polymerase by streptolydigin. *Nat New Biol* **230**, 197-200 (1971).

132. McClure, W.R. On the mechanism of streptolydigin inhibition of *Escherichia coli* RNA polymerase. *J Biol Chem* **255**, 1610-6 (1980).

133. Bushnell, D.A., Cramer, P. & Kornberg, R.D. Structural basis of transcription: α-Amanitin–RNA polymerase II cocrystal at 2.8 Å resolution. *Proc Natl Acad Sci U S A* **99**, 1218-1222 (2002).

134. Brueckner, F. & Cramer, P. Structural basis of transcription inhibition by alpha-amanitin and implications for RNA polymerase II translocation. *Nat Struct Mol Biol* **15**, 811-8 (2008).

135. Bartolomei, M.S. & Corden, J.L. Localization of an alpha-amanitin resistance mutation in the gene encoding the largest subunit of mouse RNA polymerase II. *Mol Cell Biol* **7**, 586-94 (1987).

136. Chen, Y., Weeks, J., Mortin, M.A. & Greenleaf, A.L. Mapping mutations in genes encoding the two large subunits of *Drosophila* RNA polymerase II defines domains essential for basic transcription functions and for proper expression of developmental genes. *Mol Cell Biol* **13**, 4214-22 (1993).

137. Bartolomei, M.S. & Corden, J.L. Clustered alpha-amanitin resistance mutations in mouse. *Mol Gen Genet* **246**, 778-82 (1995).

138. Izban, M.G. & Luse, D.S. The RNA polymerase II ternary complex cleaves the nascent transcript in a 3'----5'direction in the presence of elongation factor SII. *Genes Dev* **6**, 1342-1356 (1992).

139. Chafin, D.R., Guo, H. & Price, D.H. Action of alpha-amanitin during pyrophosphorolysis and elongation by RNA polymerase II. *J Biol Chem* **270**, 19114-9 (1995).

140. Rudd, M.D. & Luse, D.S. Amanitin greatly reduces the rate of transcription by RNA polymerase II ternary complexes but fails to inhibit some transcript cleavage modes. *J Biol Chem* **271**, 21549-21558 (1996).

141. Gong, X.Q., Nedialkov, Y.A. & Burton, Z.F. Alpha-amanitin blocks translocation by human RNA polymerase II. *J Biol Chem* **279**, 27422-7 (2004).

142. Miao, S. et al. Inhibition of bacterial RNA polymerases. Peptide metabolites from the cultures of *Streptomyces sp*. *Journal of Natural Products* **60**, 858-861 (1997).

143. Degen, D. et al. Transcription inhibition by the depsipeptide antibiotic salinamide A. *Elife* **3**, e02451 (2014).

144. Bae, B. et al. CBR antimicrobials inhibit RNA polymerase via at least two bridge-helix cap-mediated effects on nucleotide addition. *Proc Natl Acad Sci U S A* **112**, E4178-E4187 (2015).

145. Feng, Y. et al. Structural basis of transcription inhibition by CBR hydroxamidines and CBR pyrazoles. *Structure* **23**, 1470-1481 (2015).

146. Artsimovitch, I., Chu, C., Lynch, A.S. & Landick, R. A new class of bacterial RNA polymerase inhibitor affects nucleotide addition. *Science* **302**, 650-4 (2003).

147. Malinen, A.M. et al. CBR antimicrobials alter coupling between the bridge helix and the β subunit in RNA polymerase. *Nature Communications* **5**(2014).

148. Jovanovic, M. et al. Activity map of the *Escherichia coli* RNA polymerase bridge helix. *J Biol Chem* **286**, 14469-79 (2011).

149. Naji, S., Bertero, M.G., Spitalny, P., Cramer, P. & Thomm, M. Structure-function analysis of the RNA polymerase cleft loops elucidates initial transcription, DNA unwinding and RNA displacement. *Nucleic Acids Research* **36**, 676-87 (2008).

150. Lin, W. et al. Structural Basis of *Mycobacterium tuberculosis* transcription and transcription Inhibition. *Mol Cell* **66**, 169-179 e8 (2017).

151. Vassylyev, D.G., Vassylyeva, M.N., Perederina, A., Tahirov, T.H. & Artsimovitch, I. Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* **448**, 157-62 (2007).

152. Ho, M.X., Hudson, B.P., Das, K., Arnold, E. & Ebright, R.H. Structures of RNA polymerase-antibiotic complexes. *Curr Opin Struct Biol* **19**, 715-23 (2009).

153. Srivastava, A. et al. New target for inhibition of bacterial RNA polymerase:'switch region'. *Current Opinion in Microbiology* **14**, 532-543 (2011).

154. Gagneux, S. et al. The competitive cost of antibiotic resistance in *Mycobacterium tuberculosis*. *Science* **312**, 1944-6 (2006).

155. Comas, I. et al. Whole-genome sequencing of rifampicin-resistant *Mycobacterium tuberculosis* strains identifies compensatory mutations in RNA polymerase genes. *Nature Genetics* **44**, 106-110 (2012).

156. Hinkle, D.C. & Chamberlin, M.J. Studies of the binding of *Escherichia coli* RNA polymerase to DNA. II. The kinetics of the binding reaction. *J Mol Biol* **70**, 187-95 (1972).

157. Sippel, A.E. & Hartmann, G.R. Rifampicin resistance of RNA polymerase in the binary complex with DNA. *Eur J Biochem* **16**, 152-7 (1970).

158. Feklistov, A. et al. Rifamycins do not function by allosteric modulation of binding of Mg^{2+} to the RNA polymerase active center. *Proc Natl Acad Sci U S A* **105**, 14820-14825 (2008).

159. Campbell, E.A. et al. Structural, functional, and genetic analysis of sorangicin inhibition of bacterial RNA polymerase. *EMBO J* **24**, 674-82 (2005).

160. Irschik, H., Jansen, R., Gerth, K., Hofle, G. & Reichenbach, H. The sorangicins, novel and powerful inhibitors of eubacterial RNA polymerase isolated from myxobacteria. *J Antibiot (Tokyo)* **40**, 7-13 (1987).

161. Xu, M., Zhou, Y.N., Goldstein, B.P. & Jin, D.J. Cross-resistance of *Escherichia coli* RNA polymerases conferring rifampin resistance to different antibiotics. *J Bacteriol* **187**, 2783-92 (2005).

162. Ramaswamy, S. & Musser, J.M. Molecular genetic basis of antimicrobial agent resistance inMycobacterium tuberculosis: 1998 update. *Tubercle and Lung Disease* **79**, 3-29 (1998).

163. Rommele, G. et al. Resistance of *Escherichia coli* to rifampicin and sorangicin A--a comparison. *J Antibiot (Tokyo)* **43**, 88-91 (1990).

164. Oliva, B., O'Neill, A., Wilson, J.M., O'Hanlon, P.J. & Chopra, I. Antimicrobial properties and mode of action of the pyrrothine holomycin. *Antimicrob Agents Chemother* **45**, 532-9 (2001).

165. Das, K. et al. Roles of conformational and positional adaptability in structure-based design of TMC125-R165335 (etravirine) and related non-nucleoside reverse transcriptase inhibitors that are highly potent and effective against wild-type and drug-resistant HIV-1 variants. *Journal of Medicinal Chemistry* **47**, 2550-2560 (2004).

166. Cramer, P., Bushnell, D.A. & Kornberg, R.D. Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution. *Science* **292**, 1863-76 (2001).

167. Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A. & Kornberg, R.D. Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 A resolution. *Science* **292**, 1876-82 (2001).

168. Vassylyev, D.G. et al. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 A resolution. *Nature* **417**, 712-9 (2002).

169. Feklistov, A. et al. RNA polymerase motions during promoter melting. *Science* **356**, 863-866 (2017).

170. Chakraborty, A. et al. Opening and closing of the bacterial RNA polymerase clamp. *Science* **337**, 591-5 (2012).

171. Klein, B.J. et al. RNA polymerase and transcription elongation factor Spt4/5 complex structure. *Proc Natl Acad Sci U S A* **108**, 546-50 (2011).

172. Martinez-Rucobo, F.W., Sainsbury, S., Cheung, A.C. & Cramer, P. Architecture of the RNA polymerase-Spt4/5 complex and basis of universal transcription processivity. *EMBO J* **30**, 1302-10 (2011).

173. Werner, F. A nexus for gene expression-molecular mechanisms of Spt5 and NusG in the three domains of life. *J Mol Biol* **417**, 13-27 (2012).

174. Mukhopadhyay, J. et al. The RNA polymerase "switch region" is a target for inhibitors. *Cell* **135**, 295-307 (2008).

175. Belogurov, G.A. et al. Transcription inactivation through local refolding of the RNA polymerase structure. *Nature* **457**, 332 (2009).

176. Molodtsov, V. et al. X-ray crystal structures of Escherichia coli RNA polymerase with switch region binding inhibitors enable rational design of squaramides with an improved fraction unbound to human plasma protein. *J Med Chem* **58**, 3156-71 (2015).

177. Buurman, E.T. et al. Novel rapidly diversifiable antimicrobial RNA polymerase switch region inhibitors with confirmed mode of action in Haemophilus influenzae. *J Bacteriol* **194**, 5504-12 (2012).

178. Coronelli, C., White, R.J., Lancini, G.C. & Parenti, F. Lipiarmycin, a new antibiotic from Actinoplanes. II. Isolation, chemical, biological and biochemical characterization. *J Antibiot (Tokyo)* **28**, 253-9 (1975).

179. Parenti, F., Pagani, H. & Beretta, G. Lipiarmycin, a new antibiotic from Actinoplanes. I. Description of the producer strain and fermentation studies. *J Antibiot (Tokyo)* **28**, 247-52 (1975).

180. Sergio, S., Pirali, G., White, R. & Parenti, F. Lipiarmycin, a new antibiotic from Actinoplanes III. Mechanism of action. *J Antibiot (Tokyo)* **28**, 543-9 (1975).

181. Talpaert, M., Campagnari, F. & Clerici, L. Lipiarmycin: an antibiotic inhibiting nucleic acid polymerases. *Biochem Biophys Res Commun* **63**, 328-34 (1975).

182. Omura, S. et al. Clostomicins, new antibiotics produced by Micromonospora echinospora subsp. armeniaca subsp. nov. I. Production, isolation, and physico-chemical and biological properties. *J Antibiot (Tokyo)* **39**, 1407-12 (1986).

183. Arnone, A., Nasini, G. & Cavalleri, B. Structure elucidation of the macrocyclic antibiotic lipiarmycin. *Journal of the Chemical Society, Perkin Transactions 1*, 1353-1359 (1987).

184. Hochlowski, J.E. et al. Tiacumicins, a novel complex of 18-membered macrolides. II. Isolation and structure determination. *J Antibiot (Tokyo)* **40**, 575-88 (1987).

185. Theriault, R.J. et al. Tiacumicins, a novel complex of 18-membered macrolide antibiotics. I. Taxonomy, fermentation and antibacterial activity. *J Antibiot (Tokyo)* **40**, 567-74 (1987).

186. Tupin, A., Gualtieri, M., Leonetti, J.P. & Brodolin, K. The transcription inhibitor lipiarmycin blocks DNA fitting into the RNA polymerase catalytic site. *EMBO J* **29**, 2527-2537 (2010).

187. Wang, D. Ensemble fluorescence resonance energy transfer analysis of RNA polymerase clamp conformation, (Rutgers The State University of New Jersey-New Brunswick, 2008).

188. Sonenshein, A.L. & Alexander, H.B. Initiation of transcription in vitro is inhibited by lipiarmycin. *J Mol Biol* **127**, 55-72 (1979).

189. Ma, C., Yang, X. & Lewis, P.J. Bacterial Transcription Inhibitor of RNA Polymerase Holoenzyme Formation by Structure-Based Drug Design: From in Silico Screening to Validation. *ACS Infect Dis* **2**, 39-46 (2016).

190. Vassylyev, D.G. et al. Structural basis for transcription inhibition by tagetitoxin. *Nat Struct Mol Biol* **12**, 1086-93 (2005).

191. Blond, A. et al. The cyclic structure of microcin J25, a 21-residue peptide antibiotic from *Escherichia coli*. *Eur J Biochem* **259**, 747-55 (1999).

192. Bayro, M.J. et al. Structure of antibacterial peptide microcin J25: a 21-residue lariat protoknot. *J Am Chem Soc* **125**, 12382-3 (2003).

193. Rosengren, K.J. et al. Microcin J25 has a threaded sidechain-to-backbone ring structure and not a head-to-tail cyclized backbone. *Journal of the American Chemical Society* **125**, 12464-12474 (2003).

194. Wilson, K.-A. et al. Structure of microcin J25, a peptide inhibitor of bacterial RNA polymerase, is a lassoed tail. *Journal of the American Chemical Society* **125**, 12475-12483 (2003).

195. Delgado, M.A., Rintoul, M.a.R., Farías, R.N. & Salomón, R.A. *Escherichia coli* RNA polymerase is the target of the cyclopeptide antibiotic microcin J25. *J Bacteriol* **183**, 4543-4550 (2001).

196. Yuzenkova, J. et al. Mutations of bacterial RNA polymerase leading to resistance to microcin j25. *J Biol Chem* **277**, 50867-75 (2002).

197. Adelman, K. et al. Molecular mechanism of transcription inhibition by peptide antibiotic Microcin J25. *Mol Cell* **14**, 753-62 (2004).

198. Mukhopadhyay, J., Sineva, E., Knight, J., Levy, R.M. & Ebright, R.H. Antibacterial peptide microcin J25 inhibits transcription by binding within and obstructing the RNA polymerase secondary channel. *Molecular Cell* **14**, 739-751 (2004).

199. Rintoul, M.R., de Arcuri, B.F., Salomon, R.A., Farias, R.N. & Morero, R.D. The antibacterial action of microcin J25: evidence for disruption of cytoplasmic membrane energization in *Salmonella newport*. *FEMS Microbiol Lett* **204**, 265-70 (2001).

200. Bellomio, A., Vincent, P.A., de Arcuri, B.F., Farías, R.N. & Morero, R.D. Microcin J25 has dual and independent mechanisms of action in *Escherichia coli*: RNA polymerase inhibition and increased superoxide production. *J Bacteriol* **189**, 4180-4186 (2007).

201. Niklison Chirou, M. et al. Microcin J25 induces the opening of the mitochondrial transition pore and cytochrome c release through superoxide generation. *The FEBS Journal* **275**, 4088-4096 (2008).

202. Niklison-Chirou, M.V. et al. Microcin J25 triggers cytochrome c release through irreversible damage of mitochondrial proteins and lipids. *The International Journal of Biochemistry & Cell Biology* **42**, 273-281 (2010).

203. Walker, S.S. et al. Affinity Selection-Mass Spectrometry Identifies a Novel Antibacterial RNA Polymerase Inhibitor. *ACS Chem Biol* **12**, 1346-1352 (2017).

204. Mathews, D.E. & Durbin, R.D. Mechanistic aspects of tagetitoxin inhibition of RNA polymerase from Escherichia coli. *Biochemistry* **33**, 11987-92 (1994).

205. Sakai, A., Saito, C., Inada, N. & Kuroiwa, T. Transcriptional activities of the chloroplast-nuclei and proplastid-nuclei isolated from tobacco exhibit different sensitivities to tagetitoxin: implication of the presence of distinct RNA polymerases. *Plant Cell Physiol* **39**, 928-34 (1998).

206. Steinberg, T.H., Mathews, D.E., Durbin, R.D. & Burgess, R.R. Tagetitoxin: a new inhibitor of eukaryotic transcription by RNA polymerase III. *J Biol Chem* **265**, 499-505 (1990).

207. Mathews, D.E. & Durbin, R.D. Tagetitoxin inhibits RNA synthesis directed by RNA polymerases from chloroplasts and *Escherichia coli*. *J Biol Chem* **265**, 493-8 (1990).

208. Yuzenkova, Y., Roghanian, M., Bochkareva, A. & Zenkin, N. Tagetitoxin inhibits transcription by stabilizing pre-translocated state of the elongation complex. *Nucleic Acids Research* **41**, 9257-9265 (2013).

209. Aliev, A.E., Karu, K., Mitchell, R.E. & Porter, M.J. The structure of tagetitoxin. *Org Biomol Chem* **14**, 238-45 (2016).

210. Li, B., Wever, W.J., Walsh, C.T. & Bowers, A.A. Dithiolopyrrolones: biosynthesis, synthesis, and activity of a unique class of disulfide-containing antibiotics. *Nat Prod Rep* **31**, 905-23 (2014).

211. Li, B. & Walsh, C.T. Identification of the gene cluster for the dithiolopyrrolone antibiotic holomycin in *Streptomyces clavuligerus*. *Proc Natl Acad Sci U S A* **107**, 19731-5 (2010).

212. Li, B. & Walsh, C.T. *Streptomyces clavuligerus* HlmI is an intramolecular disulfide-forming dithiol oxidase in holomycin biosynthesis. *Biochemistry* **50**, 4615-4622 (2011).

213. Chan, A.N. et al. Role for dithiolopyrrolones in disrupting bacterial metal homeostasis. *Proc Natl Acad Sci U S A* **114**, 2717-2722 (2017).

214. Lauinger, L. et al. Thiolutin is a zinc chelator that inhibits the Rpn11 and other JAMM metalloproteases. *Nature Chemical Biology* **13**, 709-714 (2017).

215. Jimenez, A., Tipper, D.J. & Davies, J. Mode of action of thiolutin, an inhibitor of macromolecular synthesis in *Saccharomyces cerevisiae*. *Antimicrob Agents Chemother* **3**, 729-38 (1973).

216. Tipper, D.J. Inhibition of yeast ribonucleic acid polymerases by thiolutin. *J Bacteriol* **116**, 245-56 (1973).

217. Khachatourians, G.G. & Tipper, D.J. *In vivo* effect of thiolutin on cell growth and macromolecular synthesis in *Escherichia coli*. *Antimicrob Agents Chemother* **6**, 304-10 (1974).

218. Das, B., Butler, J.S. & Sherman, F. Degradation of normal mRNA in the nucleus of *Saccharomyces cerevisiae*. *Mol Cell Biol* **23**, 5502-15 (2003).

219. Guan, Q. et al. Impact of nonsense-mediated mRNA decay on the global expression profile of budding yeast. *PLoS Genet* **2**, e203 (2006).

220. Sivasubramanian, N. & Jayaraman, R. Thiolutin resistant mutants of *Escherichia coli* are they RNA chain initiation mutants? *Molecular and General Genetics MGG* **145**, 89-96 (1976).

221. Roza, J., Blanco, M.G., Hardisson, C. & Salas, J.A. Self-resistance in actinomycetes producing inhibitors of RNA polymerase. *J Antibiot (Tokyo)* **39**, 609-12 (1986).

222. Joshi, A., Verma, M. & Chakravorty, M. Thiolutin-resistant mutants of *Salmonella typhimurium*. *Antimicrob Agents Chemother* **22**, 541-7 (1982).

223. Pelechano, V. & Perez-Ortin, J.E. The transcriptional inhibitor thiolutin blocks mRNA degradation in yeast. *Yeast* **25**, 85-92 (2008).

224. Hazelbaker, D.Z., Marquardt, S., Wlotzka, W. & Buratowski, S. Kinetic competition between RNA Polymerase II and Sen1-dependent transcription termination. *Molecular Cell* **49**, 55-66 (2013).

225. Baell, J.B. Redox-active nuisance screening compounds and their classification. *Drug Discovery Today* **16**, 840-841 (2011).

226. Johnston, P.A. Redox cycling compounds generate H_2O_2 in HTS buffers containing strong reducing reagents—real hits or promiscuous artifacts? *Current Opinion in Chemical Biology* **15**, 174-182 (2011).

227. Bergmann, R. Thiolutin inhibits utilization of glucose and other carbon sources in cells of *Escherichia coli*. *Antonie Van Leeuwenhoek* **55**, 143-52 (1989).

228. Westover, K.D., Bushnell, D.A. & Kornberg, R.D. Structural basis of transcription: nucleotide selection by rotation in the RNA polymerase II active center. *Cell* **119**, 481-9 (2004).

229. Yuzenkova, Y. et al. Stepwise mechanism for transcription fidelity. *BMC Biol* **8**, 54 (2010).

230. Fouqueau, T., Zeller, M.E., Cheung, A.C., Cramer, P. & Thomm, M. The RNA polymerase trigger loop functions in all three phases of the transcription cycle. *Nucleic Acids Research* **41**, 7048-59 (2013).

231. Fong, N. et al. Pre-mRNA splicing is facilitated by an optimal RNA polymerase II elongation rate. *Genes Dev* **28**, 2663-76 (2014).

232. Lennon, C.W. et al. Direct interactions between the coiled-coil tip of DksA and the trigger loop of RNA polymerase mediate transcriptional regulation. *Genes Dev* **26**, 2634-46 (2012).

233. Sekine, S., Murayama, Y., Svetlov, V., Nudler, E. & Yokoyama, S. The ratcheted and ratchetable structural states of RNA polymerase underlie multiple transcriptional functions. *Mol Cell* **57**, 408-21 (2015).

234. Mejia, Y.X., Nudler, E. & Bustamante, C. Trigger loop folding determines transcription rate of *Escherichia coli*'s RNA polymerase. *Proc Natl Acad Sci U S A* **112**, 743-8 (2015).

235. Fowler, D.M. & Fields, S. Deep mutational scanning: a new style of protein science. *Nat Methods* **11**, 801-7 (2014).

236. McLaughlin, R.N., Jr., Poelwijk, F.J., Raman, A., Gosal, W.S. & Ranganathan, R. The spatial architecture of protein function and adaptation. *Nature* **491**, 138-42 (2012).

237. Araya, C.L. et al. A fundamental protein property, thermodynamic stability, revealed solely from large-scale measurements of protein function. *Proc Natl Acad Sci U S A* **109**, 16858-63 (2012).

238. Kim, I., Miller, C.R., Young, D.L. & Fields, S. High-throughput analysis of *in vivo* protein stability. *Mol Cell Proteomics* **12**, 3370-8 (2013).

239. Zhai, W. et al. Synthetic antibodies designed on natural sequence landscapes. *J Mol Biol* **412**, 55-71 (2011).

240. Van den Brulle, J. et al. A novel solid phase technology for high-throughput gene synthesis. *Biotechniques* **45**, 340-3 (2008).

241. Tewhey, R. et al. Microdroplet-based PCR enrichment for large-scale targeted sequencing. *Nat Biotechnol* **27**, 1025-31 (2009).

242. Williams, R. et al. Amplification of complex gene libraries by emulsion PCR. *Nat Methods* **3**, 545-50 (2006).

243. Simchen, G., Winston, F., Styles, C.A. & Fink, G.R. Ty-mediated gene expression of the *LYS2* and *HIS4* genes of *Saccharomyces cerevisiae* is controlled by the same SPT genes. *Proc Natl Acad Sci U S A* **81**, 2431-4 (1984).

244. Cui, P., Jin, H., Vutukuru, M.R. & Kaplan, C.D. Relationships between RNA polymerase II activity and Spt elongation factors to Spt$^-$ phenotype and growth in *Saccharomyces cerevisiae*. *G3 (Bethesda)* **6**, 2489-504 (2016).

245. Greger, I.H. & Proudfoot, N.J. Poly(A) signals control both transcriptional termination and initiation between the tandem *GAL10* and *GAL7* genes of *Saccharomyces cerevisiae*. *EMBO J* **17**, 4771-9 (1998).

246. Kaplan, C.D., Holland, M.J. & Winston, F. Interaction between transcription elongation factors and mRNA 3'-end formation at the *Saccharomyces cerevisiae GAL10-GAL7* locus. *J Biol Chem* **280**, 913-22 (2005).

247. Waterhouse, A.M., Procter, J.B., Martin, D.M., Clamp, M. & Barton, G.J. Jalview Version 2--a multiple sequence alignment editor and analysis workbench. *Bioinformatics* **25**, 1189-91 (2009).

248. Niyogi, S.K., Feldman, R.P. & Hoffman, D.J. Selective effects of metal ions on RNA synthesis rates. *Toxicology* **22**, 9-21 (1981).

249. Niyogi, S.K. & Feldman, R.P. Effect of several metal ions on misincorporation during transcription. *Nucleic Acids Research* **9**, 2615-27 (1981).

250. Dyson, H.J., Wright, P.E. & Scheraga, H.A. The role of hydrophobic interactions in initiation and propagation of protein folding. *Proc Natl Acad Sci U S A* **103**, 13057-61 (2006).

251. Cheung, A.C., Sainsbury, S. & Cramer, P. Structural basis of initial RNA polymerase II transcription. *EMBO J* **30**, 4755-63 (2011).

252. Rudolph, H.K. et al. The yeast secretory pathway is perturbed by mutations in *PMR1*, a member of a Ca^{2+} ATPase family. *Cell* **58**, 133-45 (1989).

253. Mandal, D., Woolf, T.B. & Rao, R. Manganese selectivity of *pmr1*, the yeast secretory pathway ion pump, is defined by residue gln783 in transmembrane segment 6. Residue Asp778 is essential for cation transport. *J Biol Chem* **275**, 23933-8 (2000).

254. Hirata, A., Klein, B.J. & Murakami, K.S. The X-ray crystal structure of RNA polymerase from Archaea. *Nature* **451**, 851-4 (2008).

255. Da, L.T. et al. Bridge helix bending promotes RNA polymerase II backtracking through a critical and conserved threonine residue. *Nat Commun* **7**, 11244 (2016).

256. Engel, C., Sainsbury, S., Cheung, A.C., Kostrewa, D. & Cramer, P. RNA polymerase I structure and transcription regulation. *Nature* **502**, 650-5 (2013).

257. Opalka, N. et al. Structure and function of the transcription elongation factor GreB bound to bacterial RNA polymerase. *Cell* **114**, 335-45 (2003).

258. de la Mata, M. et al. A slow RNA polymerase II affects alternative splicing *in vivo*. *Mol Cell* **12**, 525-32 (2003).

259. Coulter, D.E. & Greenleaf, A.L. A mutation in the largest subunit of RNA polymerase II alters RNA chain elongation in vitro. *J Biol Chem* **260**, 13190-8 (1985).

260. Koyama, H., Ueda, T., Ito, T. & Sekimizu, K. Novel RNA polymerase II mutation suppresses transcriptional fidelity and oxidative stress sensitivity in rpb9Delta yeast. *Genes Cells* **15**, 151-9 (2010).

261. Archambault, J. et al. Stimulation of transcription by mutations affecting conserved regions of RNA polymerase II. *J Bacteriol* **180**, 2590-8 (1998).

262. Arndt, K.T., Styles, C.A. & Fink, G.R. A suppressor of a *HIS4* transcriptional defect encodes a protein with homology to the catalytic subunit of protein phosphatases. *Cell* **56**, 527-37 (1989).

263. Walmacq, C. et al. Mechanism of translesion transcription by RNA polymerase II and its role in cellular resistance to DNA damage. *Mol Cell* **46**, 18-29 (2012).

264. Jin, H. & Kaplan, C.D. Relationships of RNA polymerase II genetic interactors to transcription start site usage defects and growth in *Saccharomyces cerevisiae*. *G3 (Bethesda)* **5**, 21-33 (2014).

265. Winston, F., Dollard, C. & Ricupero-Hovasse, S.L. Construction of a set of convenient *Saccharomyces cerevisiae* strains that are isogenic to S288C. *Yeast* **11**, 53-5 (1995).

266. Qiu, C. et al. High-Resolution Phenotypic landscape of the RNA Polymerase II trigger loop. *PLoS Genet* **12**, e1006321 (2016).

267. Puig, O. et al. The tandem affinity purification (TAP) method: a general procedure of protein complex purification. *Methods* **24**, 218-29 (2001).

268. Schmitt, M.E., Brown, T.A. & Trumpower, B.L. A rapid and simple method for preparation of RNA from *Saccharomyces cerevisiae*. *Nucleic Acids Research* **18**, 3091-2 (1990).

269. Ranish, J.A. & Hahn, S. The yeast general transcription factor TFIIA is composed of two polypeptide subunits. *J Biol Chem* **266**, 19320-7 (1991).

270. de Hoon, M.J., Imoto, S., Nolan, J. & Miyano, S. Open source clustering software. *Bioinformatics* **20**, 1453-4 (2004).

271. Li, W., Jaroszewski, L. & Godzik, A. Tolerating some redundancy significantly speeds up clustering of large protein databases. *Bioinformatics* **18**, 77-82 (2002).

272. Edgar, R.C. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Research* **32**, 1792-7 (2004).

273. Monje-Casas, F., MICHáN, C. & Pueyo, C. Absolute transcript levels of thioredoxin-and glutathione-dependent redox systems in *Saccharomyces cerevisiae*: response to stress and modulation with growth. *Biochemical Journal* **383**, 139-147 (2004).

274. Smith, A.M. et al. Quantitative phenotyping via deep barcode sequencing. *Genome Res* **19**, 1836-42 (2009).

275. Lee, A.Y. et al. Mapping the cellular response to small molecules using chemogenomic fitness signatures. *Science* **344**, 208-211 (2014).

276. Parsons, A.B. et al. Exploring the mode-of-action of bioactive compounds by chemical-genetic profiling in yeast. *Cell* **126**, 611-25 (2006).

277. Janke, C. et al. A versatile toolbox for PCR-based tagging of yeast genes: new fluorescent proteins, more markers and promoter substitution cassettes. *Yeast* **21**, 947-62 (2004).

278. Chee, M.K. & Haase, S.B. New and redesigned pRS plasmid shuttle vectors for genetic manipulation of *Saccharomyces cerevisiae*. *G3: Genes, Genomes, Genetics* **2**, 515-526 (2012).

279. Huang, Z. et al. A functional variomics tool for discovering drug-resistance genes and drug targets. *Cell Rep* **3**, 577-85 (2013).

280. Robinson, D.G., Chen, W., Storey, J.D. & Gresham, D. Design and analysis of Bar-seq experiments. *G3 (Bethesda)* **4**, 11-8 (2014).

281. Birkeland, S.R. et al. Discovery of mutations in *Saccharomyces cerevisiae* by pooled linkage analysis and whole-genome sequencing. *Genetics* **186**, 1127-37 (2010).

282. Kuge, S., Jones, N. & Nomoto, A. Regulation of yAP-1 nuclear localization in response to oxidative stress. *EMBO J* **16**, 1710-20 (1997).

283. Wemmie, J.A., Steggerda, S.M. & Moye-Rowley, W.S. The *Saccharomyces cerevisiae* AP-1 protein discriminates between oxidative stress elicited by the oxidants H_2O_2 and diamide. *J Biol Chem* **272**, 7908-7914 (1997).

284. Coleman, S.T., Epping, E.A., Steggerda, S.M. & Moye-Rowley, W.S. Yap1p activates gene transcription in an oxidant-specific fashion. *Mol Cell Biol* **19**, 8302-13 (1999).

285. Kuge, S. et al. Regulation of the yeast Yap1p nuclear export signal is mediated by redox signal-induced reversible disulfide bond formation. *Molecular and Cellular Biology* **21**, 6139-6150 (2001).

286. Gulshan, K., Rovinsky, S.A., Coleman, S.T. & Moye-Rowley, W.S. Oxidant-specific folding of Yap1p regulates both transcriptional activation and nuclear localization. *J Biol Chem* **280**, 40524-40533 (2005).

287. Alarco, A.M., Balan, I., Talibi, D., Mainville, N. & Raymond, M. AP1-mediated multidrug resistance in Saccharomyces cerevisiae requires FLR1 encoding a transporter of the major facilitator superfamily. *J Biol Chem* **272**, 19304-13 (1997).

288. Coleman, S.T., Tseng, E. & Moye-Rowley, W.S. *Saccharomyces cerevisiae* basic region-leucine zipper protein regulatory networks converge at the *ATR1* structural gene. *J Biol Chem* **272**, 23224-23230 (1997).

289. Gulshan, K. & Moye-Rowley, W.S. Multidrug resistance in fungi. *Eukaryot Cell* **6**, 1933-42 (2007).

290. Meyer, Y., Buchanan, B.B., Vignols, F. & Reichheld, J.P. Thioredoxins and glutaredoxins: unifying elements in redox biology. *Annu Rev Genet* **43**, 335-67 (2009).

291. Winzeler, E.A. et al. Functional characterization of the *S. cerevisiae* genome by gene deletion and parallel analysis. *Science* **285**, 901-6 (1999).

292. Machado, A.K., Morgan, B.A. & Merrill, G.F. Thioredoxin reductase-dependent inhibition of MCB cell cycle box activity in *Saccharomyces cerevisiae*. *J Biol Chem* **272**, 17045-54 (1997).

293. Carmel-Harel, O. et al. Role of thioredoxin reductase in the Yap1p-dependent response to oxidative stress in *Saccharomyces cerevisiae*. *Mol Microbiol* **39**, 595-605 (2001).

294. Trotter, E.W. & Grant, C.M. Thioredoxins are required for protection against a reductive stress in the yeast *Saccharomyces cerevisiae*. *Mol Microbiol* **46**, 869-78 (2002).

295. Hacioglu, E., Esmer, I., Fomenko, D.E., Gladyshev, V.N. & Koc, A. The roles of thiol oxidoreductases in yeast replicative aging. *Mech Ageing Dev* **131**, 692-9 (2010).

296. Ragu, S. et al. Loss of the thioredoxin reductase Trr1 suppresses the genomic instability of peroxiredoxin *tsa1* mutants. *PLoS One* **9**, e108123 (2014).

297. Grant, C.M., Perrone, G. & Dawes, I.W. Glutathione and catalase provide overlapping defenses for protection against hydrogen peroxide in the yeast *Saccharomyces cerevisiae*. *Biochemical and Biophysical Research Communications* **253**, 893-898 (1998).

298. Castro, F.A., Mariani, D., Panek, A.D., Eleutherio, E.C. & Pereira, M.D. Cytotoxicity mechanism of two naphthoquinones (menadione and plumbagin) in *Saccharomyces cerevisiae*. *PLoS One* **3**, e3999 (2008).

299. Toledano, M.B., Delaunay-Moisan, A., Outten, C.E. & Igbaria, A. Functions and cellular compartmentation of the thioredoxin and glutathione pathways in yeast. *Antioxid Redox Signal* **18**, 1699-711 (2013).

300. Jung, U., Zheng, X., Yoon, S.O. & Chung, A.S. Se-methylselenocysteine induces apoptosis mediated by reactive oxygen species in HL-60 cells. *Free Radic Biol Med* **31**, 479-89 (2001).

301. Ravi, D. & Das, K.C. Redox-cycling of anthracyclines by thioredoxin system: increased superoxide generation and DNA damage. *Cancer Chemother Pharmacol* **54**, 449-58 (2004).

302. Bandara, P.D., Flattery-O'Brien, J.A., Grant, C.M. & Dawes, I.W. Involvement of the *Saccharomyces cerevisiae UTH1* gene in the oxidative-stress response. *Curr Genet* **34**, 259-68 (1998).

303. Wu, C.-Y., Bird, A.J., Winge, D.R. & Eide, D.J. Regulation of the yeast *TSA1* peroxiredoxin by *ZAP1* is an adaptive response to the oxidative stress of zinc deficiency. *J Biol Chem* **282**, 2184-2195 (2007).

304. Wu, C.-Y. et al. Repression of sulfate assimilation is an adaptive response of yeast to the oxidative stress of zinc deficiency. *J Biol Chem* **284**, 27544-27556 (2009).

305. Wu, C.-Y., Steffen, J. & Eide, D.J. Cytosolic superoxide dismutase (*SOD1*) is critical for tolerating the oxidative stress of zinc deficiency in yeast. *PloS one* **4**, e7061 (2009).

306. Acosta-Alvear, D. et al. Paradoxical resistance of multiple myeloma to proteasome inhibitors by decreased levels of 19S proteasomal subunits. *Elife* **4**, e08153 (2015).

307. Tsvetkov, P. et al. Compromising the 19S proteasome complex protects cells from reduced flux through the proteasome. *Elife* **4**, e08467 (2015).

308. Tsvetkov, P. et al. Suppression of 19S proteasome subunits marks emergence of an altered cell state in diverse cancers. *Proc Natl Acad Sci U S A* **114**, 382-387 (2017).

309. D'Aurora, V., Stern, A.M. & Sigman, D.S. 1,10-Phenanthroline-cuprous ion complex, a potent inhibitor of DNA and RNA polymerases. *Biochem Biophys Res Commun* **80**, 1025-32 (1978).

310. Perrin, D.M., Pearson, L., Mazumder, A. & Sigman, D.S. Inhibition of prokaryotic and eukaryotic transcription by the 2:1 2,9-dimethyl-1,10-phenanthroline-cuprous complex, a ligand specific for open complexes. *Gene* **149**, 173-8 (1994).

311. Reimann, C.W., Block, S. & Perloff, A. The Crystal and molecular structure of Dichloro (1, 10-phenanthroline) zinc. *Inorganic Chemistry* **5**, 1185-1189 (1966).

312. Lin, S.-J. & Culotta, V.C. Suppression of oxidative damage by *Saccharomyces cerevisiae ATX2*, which encodes a manganese-trafficking protein that localizes to Golgi-like vesicles. *Molecular and Cellular Biology* **16**, 6303-6312 (1996).

313. Portnoy, M.E., Liu, X.F. & Culotta, V.C. *Saccharomyces cerevisiae* expresses three functionally distinct homologues of the nramp family of metal transporters. *Mol Cell Biol* **20**, 7893-902 (2000).

314. Jensen, L.T., Ajua-Alemanji, M. & Culotta, V.C. The *Saccharomyces cerevisiae* high affinity phosphate transporter encoded by *PHO84* also functions in manganese homeostasis. *J Biol Chem* **278**, 42036-42040 (2003).

315. Mei, Y., Jensen, L.T., Gardner, A.J. & Culotta, V.C. Manganese toxicity and *Saccharomyces cerevisiae* Mam3p, a member of the ACDP (ancient conserved domain protein) family. *Biochemical Journal* **386**, 479-487 (2005).

316. Reddi, A.R. et al. The overlapping roles of manganese and Cu/Zn SOD in oxidative stress protection. *Free Radic Biol Med* **46**, 154-62 (2009).

317. Pecci, L., Montefoschi, G., Musci, G. & Cavallini, D. Novel findings on the copper catalysed oxidation of cysteine. *Amino Acids* **13**, 355-367 (1997).

318. Prudent, M. & Girault, H.H. The role of copper in cysteine oxidation: study of intra- and inter-molecular reactions in mass spectrometry. *Metallomics* **1**, 157-65 (2009).

319. Fetherolf, M.M. et al. Copper-zinc superoxide dismutase is activated through a sulfenic acid intermediate at a copper ion entry site. *J Biol Chem* **292**, 12025-12040 (2017).

320. Lockless, S.W. & Ranganathan, R. Evolutionarily conserved pathways of energetic connectivity in protein families. *Science* **286**, 295-9 (1999).

321. Suel, G.M., Lockless, S.W., Wall, M.A. & Ranganathan, R. Evolutionarily conserved networks of residues mediate allosteric communication in proteins. *Nat Struct Biol* **10**, 59-69 (2003).

322. Halabi, N., Rivoire, O., Leibler, S. & Ranganathan, R. Protein sectors: evolutionary units of three-dimensional structure. *Cell* **138**, 774-86 (2009).

323. Russ, W.P., Lowery, D.M., Mishra, P., Yaffe, M.B. & Ranganathan, R. Natural-like function in artificial WW domains. *Nature* **437**, 579-83 (2005).

324. Reynolds, K.A., McLaughlin, R.N. & Ranganathan, R. Hot spots for allosteric regulation on protein surfaces. *Cell* **147**, 1564-75 (2011).

325. Stiffler, M.A., Hekstra, D.R. & Ranganathan, R. Evolvability as a function of purifying selection in TEM-1 β-lactamase. *Cell* **160**, 882-892 (2015).

326. Raman, A.S., White, K.I. & Ranganathan, R. Origins of allostery and evolvability in proteins: A Case Study. *Cell* **166**, 468-480 (2016).