IDENTIFICATION OF CANDIDATE NUCLEASE MUTANTS IN RNA POLYMERASE II

by

SASKIA BYERLY NEHER

A THESIS

Presented to the Department of Chemistry
and the Honors College at the University
of Oregon in partial fulfillment of the requirements
for the degree of
Bachelor of Science

May 2000

Approved : [redacted]

Diane K. Hawley

An Abstract of the Thesis of

Saskia Byerly Neher for the degree of Bachelor of Science

in the Department of Chemistry to be taken June 2000

Title: IDENTIFICATION OF CANDIDATE NUCLEASE MUTANTS IN RNA POLYMERASE II

Approved: [redacted]

Diane K. Hawley

Current understanding of the process by which RNA polymerase II transcribes a DNA template is incomplete. Many unanswered questions remain regarding the elongation and termination stages of transcription. These processes could be further elucidated with a greater understanding of a nuclease activity known to be intrinsic to the polymerase. To address these questions, we have designed and executed a genetic screen with the goal of identifying the active site of nuclease activity in the polymerase. The basis of our screen is the expectation that strains of the yeast *Saccharomyces cerevisiae* with polymerases defective in nuclease activity will be unable to grow in the presence of cordycepin, a chain terminating nucleoside analog.

We report the identification of five candidate nuclease mutants in RPB2, the second largest subunit of RNA polymerase II. Each candidate mutant has one or more amino acid changes in RPB2 that may cause the cordycepin-sensitive phenotype. Mapping these mutations to the tertiary structure of an RNA polymerase reveals that some fall into regions previously reported to affect nuclease activity.

ACKNOWLEDGMENTS

I wish to express sincere appreciation to Professor Diane Hawley for her guidance in my research and in the preparation of this manuscript. Thanks are also due to members of the Hawley lab, especially Matt Miller, for their patience and wisdom. Additionally, I would like to thank the Sprague lab for their generous contribution of advice and materials.

TABLE OF CONTENTS

LIST OF FIGURES

I. INTRODUCTION

In an event often called the birth of modern biology, James Watson and Francis Crick built a model for the structure of DNA based on the experimental data of their contemporaries. The arrangement they proposed, a double helix, not only explained how DNA could carry genetic information but also how it could replicate. This discovery, coupled with the later elucidation of the process by which DNA is transcribed to RNA, and RNA translated to protein, became Crick's "central dogma" of molecular biology. An enormous amount of scientific research is now devoted to understanding how genetic information is correctly expressed. Areas of research such as gene regulation take on an increasing importance as we learn that many human disorders can be attributed to abnormalities in genetic processes. Such studies are often built on basic research that aims to identify and understand components involved in the chain of gene expression.

While an organism's DNA is a repository for its genetic information, RNA is the intermediary between this information and its use in molecules such as enzymes. DNA-dependent RNA polymerase is the enzyme responsible for transcription, the first step in the expression of a gene. The focus of this work is an investigation of a particular feature of RNA polymerase II, the enzyme that synthesizes mRNA used as protein templates. This feature is a nuclease activity known to be intrinsic to the enzyme, but of unknown location. The thesis describes a genetic screen for the catalytic site of nuclease activity in RNA polymerase II. It begins with a discussion of background information and motives for undertaking the screen before experiments and results are described.

Transcription

A gene is a piece of DNA that contains information encoding a specific RNA molecule or protein, as well as information about how its *expression* is regulated. Before a gene can be expressed as a protein, it must be copied to a similar molecule, RNA. Genes which encode information for proteins are copied from DNA to RNA by a large, multisubunit enzyme known as RNA polymerase II. This process, called *transcription*, occurs in three stages. In initiation, the polymerase, with the help of other proteins, recognizes a specific DNA sequence called a promoter located at the beginning of the gene. The polymerase binds to the DNA and begins to incorporate RNA nucleotides to synthesize a chain that is *complimentary* to the DNA strand. Once the first few RNA molecules have been incorporated, the initiation phase of transcription gives way to elongation. In this stage, RNA nucleotides are processively added to the growing RNA chain until the polymerase terminates. Upon termination, the nascent RNA is released and the polymerase disassociates from the DNA template.

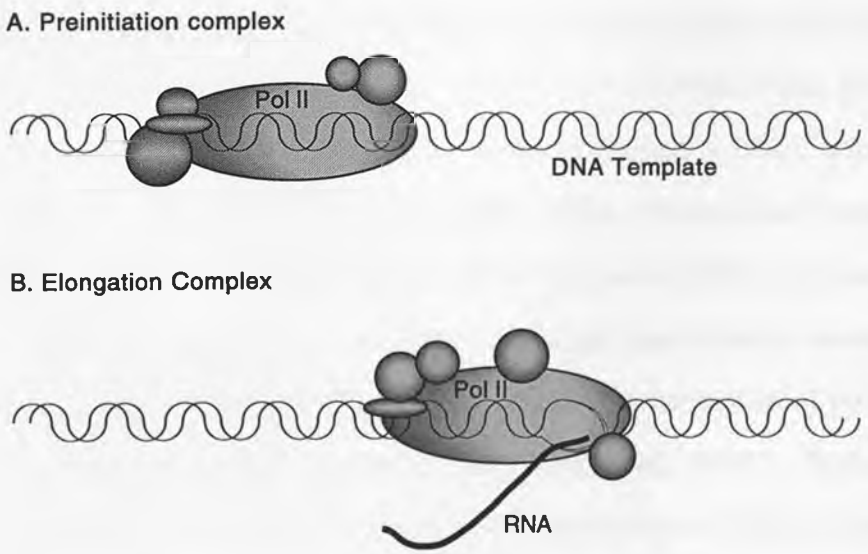


Figure 1: RNA polymerase initiates transcription after the formation of the preinitiation complex.

A variety of protein factors can bind to sequence elements in the promoter to influence the probability that a gene will be transcribed. As such, studies on the regulation of transcription have traditionally focused on the initiation stage. However, regulation during elongation is well established in prokaryotes, with examples such as attenuation in bacteria and anti-termination in lamba phage (Yanofsky, 1988, Yang et al., 1989). Although regulation during elongation is not as well characterized in eukaryotes, several recent studies have shown that elongation may be an important in regulation. Genes proposed to be regulated at the transcriptional level include Drosophila heat shock genes, the human adenosine deaminase gene, and proto-oncogenes such as c-myc, c-fos, c-myb and c-mos (reviewed in Kerppola and Kane, 1991, and Shilatifard, 1998. Pinaud and Mirkovitch, 1998). These studies underscore the importance of further elucidation of the elongation properties of RNA polymerase II.

Arrest/SII Rescue

One way to better understand cellular processes is to isolate and identify the components and reconstitute the process outside of the cell. These experiments are termed *in vitro* assays. Transcription can occur in *in vitro* assays when the necessary components, including a DNA template, RNA polymerase, ribonucleotides and other cellular factors are included. Such assays reveal that during transcription elongation RNA polymerase can encounter certain sequences on the DNA template that hinder continued transcription. These blocks to elongation include transcriptional pauses, arrest, and termination sites (reviewed in Uptain et al., 1997.) Pausing and arrest differ from termination in that the polymerase and RNA chain remain bound to the template and are still catalytically active. Arrested complexes are of particular interest because, unlike paused complexes, they require the

addition of specific proteins in order to continue elongation. Without these factors arrested complexes will remain stuck on the template, unable to continue chain elongation.

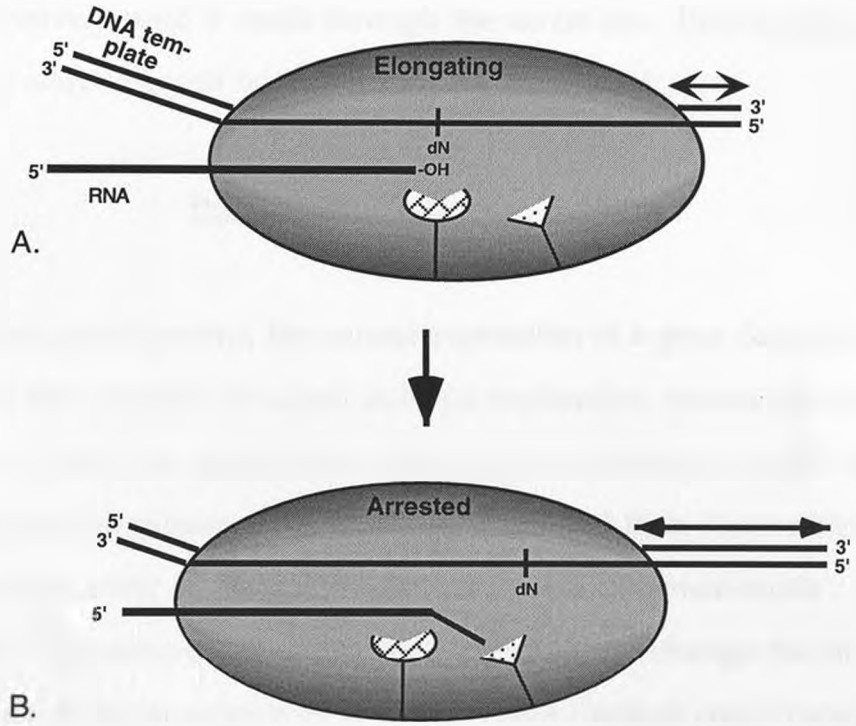


Figure 2: A model for transcriptional arrest. In part A, the 3' end of the RNA transcript is positioned correctly for polymerization. In an arrested complex (part B) the 3' end of the RNA is displaced.

It has been shown by *in vitro* transcription assays at well characterized arrest sites that a transcription elongation factor known as SII enables arrested polymerases to continue elongation (Wiest et al., 1992). Additional studies have shown that SII stimulates a cleavage activity intrinsic to the polymerase (Reines, 1992, Izban and Luse, 1992, Wang and Hawley, 1993). This activity removes nucleosides from the 3' end of the transcript which helps to rescue the stalled complex. One model for arrest suggests that the 3' end of the RNA becomes improperly positioned for continued elongation (Hawley et al., 1993). The SII stimulated cleavage activity would remove this end, allowing the

proper alignment of the end of the RNA chain with the active site. This gives the polymerase another opportunity to elongate through the arrest site. Because arrest sites aren't 100% efficient, the polymerase elongates and cleaves iteratively until it reads through the arrest site. Polymerases deficient in nuclease activity could be utilized to test this model.

Proofreading/Nuclease Activity

In biological systems, the correct expression of a gene depends upon the accuracy of the enzymes involved in DNA replication, transcription and translation. While the mechanisms affecting the accuracy of DNA replication and translation have been more extensively studied than those affecting transcription, fidelity at the transcriptional level is of considerable importance. Transcriptional errors have the ability to change the structure and function of the proteins they encode. These changes could range from an innocuous silent mutation, to an energetically costly nonfunctional protein, to a protein with an altered, deleterious function. Because each mRNA is translated multiple times, an incorrect transcript could lead to the production of many defective proteins.

The mechanisms responsible for accuracy in DNA polymerases have been extensively studied in prokaryotic systems and provide important background for understanding the fidelity of RNA polymerases. DNA replication occurs with a very low rate of error, which is achieved in three ways. These include selection of the correct dNTP substrate, cleavage of an incorrect substrate from the end of the growing chain, and excision of incorrectly inserted bases after the polymerase has finished extending the DNA chain (Echols and Goodman, 1991.) The first two mechanisms are of

particular relevance in understanding the accuracy of the RNA polymerase and will be discussed more fully.

A remarkable feature of the A-T and G-C "Watson-Crick" base pairs is their geometric equivalence. An A-T pair can be exchanged for a G-C pair without distorting the double helix because the C1' bond angles and distances in the sugar-phosphate backbone are equivalent for both pairs. Any other combination of bases, such as a G-T base pair, would reorient the sugar-phosphate backbone causing distortion of the double helix. Echols and Goodman suggest that the demand for equivalence geometry of the Watson-Crick base pairs is a major component in correct base selection for DNA polymerases. It is thought that the enzyme active site imposes geometric constraints favoring Watson-Crick over non-Watson-Crick base pairs (Kuchta et al., 1988). Additionally, there is an energetic advantage for correct over incorrect incoming base due to hydrogen bonding and base stacking interactions. RNA polymerases are thought to select the correct base by similar methods.

The second mechanism used by many DNA polymerases for maintaining fidelity is a nuclease activity involved in proofreading of the newly added nucleotide. Many DNA polymerases have an *active site* for polymerization as well as an active site for nuclease activity (Wang et al., 1996, Besse and Steitz, 1991, Steitz, 1999). There is a balance between the polymerization and nuclease functions that generally favors the addition of nucleotides. This is due to the fast rate of polymerization relative to that of transferof the growing end of the DNA chain to the nuclease site (Patel et al., 1991). However, the rate of extension from a mismatched base is slowed allowing transfer of the DNA to the nuclease site (Wong et al., 1991). This model for proofreading is termed kinetic partitioning, and allows for the

cleavage of mismatched bases without indiscriminate removal of correct bases (Donlin et al., 1991).

The discovery of a 3' to 5' nuclease activity in RNA polymerase raised the possibility that a similar mechanism could promote the fidelity of transcription (Izban and Luse, 1992, Reines, 1992, Wang and Hawley, 1993). Subsequently, it was demonstrated that the nuclease activity could act to remove misincorporated bases in the presence of transcription factor SII (Jeon and Agarwal, 1996). *In vitro* studies showed that the polymerase could not only discriminate between correct and incorrect bases, but also remove incorrect bases rapidly enough to permit proofreading during steady state transcription (Thomas et al., 1998).

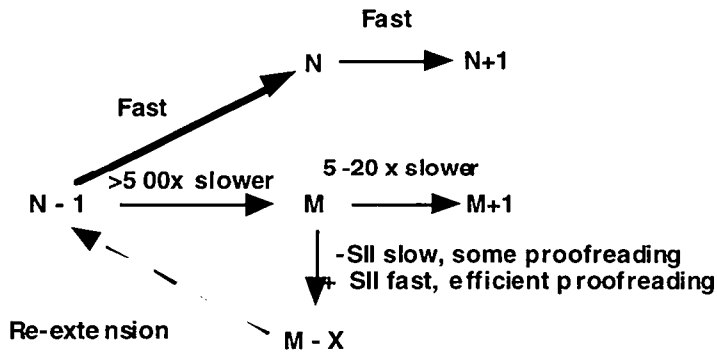


Figure 3: A model for kinetic partitioning in RNA polymerase II adapted from Thomas et al., 1998. When the elongation complex is at postition N-1, it can add the next correct nucleotide, N, or a misincorporation (M). After a misincorporation, the addition of the next nucleotide is slowed, allowing time for cleavage and re-extension.

These studies suggest that the nuclease activity intrinsic to RNA polymerase II could, in conjunction with transcription factor SII, constitute a proofreading function for transcription similar to that observed in DNA replication. However, transcriptional proofreading has not been demonstrated *in vivo*. For DNA replication in prokaryotes, mutator

polymerases correlated with decreased nuclease activity and antimutator polymerases correlated with increased nuclease activity (reviewed in Reha-Krantz, 1998). Identification and characterization of nuclease mutants in RNA polymerase II would allow us to make a similar correlation. A system for the identification of accuracy mutants in the *E. coli* RNA polymerase has been described (Blank et al., 1986) which could be adapted for use in *S. cerevisiae.* A goal of this project is to isolate potential mutants on the basis of cordycepin sensitivity for future *in vitro* and *in vivo* characterization of proofreading activity.

CHAPTER II: A NUCLEASE ACTIVITY SCREEN

Introduction

Variants of RNA Polymerase II lacking nuclease activity have the potential to further our understanding of many aspects of transcription. However, such mutants have as of yet proved elusive. Attempts to identify specific protein residues involved in nuclease activity based on sequence similarity to bacterial ribonucleases such as barnase were unsuccessful (Shirai and Go, 1991, and Powell et al, 1997). A group of altered termination mutants in *ret1* of RNA Polymerase III were tested for nuclease activity. Most exceeded the wild type level of activity, and none were lacking nuclease activity (Bobkova et al., 1999). Use of screening agents has proved equally unfruitful. A series of mutations were identified on the basis of increased sensitivity to 6-azauracil, a drug that reduces cellular GTP and UTP levels (Scafe et al., 1990). In *in vitro* assays, these mutants exhibited normal nuclease activity (Powell and Reines, 1996).

To our knowledge, no screen for nuclease activity has as of yet been undertaken using the nucleoside analog cordycepin. This drug is similar to the nucleoside adenosine, except that it lacks the 3' OH group that is chemically essential to elongation. Because of the nature and directionality of nucleotide chain elongation, cordycepin's addition to the 3' end of RNA will prevent further chain elongation. Yeast with wild type nuclease activity will be able to remove the cordycepin and continue RNA chain elongation. Thus, it is our expectation the cordycepin will be an effective way to identify nuclease mutants, as these mutants should be able to grow on normal media,

but not on media that contains cordycepin. Cordycepin's inhibitory effect on RNA chain growth is shown in figure 4.

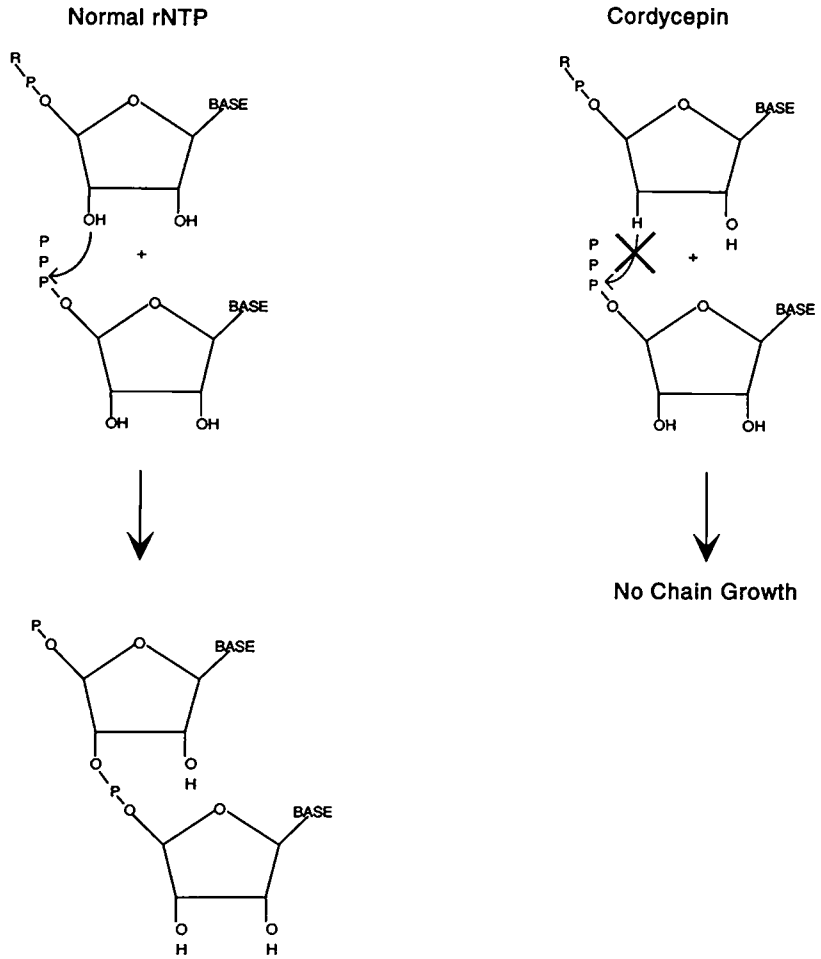


Figure 4: Cordycepin terminates the growing RNA chain unless removed.

Many previous studies have been done on uptake of cordycepin by yeast cells and its incorporation into mRNA, mostly in the context of polyadenylation. Importantly, it has been established that cordycepin can be converted into its triphosphate analog and incorporated into mRNA (Rose et al, 1997). Additional studies have shown that the vitamin thiamin, when present in the growth media, prevents the uptake of cordycepin by yeast

(Iwashima at al, 1992). Media without thiamin is available commercially, making use of cordycepin an effective method to screen for nuclease activity in the polymerase.

We elected to conduct the genetic screen in the first and second largest subunits of RNA pol II, known as RBP1 and RBP2, respectively. Screening in these two subunits should ensure the highest probability of identifying nuclease mutants for a variety of reasons. Chief among these is the sequence conservation seen in the first and second largest subunits in many polymerases known to have nuclease activity. These include the other eukaryotic DNA dependent RNA polymerases and the ß and ß' subunits of the bacterial RNA polymerase (Powell et al., 1997). Additionally, a recent crystal structure of the *Thermus aquaticus* RNA polymerase shows that RPB1 and RPB2 constitute the catalytic center of the enzyme (Zhang et al., 1999). Indeed, many mutations that affect transcriptional pausing, elongation, and termination map to these two subunits (Archambault and Friesen, 1993). Included among these are a series of mutations in the second largest subunit that cause increased arrest *in vitro* as well as increased sensitivity to the drug 6-azauracil (Powell and Reines, 1996). While these mutations do not affect cleavage activity *in vitro*, it has been suggested that the nucleotide binding region, polymerase active site and nuclease active site may be in close proximity in the enzyme (Nudler et al., 1995).

To identify candidate nuclease mutants we elected to use a classic genetic screen. In this type of screen one looks for a specific phenotype which presumably corresponds to the genotype of interest. When candidates are identified, they are examined to determine the gene or area of a gene that is responsible for the phenotype. In this screen, it was reasoned that a mutation

in an RNA polymerase II subunit that caused poor growth on cordycepin could indicate a decrease in nuclease activity.

Elements of the screen were designed to make it as easy and efficient as possible to identify mutants and determine the genetic basis for the phenotype. The search was narrowed down to two genes, RPB1 and RPB2, for reasons outlined above. These genes were chemically mutagenized to increase the likelihood of finding a nuclease mutant. Using RBP1 as an example, it was assumed that any of the ~5000 base pairs in the gene could be an essential element for nuclease activity. The mutation frequency was approximately 1 mutation event per 5000 base pairs, giving a mean frequency of 1 mutation event per RPB1 gene. For m=1 in the Poisson distribution [$P(n) = e^{-m}m^n/n!$)], the probability of n mutational events is as follows: P(0)=.367, P(1)=.367 and P(2)=.287. In order to be 98% certain of a mutational event at every base pair, the mean number of mutations per base pair over the course of the screen should be 3.91 [$P(0) = .02 = (e^{-m}m^0/0!)$]. Thus, a total number of ~53,000 colonies should be screened in the 5000 base pair gene if the genes with 1 mutational event are considered [(3.91*5000)/.367]. For the slightly smaller gene RPB2 (4 Kb), an estimated 43,000 colonies should be screened.

We expected to screen through a large number of candidates with normal properties in order to identify candidates of interest. As such, it was beneficial to make the screen as efficient as possible and also yield the desired sort of mutants. We were able to undertake such a measure by deleting the gene that encodes the transcription elongation factor SII. Because of SII's ability to enhance the polymerase's intrinsic nuclease activity, any mutations in the polymerase that affected interaction with this factor would likely turn up in the screen. Indeed, Wu et al. (1996) isolated mutant polymerases with a

decreased affinity for SII. While these polymerases exhibited normal cleavage activity, this activity was not stimulated by SII in the mutants as it was in the wild type. Thus, to find the desired type of mutants, those with impaired intrinsic nuclease activity, a strain of yeast without SII was needed for the screen.

Materials and Methods

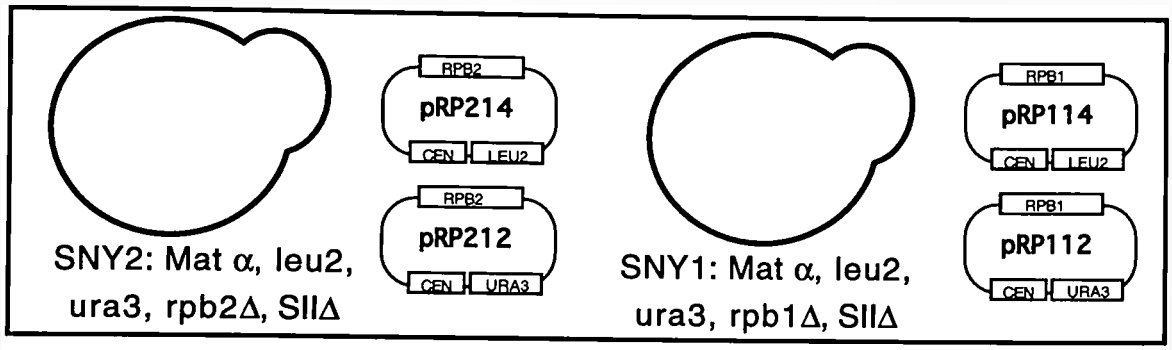
Strains

Yeast strains Z26 and Z24 as well as plasmids pRP212, pRP214, pRP112, and pRP114, were obtained from Richard Young of the Whitehead Institute (Scafe et al., 1990, and Nonet et al., 1987). Yeast strains Z26 and Z24 lack the chromosomal copy of RPB1 or RPB2, respectively. The deletion of these essential genes is complemented by either plasmid pRP112 or plasmid pRP114. Both plasmids contain the URA3 auxotrophic marker and RPB1 (pRP112) or RPB2 (pRP212). Plasmids pRP114 and pRP214 are similar but contain the LEU2 auxotrophic marker. See figure 5 for strain and plasmid details.

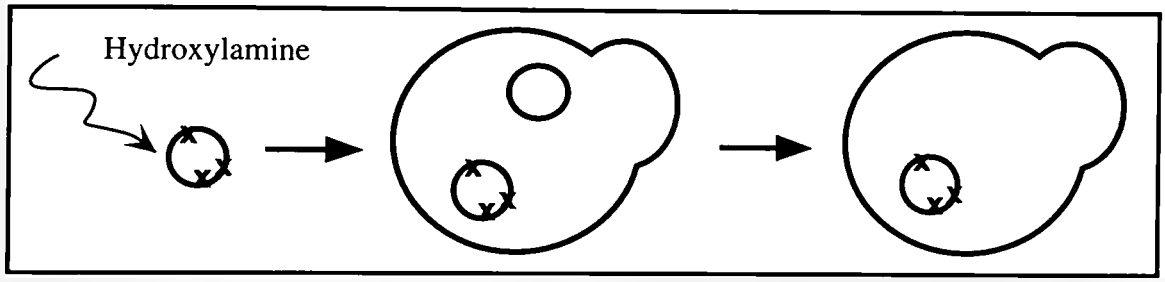
Deletion of SII

Plasmid pRS306:Δ SII KanR was previously created for the deletion of the gene PPR2 (Matt Miller, unpublished work.) It contains a gene for resistance to the drug kanamycin from the plasmid pKan^{r}MX4 flanked by a ~500 bp 5' region and a ~400 bp 3' region of PPR2 (Wach et al., 1994). Plasmid pRS306:Δ SII KanR was cut with KpnI and SacI to yield a linear fragment that was transformed into yeast strains Z26 and Z24. The 5'and 3' flanking regions of PPR2 encouraged site specific recombination, replacing the PPR2 gene with the gene for kanamycin resistance. Candidate strains were selected on the basis of their ability to grow on the drug kanamycin, as well as for lack of

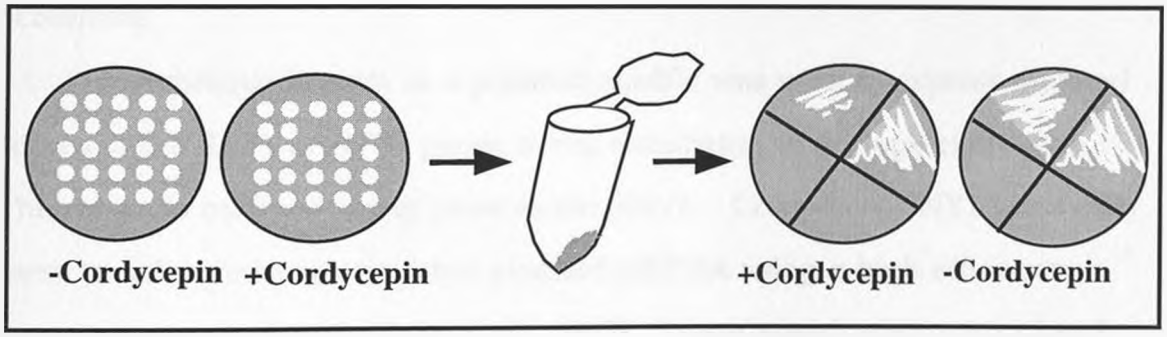
Figure 5: Yeast Strains and Screen Outline



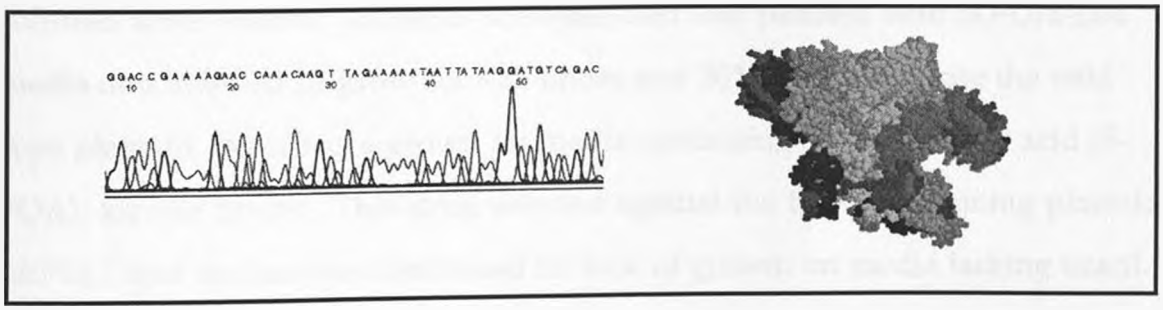
Panel 1: Yeast strain and plasmid details. Yeast strain SNY2 lacks the gene for the essential polymerase subunit RPB2, which is complemented by pRP214 or pRP212.



Panel 2: Yeast are transformed with a mutated plasmid. The wild type plasmid is selected against, and only the mutated version is expressed.



Panel 3: Yeast are tested on media with and without Cordycepin. Plasmid DNA from promising colonies is isolated, transformed back into yeast and retested.



Panel 4: Strong candidates are selected for sequencing on the basis of phenotype. Mutations are placed on the T. aquaticus RNA polymerase structure.

growth on the drug Mycophenolic acid. This drug is an inhibitor of GTP synthesis, and has been shown to negatively affect growth of yeast strains lacking SII. Genomic DNA was prepared from yeast with these characteristics and the PPR2 deletion was confirmed by PCR. After the deletion of SII, yeast strain Z26 was renamed SNY1 and Z24 was renamed SNY2.

Mutagenesis

Hydroxylamine, which causes GC to AT mutations, was selected as the mutagen. Plasmid DNA was mutagenized until a 200 fold loss of viability (as measured by loss of resistance to the drug ampicillin) was observed in E. coli. Plasmid DNA (15 μg) was mutagenized at 70° C using 90 μl of hydroxylamine solution (.36 M H_3NO HCl, .05 M KH_2 PO_4, pH 6) for either 4 hours (RPB2) or 4.5 hours (RPB1.) The plasmid DNA was electroporated into either competent TG or XL2 Blue E. coli cells. Mutated DNA was harvested from approximately 50,000 individual E. coli transformants.

Screening

A technique known as a plasmid shuffle was used to express mutated copies of the RPB1 or RPB2 genes while eliminating wild type expression. The screen is outlined using yeast strain SNY1. Competent SNY1 yeast cells were transformed with mutated plasmid pRP114 using a high efficiency transformation protocol (Geitz et al., 1995). Transformed yeast were plated onto SD-Ura-Leu media and allowed to grow at 30° C for 3-4 days until colonies were visible. Colonies were selected and patched onto SD-Ura-Leu media and allowed to grow for ~24 hours and 30° C. To eliminate the wild type plasmid, yeast were grown on media containing 5-fluoro-orotic acid (5-FOA), for ~48 hours. This drug selected against the URA3 containing plasmid pRP112, and its loss was confirmed by lack of growth on media lacking uracil. After yeast were cured of the wild type plasmid, they were replica plated onto

solid media with and without 5 μg/mL cordycepin. This concentration was slightly below that at which negative growth effects were seen in yeast with wild type plasmids. Plates were compared with the aim of finding colonies that grew on the normal media and did not grow on the cordycepin containing media. Plasmids from these yeast strains were recovered and retransformed to confirm that the observed phenotype was plasmid-born.

A small number (3-5%) of colonies were not viable when cured of the wild type plasmid. This inviability could be due to a variety of mutations. It is possible that nuclease activity is essential for growth on normal media, and loss of nuclease activity would cause a dominant lethal phenotype on cordycepin containing media. Colonies that did not grow with only a mutated plasmid were isolated from the original plates and tested on cordycepin containing media.

Further testing

Promising candidate mutants were subjected to a variety of tests, including competition assays, growth curves, and spot tests. A simple test proved to be the most reliable and reproducible. Cells were grown in liquid media for ~48 hours then diluted to an OD_{600} of .200. Aliquots (10 μl) were streaked out onto plates with and without cordycepin, and mutant growth was compared to wild type growth. Promising candidate mutants were also transformed into the SII+ parent yeast strain (Z24) and tested on media containing 25, 50, and 75 μ/mL mycophenolic acid (MPA.)

RESULTS AND DISCUSSION

Approximately 25,000 colonies were screened for defects in nuclease activity; of these, 15,000 were in RPB2, and 10,000 in RPB1. A total of 15 putative mutants were identified in the primary screen, all in the second

largest subunit. Plasmids were rescued from fourteen of the putative mutants for further testing. One proved irrecoverable due to the extremely poor growth phenotypes of the yeast. Of the fourteen recovered plasmids, seven conferred an enhanced degree of cordycepin sensitivity upon host yeast strains when retransformed. The remaining seven showed no discernible growth effect on cordycepin-containing media. Mutant growth phenotypes are shown in figure 6.

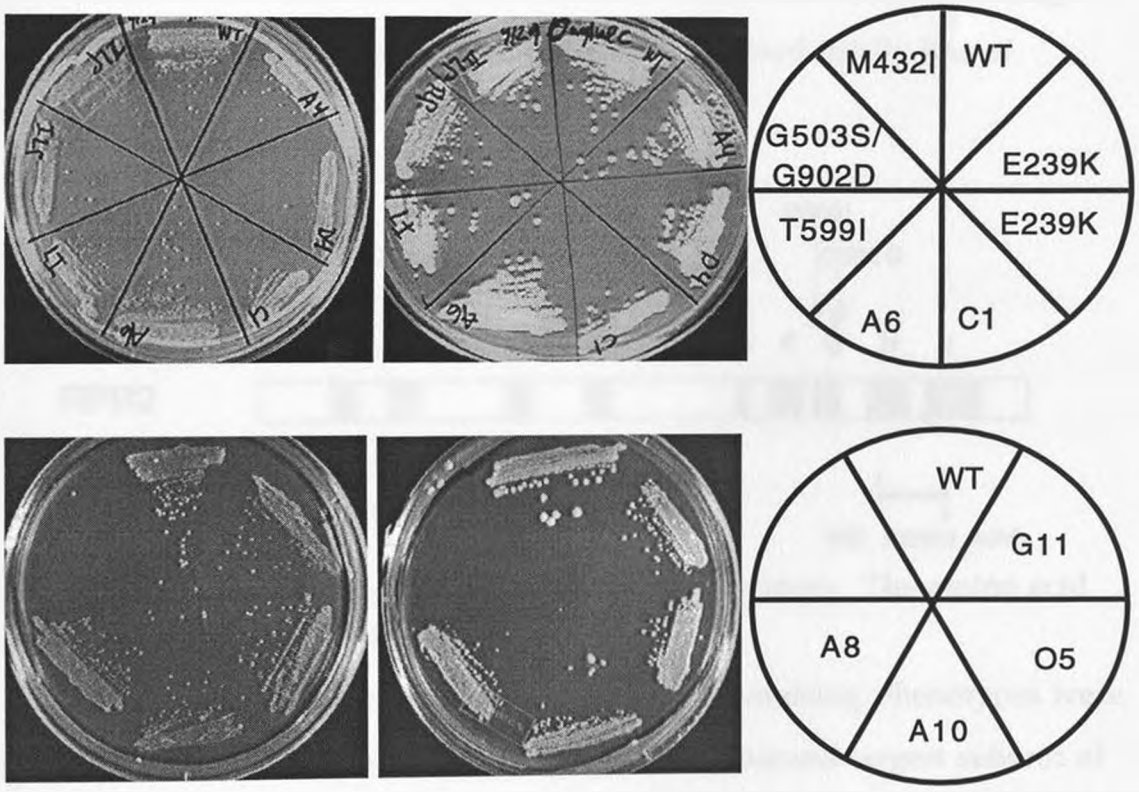


Figure 6: Growth of candidate mutants on cordyceping plates. Saturated cultures were diluted to an identical optical density and streaked onto plates with and without cordycepin.

Mutant plasmids were also transformed into the parent strain (Z24) and tested on mycophenolic acid (MPA) containing media. The deletion of SII renders yeast cells sensitive to MPA, an enzyme inhibitor that lowers the intracellular GTP pools (Exinger and Lacroute, 1992). Mutations in RNA polymerase II causing increased arrest *in vitro* also caused sensitivity to the presence of MPA or 6-Azauracil in the growth media (Lennon et al., 1998). It was reasoned that mycophenolic acid might be useful as a secondary screening agent. However, none of the yeast strains with increased sensitivity to cordycepin exhibited altered growth characteristics on MPA-containing media. This suggests that none of the mutations caused totally loss of nuclease activity.

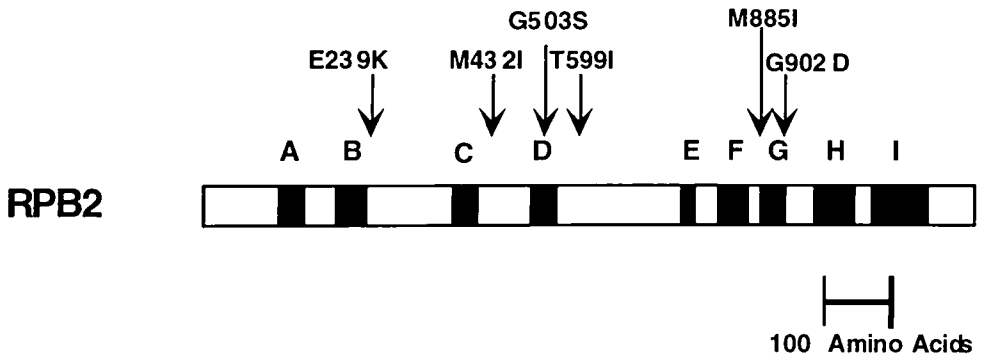


Figure 7: RPB2 has nine regions of sequence conservation. The amino acid substitutions identified in this study are show.

The five candidate mutants with the most promising phenotypes were sequenced and all had one or more mutation. The second largest subunit of RNA Polymerase II has nine regions of sequence conservation shared with other multisubunit RNA polymerases, and the mutations fell into or between several of these regions. Mutations and conserved regions are shown in figure 7. The recent publication of a high resolution structure for *T. aquaticus* RNA polymerase enabled us visualize the three dimensional

location of the mutations on an RNA polymerase (Zhang et al., 1999). While the sequence conservation between the *T. aquaticus* and *S. cerevisiae* RNA polymerases is not absolute, the presence of conserved regions allowed us to approximate the positions of our mutations. These approximations are likely most accurate for mutations that fell in highly conserved regions. Interestingly, some mutations that mapped to different regions in the primary sequence were in close proximity when shown in the protein structure.

Of particular interest were mutants E239K and the first mutation in G503S, G902D (these two mutations have not yet been separated and will be discussed individually.) The first mutation, E239K, proved to be responsible for the mutant phenotype of two separate candidates. Although E239K is located between conserved regions B and C, and G503S is located in conserved region D, these two mutations appear to be in close proximity on the *Thermus aquaticus* structure. On the primary sequence, the G503S mutation is adjacent to a region shown to influence nuclease activity in ret1, the second largest subunit of RNA polymerase III. This region will be referred to as ret1 476-512. Mutations in ret1 476-512 were first identified by altered termination properties and were also shown to modify the specificity of nuclease cutting (Shaaban et al., 1995, Bobkova et al., 1999). It has been suggested that the ret1 476-512 region could interact with the 3' end of the nascent RNA (Bobkova et al., 1999).

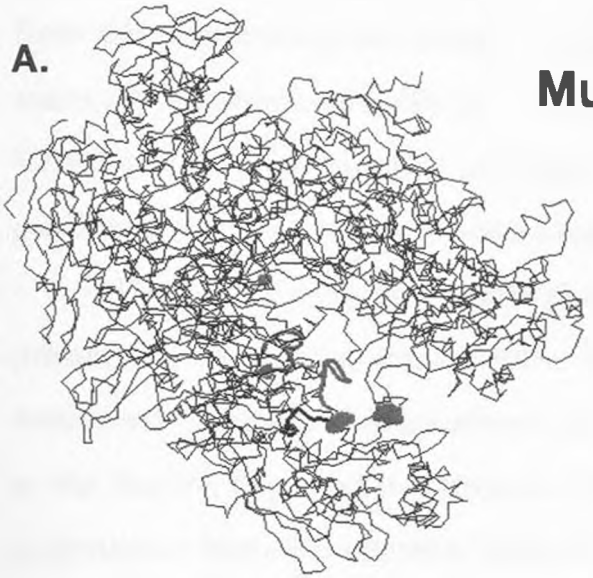
Projecting mutations onto the *T. aquaticus* structure reveals that three additional domains influencing nuclease activity are in the vicinity of the E239K and G503S mutations. The nuclease cluster is shown in figure 8a. Mutations in the first domain, referred to as ret1 367-376, caused a dramatic increase in nuclease activity (Bobkova et al., 1999). While Bobkova et al.

maintain that the extreme changes they made to residues in this region do not support a catalytic role, they suggest that the nuclease active site may be close to these residues. Interestingly, a mutation in region F of the largest subunit of RNA polymerase III that exhibited increased nuclease activity is located in the same region. This mutation, C160-270, had a decreased rate of elongation, extended pausing and a ten fold increase in transcript cleavage (Thuillier et al., 1996). A mutation in the ß subunit of *E. coli* RNA polymerase with decrease transcriptional accuracy clusters with the aforementioned mutations when projected on the *Thermus aquaticus* polymerase. However, this mutation was shown to be associated with the loss of an NTPase activity that converts non-cognate rNTPs into the corresponding rNDPs (Libby et al., 1989, Libby and Gallant, 1991).

The proximity of the four previously identified regions with altered accuracy or nuclease activity to the mutants identified in this study suggest that there is a possibility that the nuclease active site is nearby. This possibility is especially exciting in light of the fact that the mutations come from not only different regions of the second largest subunit but also from the largest subunit. However, the suggestion should be taken with the following caveats. First, we've taken the liberty of placing mutations found in four different RNA polymerases onto the structure of another. For example, mutant E239K lies outside of the highly conserved regions of the RNA polymerase. As such, it is difficult to accurately place the residue on the *T. aquaticus* structure. It should also be noted that there are two mutations in the sequence of G503S, G902D, and it is highly unlikely that both are of consequence to the nuclease activity. Nonetheless, the fact that six separate mutations map to the same area on the tertiary structure of Thermus aquaticus is significant. Additionally, the nuclease cluster is located ~30-40 Å

Figure 8: Location of Mutations on *T. aquaticus* RNA Polymerase

A.

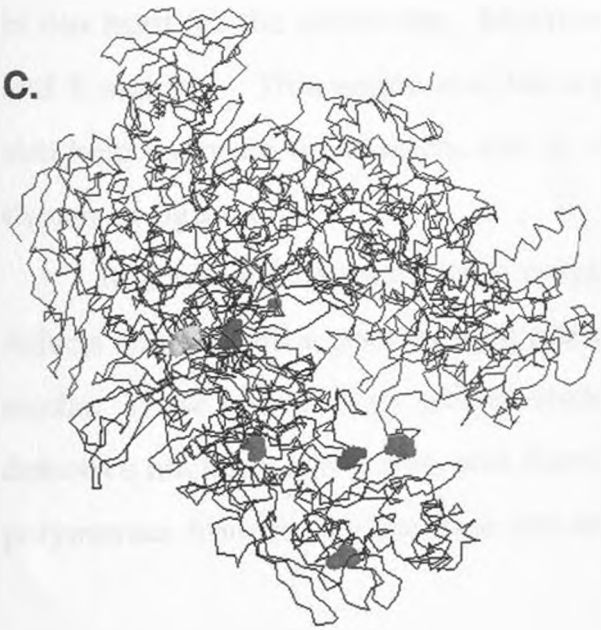


A. Accuracy and nuclease mutations from several RNA polymerases cluster with E239K and G503S on the *T. aquaticus* structure. Color scheme: Magnesium active center magenta, E239K purple, G503S green, ret1 476-512 cyan, ret1 367-376 blue, c160-270 orange, *E. Coli* Ack1 red. See text for details.

B.

B. Two mutations fall into the flexible flap region. The polymerase was crystalized in the open position, and the flexible flap is thought to close during processive elongation. Color scheme: Magnesium active center magenta, flexible flap region blue, M885I yellow, G902D red. The coordinates for the *T. aquaticus* structure are from Zhang et al. 1999.

C.



C. All of the mutations identified in this study are shown on the β subunit. Color scheme: Magnesium active center magenta, β subunit blue, E239K purple, M432I cyan, G502S bright green, T599I blue green, M885I yellow, G902D red.

from the polymerization center, a spatial relationship not unlike that seen in some DNA polymerases (Steitz, 1999). Consequently, mutant polymerases E239K and the first mutation of G503S, G1002D are high priorities for purification and *in vitro* cleavage assays.

The second mutation of G503S, G902D is in close proximity in the primary structure to mutant M885I. These two mutations fall into the highly conserved G region, and are shown in figure 8b. The G region is also known as the flexible flap, and it is proposed to have a role in conversion of the polymerase from the unstable initiating complex to the stable transcription elongation complex (Mooney and Landick, 1999). It is suggested that this flap might close around the upstream RNA to form an RNA exit tunnel in the transcription elongation complex (Zhang et al., 1999). This region could have some involvement in the transition from polymerization to cleavage.

Mutation M432I lies slightly outside of the conserved C region, and is approximately 20 amino acids away from the ret1 367-376 region mentioned earlier. It is unlikely that this position itself is involved in catalysis, as it affects the nonpolar methionine. However, could be involved in positioning the RNA or other functions important to cleavage, so it is an important clue in our hunt for the active site. Mutation T699I lies between the conserved D and E regions. This amino acid has a polar, uncharged side chain, and could also help us locate the catalytic site of cleavage. All of the mutations are shown in figure 8c.

It is reasonable that a RNA polymerases with a variety of elongation defects could exibit a poor growth phenotype on cordycepin-containing media. These include, but are not limited to, polymerases with a catalytically defective nuclease active site, and those with a reduced ability to move the polymerase from the polymerase site to the nuclease site or align it for

cleavage. Additionally, it is possible that we've identified mutants with an increased affinity for cordycepin. However, each mutant could help us to understand a different aspect of elongation and nuclease activity. As such, we look forward to purifying the mutant polymerases and characterizing their elongation and cleavage properties. Additionally the mutants give clues as to areas of the polymerase involved in nuclease activity. They define more specific targets for further screens.

CONCLUSIONS

The five variants of the RPB2 gene identified in this screen on the basis of increased sensitivity to cordycepin represent a new class of mutants. In the coming months, we plan to purify and characterize these mutants. It is our hope that they will prove to have defects in nuclease activity or other elongation properties. Of the variants identified, four had a single point mutation, and one had two point mutations.

Because the coordinates for the *T. aquaticus* polymerase, an RNA Polymerase II-like structure, are available, we were able to place our mutations in a structural context. This modeling revealed that two of our mutants, E239K and G503S, cluster on the tertiary structure with four nuclease or accuracy mutants from other studies. In conjunction with the previously identified mutants, our mutants strongly implicate this area in some aspect of nuclease activity. This area should be a target in further studies by random or site directed mutagenesis.

None of the mutants displayed any altered growth characteristics when tested on mycophenolic media. Because MPA is harmful to cells lacking SII, it was our expectation that mutants devoid of nuclease activity would grow poorly in the presence of MPA. Normal growth on MPA, along with the observation that the mutants grow, although poorly, on cordycepin, suggests that none of the mutants identified were completely lacking nuclease activity. It is possible that nuclease activity is essential in yeast, causing such mutants to have a null phenotype. Bobkova et al. suggest that this may be the case for RNA polymerase III (1999). We took this possibility into account in our screen, and would have identified any null mutants that confered a dominate negative phenotype when grown on cordycepin.

The screen was far from comprehensive. We calculated that 43,000 colonies should be screened in RPB2 to ensure with 98% certainty a mutational event at every base. However, only 15,000 colonies were screened. Nonetheless, the screen as described in this work has in some respects reached its full potential. Because of the nature of the phenotype, slow growth on cordycepin-containing media, the screen was slow and labor intensive. The identification of regions causing cordycepin sensitivity will enable us to design smaller, more efficient screens that target a specific area. No variants were identifed in the RPB1 gene, suggesting that future work is best targeted to the RPB2 gene.

Isolation of nuclease mutants opens a new door for understanding transcriptional pausing, elongation, proofreading, and termination. In this screen, we have identified five variants that may have impaired nuclease activity, as well as regions of the polymerase that warrant further study. It is a starting point for further studies on the basic mechanisms of transcription.

GLOSSARY

Catalytic site: Specific amino acid residues of an enzyme involved in substrate catalysis.

Complementary: Refers to two nucleic acid strands or molecules that fit together. In DNA, guanine is complementary to cytosine, and adenosine is complementary to thymine.

Expression: Process by which the information in a gene produces an observable phenotype.

In vitro: Biochemical studies that take place outside the organism, generally using isolated proteins or cell free extracts.

In vivo: In vivo studies are carried out in an intact cell or organism.

Plasmid: A small, circular molecule of extrachromosomal DNA. Plasmids are used to introduce genetic information into yeast and E. coli in this study.

Transcription: Process where one strand of DNA is used as a template for the synthesys of a complementary RNA.

Transformation: Introduction of foreign DNA into a cell. Yeast and E. coli are routinely transformed with plasmids in this study.

NUCLEASE SCREEN SUMMARY

Mutant	Isolated in Original Screen	Amino Acid side chain	Located in Conserved Region:	Cordycepin Phenotype
E239K	A4	Glu to Lys acidic to basic	Between B and C	Sensitive
M432I	J7II	Met to Ile both nonpolar	Immediately after region C	Sensitive
G503S/ G902D	J7I	gly to ser nonpolar to polar gly to asp nonpolar to acidic	D G	mutations not separated, yeast are sensitive
T599I	I7	tyr to ile polar to nonpolar	between D and E	Sensitive
M885I	B7	Met to Ile both nonpolar	G	Sensitive
E239K	D4	Glu to Lys acidic to basic	Between B and C	Sensitive
unknown	A10	unknown	unknown	Sensitive
	C1			Not Sensitive
	A6			Not Sensitive
	O5			Not Sensitive
	A8			Not Sensitive
	G11			Not Sensitive
	K11			Not Sensitive
	R5	single plasmid not viable, no growth effect with two plasmids mutants of this sort are not considered candidates		
	B5	sensitive in primary screen, could not be recovered for further testing		

References

Archambault, J. and Friesen., J. Genetics of Eukaryotic RNA Polymerases I, II, and III. Microbiological Reviews. 703-724. Sept. 1993.

Beese, L., and Steitz, T. Structural basis for the 3'-5' exonuclease activity of Escherichia coli DNA polymerase I: a two metal ion mechanism. EMBO Journal. 10 (1), 25-33. 1991

Blank, A., Gallant, J., Burgess, R., and Loeb, L. An RNA polymerase mutant with reduced accuracy of chain elongation. Biochemistry 25 (20), 5920-8. 1986.

Bobkova, E., Habib, N., Alexander, G., and Hall, B. Mutational Analysis of the Hydrolytic Activity of yeast RNA Polymerase III. Journal of Biological Chemistry. 274, 21342-21348. 1999.

Donlin, M., Patel, S., and Johnson., K., Kinetic partitioning between the exonuclease and polymerase sites in DNA correction. Biochemistry. 30, 538-546. 1991.

Echols, H., and Goodman, M., Fidelity Mechanisms and DNA replication. Annu. Rev. Biochem. 477-511. 1991.

Eichhorn , G., Chuknyisky, Pl, Butzow, J., Beal, R., Garland, C., Janzen, C., Clark, P., and Tarien, E. A structural model for fidelity in transcription. Proceedings of the National Academy of Sciences. 91, 7613-7617. 1994.

Exinger, F. and Lacroute, F. 6-Azauracil inhibition of GTP biosynthesis in Saccharomyces cerevisiae. Current Genetics. 22 (1), 9-11. 1992.

Gietz, R., Schiestl, R., Willems, A. and Woods, R. Studies on the transformation of intact yeast cells by the LiAc/SS-DNA/PEG procedure. Yeast 11 (4), 355-60. 1995.

Hawley, D., Wiest, D., Holtz, M., and Wang, D. Transcriptional pausing, arrest, and readthrough at the adenovirus major late attenuation site. Cellular and Molecular Biology Research. 39 (4): 339-348. 1993

Iwashima, A., Kawasaki, Y., Nosaka, K., and Nishimura, H. Effect of thiamin on cordycepin sensitivity in Saccharomyces cerevisiae. FEBS. 1, 60-62. 1992.

Izban., M. and Luse, D. The RNA polymerase II ternary complex cleaves the nascent transcript in a 3' to 5' direction in the presence of elongation factor SII. Genes and Development. 6, 1342-1356. 1992.

Jeon, C., and Agarwal, K. Fidelity of RNA polymerase transcription controlled by elongation factor TFIIS. Proceedings of the National Academy of Science. 93, 13677-13682. 1996.

Kerppola, T., and Kane, C. RNA polymerase: regulation of transcript elongation and termination. FASEB JOURNAL 5 (13), 2833-42. 1991.

Kuchta R., Benkovic P., Benkovic, S. Kinetic mechanism whereby DNA polymerase I (Klenow) replicates DNA with high fidelity. Biochemistry 27 (18): 6716-25 1988.

Lennon, J., Wind, M., Saunders, L., Hock, M., and Reines, D. Mutations in RNA polymerase II and elongation factor SII severely reduce mRNA levels in *Saccharomyces cerevisiae*. Molecular and Cellular Biology. 18 (10) 5771-5779. 1998.

Libby, R., and Gallant, J. The role of RNA polymerase in transcriptional fidelity. Molecular Microbiolgy. 5, 999-1004. 1991.

Libby, R., Nelson, J., Clavo, J., and Gallant, J. Transcriptional proofreading in Escherichia coli. EMBO Journal. 8, 3153-3158. 1989.

Mooney, R. and Landick, R. RNA polymerase unveiled. Cell 98 (6), 687-90. 1999.

Nonet, M., Sweester, D., and Young, R., Functional Redundancy and sturctural polymorhpism in the large subunit of RNA polymerase II. Cell., 50, 909-915. 1987.

Nudler, E., Kashlev, M., Nikiforov, V., Goldfarb, A., Coupling between transcription termination and RNA polymerase inchworming. Cell. 81, 351-7 1995.

Patel, S., Wong, I., and Johnson., K., Pre-steady state kinetic analysis of processive DNA replication including complete caharacterization of an exonuclease-deficent mutant. Biochemistry, 30, 511-525. 1991.

Pinaud, S., and Mirkovitch, J., Regulation of c-fos expression by RNA polymerase elongation competence. Journal of Molecular Biology. 280(5), 785-798. 1998.

Powell, W., and Reins, D., Mutations in the second largest subunit of RNA polymerase II cause 6-Azauracil sensitivity in yeast and increased transcriptional arrest *in vitro*. Journal of Biol. Chem. 271, 6866-6873. 1996.

Powell, W., Lennon, J., Elsevier, J., Reines., D. Glutamic acid 371 of teh barnase homology domain in RNA polymersae II is not required for SII-activated RNA cleavage. Mol. Gen. Genetics. 253, 507-511. 1997.

Reha-Krantz, L., Regulation of DNA polymerase exonucleolytic proofreading activity: Studies of bacteriophage T4 "antimutator" DNA polymerases. Genetics. 148, 1551-1557. 1998.

Reines, D. Elongation factor-dependent transcript shortening by template-engaged RNA polymerase II J. Biol. Chem. 267, 3795-3800. 1992.

Rose, K., Bell, L., and Jacob, S. Specific inhibition of chromatin-associated poly(A) synthesis *in vitro* by cordycepin 5' triphosphate. Nature. 267, 178-180 1977

Scafe C., Martin C., Nonet M., Podos S., Okamura S., Young R.A., Conditional mutations occur predominantly in highly conserved residues of RNA polymerase II subunits. Mol Cell Biol. 10 1270-5. 1990.

Shaaban S., Krupp B., Hall B., Termination-altering mutations in the second-largest subunit of yeast RNA polymerase III. Molecular and Cellular Biology. 15 (3). 1467-78. 1995.

Shirai, T., and Go, M., RNase-like domain in DNA-directed RNA polymerase II. Proc. Natl. Acad. Sci. 88, 9056-9060.

Shilatifard, A. The RNA polymerase II general elongation complex. Biological Chemistry. 379. 27-31. 1998.

Steitz, T., DNA polymerases: structural diversity and common mechanisms. Journal of Biological Chemistry. 274(25) 17395-17398. 1999.

Sweester, D., Nonet, M., and Young, R., Prokaryotic and eukaryotic RNA polymerases have homologous core subunits. Proc. Natl. Acad. Sci. 84, 1192-1196, 1987.

Thomas, M., Platas, A., and Hawley, D. Transcriptional fidelity adn proofreading by RNA polymerase II. Cell, 93, 627-637, 1998.

Thuillier, V., Brun. I., Sentenac, A., and Werner, M. Mutations in the α-amanitin conserved domain of the largest subunit of yeast RNA polymerase III affect pausing, RNA cleavage and transcriptional transitions. EMBO Journal. 15 (3), 618-619. 1996.

Uptain, S., Kane, C., and Chamberlin, J. Basic Mechainsms of Transcript elongation and its regulation. Annu. Rev. Biochem. 66, 117-172. 1997.

Wang, D., and Hawley, D. Identification of a 3' to 5' exonuclease activity associated with human RNA polymerase II. Proc. Natl. Acad. Sci. 90, 843-847. 1993.

Wang, J., Yu, P., Lin, T., Konigsberg, W., and Steitz, T., Crystal structures of an NH_2-terminal fragment of T4 DNA polymerase and its complexes with single-stranded DNA and with divalent metal ions. Biochemistry. 35, 8110-8119. 1996.

Wiest, D., Wang., D. and Hawley, D. Mechanistic Studies of transcription arrest at the Adenovirus major late attenuation site. Journal of Biol. Chem. 1992.

Wong, I., Patel, S., and Johnson, K., An induced fit mechanism for DNA replication fidelity: Direct measurement by single-turnover kinetics. Biochemistry. 30, 526-537. 1991.

Wu, J., Awrey, D., Edwards, A., Archambault, J., Friesen, J., In vitro characterization of mutant yeast RNA polymerase II with reduced binding for elongation factor TFIIS. Proc. Natl. Acad. Sci. 93, 11552-7. 1996.

Yang, X. and Roberts, J. Gene Q antiterminator proteins of Escherichia coli phages 82 and lambda suppress pausing by RNA polymerase at a rho-dependent terminator and at other sites. Proceedings of the National Academy of Science. 86 (14), 5301-5305. 1989.

Yanofsky, C., Transcription attenuation. Journal of Biological Chemistry. 263, 609-612. 1988.

Zhang, G., Campbell, E., Minakhin, L., Richter, C., Severinov, K., and Darst, S. Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 Å resolution. Cell. 89, 811-824. 1999.