Functional Analysis of Read-through Transcription in *Saccharomyces cerevisiae* Genes Transcribed by RNA Polymerases I and II

Submitted for the degree of
Doctorate of Philosophy

Trinity Term, 2002

Elizabeth Marie Prescott

Sir William Dunn School of Pathology
Balliol College
University of Oxford

RADCLIFFE SCIENCE LIBRARY OXFORD
11 NOV 2002

Abstract

The molecular consequences of read-through transcription in *Saccharomyces cerevisiae* genes transcribed by RNA Polymerases I and II (Pol I and Pol II) are investigated in the two results sections of this thesis (Part I and Part II). In these studies, nascent and steady state RNA analyses are employed to evaluate transcriptional termination.

Part I focuses on the impact that co-expressed Pol II transcribed genes can have on each other when placed in a convergent orientation. Transcriptional interference (TI), previously shown to occur between tandemly arranged genes, inhibits initiation of transcription on a downstream gene. These studies demonstrate that TI can also occur with *S.cerevisiae* Pol II genes arranged convergently. When the *GAL10* and *GAL7* genes are rearranged in a convergent orientation, transcriptional initiation occurs at full levels. However as soon as the two transcripts begin to overlap, elongation is restricted resulting in a severe reduction in steady state mRNA accumulation. This effect is only observed in the *cis* arrangement arguing against RNAi effects acting on the potential generation of antisense transcripts. These data reinforce the necessity to separate adjacent Pol II transcription units by efficient termination signals.

Part II describes experiments that define the nascent termination site of Pol I transcribed rRNA genes. Furthermore, transcription run-on (TRO) analysis was performed on yeast strains in which specific factors involved in the Pol I transcription and RNA processing had been deleted. In particular, *S.cerevisiae* strains lacking the RPA12p subunit of Pol I have a severe defect in transcriptional termination at the nascent level indicating a direct role for this subunit in termination. Additionally, a yeast strain lacking Rnt1p, a protein involved in early rRNA processing, shows an altered transcriptional termination profile suggesting a link between processing and termination efficiency. In contrast to TI in Pol II, read-through transcription in Pol I may not significantly reduce rRNA gene synthesis.

Acknowledgments

First, I would like to thank Professor Nick J. Proudfoot for all of his support, scientific and otherwise, over the years. I would also like to express my appreciation (beyond the copious amounts of chocolate!) for the NJP lab past and present. Special thanks go to Dr. Kate Plant, Dr. Justin O'Sullivan, Dr. André Furger and Dr. Claudia Alén for ploughing through drafts of this thesis. Thanks also go to Dr. Ingo Gregor, Dr. Agustin Aranda and Barbara Lee for early advice and ideas. In addition, I have been greatly assisted by others in the department and beyond and for that I will be forever grateful!

Life outside of lab was a key ingredient in maintaining my sanity over these years. Balliol College supplied numerous avenues to mature in my field and beyond. Piers Nye was a wonderful College advisor and Andrew Graham has given me great inspiration to pursue interdisciplinary interests (as well as getting me out of bed in the morning!). Moreover, I could not have made it this far without the vibrant and supportive graduate community supplied by Balliol in the form of 'The Manor'. I have come in contact with so many amazing people in the past four years. I am not able to acknowledge you all but I must highlight a few people who kept my sprits up when the science was dragging me down: Alana McMahon, Matthew Glossop, Geoff Cross, Beth Shapiro, Antwaun Smith, Erin Bohula, Erin Whalen, Jayne Hooper, Abigail Cox, Lina Nerlander and the HM dancing troupe. Thank you for everything over the years!

To my brother, who continues to believe in me when I need it most, your unquestioning support will never be forgotten. To my Mom, your love and support made it possible for me to venture so far from the nest. The 'beast' belongs to you both.

To Dr. Howie

and unfinished adventures

Contents

Title.....i
Abstract.....ii
Acknowledgments.....iii
Dedication.....iv
Contents.....v
Abbreviations.....ix

Introduction.....1
Preface.....1
1.1 *S.cerevisiae* as a model organism.....1
Genome.....1
ORF orientation and proximity.....2
1.2 Transcription.....2
Different eukaryotic polymerases.....3
[Table A].....4
Chromatin.....4
1.3 Transcription Phases.....5
(a) Initiation.....6
Cis-acting sequence elements.....6
Pol I.....6
Pol II.....8
General Transcription Factors.....9
Pol I.....9
[Figure A].....10
Pol I formation of preinitiation complex.....12
Pol I holoenzyme.....12
Pol II.....12
Pol II formation of preinitiation complex.....13
[Figure B].....14
Pol II holoenzyme.....15
Pol II transcription factors.....16
Pol II transcription co-activators.....16
Pol II transcription repressors.....17
(b) Open complex formation.....17
Pol I.....17
Pol II.....17
(c) Elongation.....18
Mechanism of elongation.....19
Promoter escape and early elongation.....19
Pol I.....20
Pol II.....20
Phosphorylation of CTD.....20
Elongation Factors.....20
Pol I.....21
Pol II.....21
DNA dependent pause, arrest and transcription blocks.....22

(d) Processing....23
Pol I....23
Pol II....24
[Figure C]....26
(e) Termination and Release....27
Pol I....27
[Figure D]....29
Pol II....30
1.4 rDNA repeat: Model Pol I system....32
Sequence specifics....32
[Figure E]....33
Regulation....34
1.5 Galactose metabolising genes: Model Pol II system....35
Genes involved in galactose metabolism....35
GAL gene repression....36
Conversion of galactose to glucose-6-phosphate....36
1.6 Aims of this study....36

Materials....38
Reagents....38
Enzymes....39
Radioisotopes....39
Kits....40
Organic reagents....40
Molecular weight markers....40
Media....40
Solutions and buffers....42
Strains....46
Plasmid constructs....47
M13 probe constructs....48
[Table B]....49
Primers....50
Northern probes....50

Methods....51
Phenol extraction/ethanol precipitation....51
Restriction enzyme digests....51
Phosphotasing....51
Flushing ends....51
Polymerase chain reactions....52
Colony screening....53
Plaque screening....53
Kinasing primers....53
Ligation....53
Plasmid preparation....53
Electrophoresis....54
Sequencing....55
Transformation and growth....55
Blotting....58
In vitro transcription....60
Transcription Run-on Analysis (TRO)....61

Part I: Convergent Transcription.....64
Preface.....64
Introduction.....64

Section 1: Convergent nascent transcription in *GAL10* and *GAL7*.....67
1.1 TRO analysis of nascent transcription.....67
[Figure I 1.1].....68
1.2 Transcription of WT *GAL* genes with redesigned probes.....68
[Figure I 1.2].....69
[Figure I 1.3].....70
[Figure I 1.4].....71
1.3 Transcription of *GAL* genes in a convergent orientation.....71
[Figure I 1.5].....72
[Figure I 1.6].....72
[Figure I 1.7].....73
[Figure I 1.8].....74
[Figure I 1.9].....75
[Figure I 1.10].....75

Discussion.....76

Section 2: Transcriptional Collision.....78
2.1 Transcriptional collision in non-terminating convergent transcription... 78
[Figure I 2.1].....79
[Figure I 2.2].....79
2.1.2 Nascent transcription from Fus construct.....79
[Figure I 2.3].....80
[Figure I 2.4].....81
[Figure I 2.5].....82
2.2 Separate transcription from individual promoters.....82
2.2.2 Transcription of *GAL10* with *GAL7* promoter deletion (Fus-Δ7P)....82
[Figure I 2.6].....83
[Figure I 2.7].....84
2.2.3 Transcription of *GAL7* with *GAL10* promoter deletion (Fus-Δ10P)..84
[Figure I 2.8].....85
[Figure I 2.9].....85
[Figure I 2.10] 86
2.2.4 Transcription with both *GAL* promoters deleted (Fus-Δ7P/Δ10P).....86
[Figure I 2.11].....87
2.2.5 Separate *GAL10* and *GAL7* transcription in Fus (Fus-Δ7P & Δ10P)..87
[Figure I 2.12].....88
Discussion.....88

Publication – PNAS 99 (13): 8796-8801 JUN 25 2002

Part II: rRNA Gene Termination.....91
Preface.....91
Introduction.....91
[Figure II 0.1].....95
[Figure II 0.2].....96

Section 1: Transcriptional termination in rRNA genes.....97
1.1 Probes for rRNA genes.....97
[Figure II 1.1].....98
[Figure II 1.2].....99
1.2 Synthetic RNA demonstrates probe hybridisation.....99
[Figure II 1.3].....100
1.3 Identification of sites of Pol I nascent rRNA gene termination.....100
1.3.1 *S.cerevisiae* strain *ΔGAL* transcription termination profile.....101
[Figure II 1.4].....101
[Figure II 1.5].....102
1.3.2 *S.cerevisiae* strains MWK028 and *RNT1* TRO profiles.....103
[Figure II 1.6].....103
[Figure II 1.7].....104
[Figure II 1.8].....105
Discussion.....106

Section 2: Deletion of Pol I subunit Rpa12p eliminates termination.....108
2.1 Effect of *RPA12* disruption on Pol I termination.....108
[Figure II 2.1].....109
[Figure II 2.2].....109
2.2 Over expression of Rpa190p does not restore termination.....110
[Figure II 2.3].....110
[Figure II 2.4].....111
2.3 Composite of MWK028, RLY10 and RLY11.....111
[Figure II 2.5].....112
Discussion.....113

Section 3: Deletion of *RNT1* on rRNA gene termination.....116
3.1 Rnt1p deletion strain termination profile.....116
[Figure II 3.1].....117
[Figure II 3.2].....117
[Figure II 3.3].....118
Discussion.....119

Conclusion.....121
Future Prospects.....123

References.....124

Abbreviations

In addition to standard SI units and chemical symbols the following abbreviations were used:

A	adenine
AP	alkaline phosphatase
ARS	autonomously replicating sequence
ATP	adenosine 5'-triphosphate
bp	base pairs
BSA	bovine serum albumin
C	cytidine
CBC	cap binding complex
cDNA	complimentary DNA
CEN	centromere element
Ci	Curie
CF	cleavage factor or core factor
CoTC	co-transcriptional cleavage
CPSF	cleavage and polyadenylation specificity factor
CstF	cleavage stimulation factor
CTD	C-terminal domain of RNA Pol II
CTP	cytidine 5'-triphosphate
DEPC	diethyl pyrocarbonate
d	2'-deoxy
dH_2O	distilled water
DNA	deoxyribonucleic acid
DNase	deoxyribonuclease
ds	double stranded
DSE	downstream element
DTT	dithiolthreitol
dTTP	deoxyrthymidine 5'-triphosphate
EDTA	ethylenediaminetetra-acetic acid
EE	efficiency element
EtOH	ethanol
ETS	external transcribed spacer
G	guanine
GT	guanylyltransferase
GTF	general transcription factor
GTP	guanosine 5'-triphosphate
HAT	histone acetyltrasnferase

HDAC	histone deacetyltransferase
IGS	intergenic spacer
Inr	initiator
IPTG	isopropyl-β-D-thiogalactoside
ITS	internal transcribed spacer
kb	kilobase
kDa	kilodaltons
LTR	long terminal repeat
MOPS	3-[N-morpholino]propanesulphonic acid
mRNA	messenger RNA
MT	methyltransferase
N	any nucleotide
NaOAc	sodium acetate
nt	nucleotides
NTP	nucleoside 5'-triphosphate
NTS	non-transcribed spacer
OD	optical density
o/n	overnight
ORF	open reading frame
PAB	poly(A) binding protein
PAGE	polyacrylamide gel electrophoresis
PAP	poly(A) polymerase
PCR	polymerase chain reaction
PE	positioning element
PEG	polyethylene glycol
PIC	pre-initiation complex
PNK	polynucleotide kinase
Pol	polymerase
poly(A)	polyadenylate
PTRF	Pol I transcript release factor
R	purine
RBD	RNA binding domain
rDNA	ribosomal DNA
RF	replicative form
RFB	replication fork barrier
rInr	ribosomal initiator
RNA	ribonucleic acid
RNAP	RNA polymerase
RNase	ribonuclease

RNP	ribonucleoprotein
rRNA	ribosomal RNA
rt	room temperature
RTP	RNA 5' triphosphatase
SC	synthetic complete
SDS	sodium dodecyl sulphate
SF	splicing factor
snoRNA	small nucleolar RNA
snRNA	small nuclear RNA
snRNP	small nuclear ribonucleoprotein
SRB	Suppressor of RNA Pol B
ss	single stranded
T	thymine
T4 Pol	(bacteriophage) T4 DNA Polymerase
TAF	TBP associated factor
TBP	TATA binding protein
TC	transcriptional collision
TEMED	N,N,N',N'-tetramethylethylenediamine
TF	transcription factor
TI	transcriptional interference
tRNA	transfer RNA
TRO	transcription run-on
ts	temperature sensitive
TTF-1	Pol I transcription termination factor (vertebrates)
U	uracil
UAF	upstream activating factor
UAS	upstream activating sequence
UPE	upstream promoter element
URS	upstream repressor sequence
USE	upstream sequence element
UTP	uridine 5'-triphosphate
UTR	untranslated region
UV	ultraviolet
X-Gal	5-bromo-4-chloro-3-indoyl β D galactoside
Y	pyrimidine
YP	yeast peptone

Introduction

Preface

The mechanism and regulation of transcriptional initiation, elongation and termination are studied in an effort to gain understanding of how organisms process internal and external stimuli. This thesis has focused on nascent transcription by RNA polymerase I (Pol I) and RNA polymerase II (Pol II) of the budding yeast *Saccharomyces cerevisiae* in situations that result in both proper and improper transcriptional termination.

1.1 *S.cerevisiae* as a model organism

Inherent characteristics of the yeast *Saccharomyces cerevisiae* make it a valuable model system for the study of eukaryotic molecular genetics. This unicellular organism has an easily controllable chemical and physical environment, a life cycle that enables traditional genetic analysis and a detailed genetic map (Goffeau *et al.* 1996). The 16 chromosomes and relatively small haploid genome have facilitated the development of techniques that enable precise manipulation of the DNA sequence. Accurate and efficient targeting of specific areas of the genome have permitted the replacement or deletion of desired genes or regions enabling greater experimental control. Numerous *S.cerevisiae* homologs to proteins involved in human disease have been identified adding to the clinical usefulness of this model system (Dujon 1996). However, a significant number of 'orphans' (putative genes that have no known function or homolog) have been identified in the *S.cerevisiae* genome so that a great deal remains to be investigated.

Genome

In 1996, as the result of a collaborative sequencing effort, the 12068 kb genome of the *S.cerevisiae* was released to public databases (Goffeau *et al.* 1996). For a modest sized genome, only four times the size of the genome of *Escherichia coli* (Blattner *et al.* 1997) and over 200 times smaller than the size of the human genome (Venter *et al.* 2001), *S.cerevisiae* has a relatively high frequency of potential protein encoding genes or open reading frames (ORFs).

Sequencing of the *S.cerevisiae* genome revealed almost 6000 potential ORFs, approximately 140 ribosomal RNA (rRNA) genes, 40 genes for small nuclear RNA (snRNA) molecules, and 275 transfer RNA (tRNA) genes (Goffeau *et al.* 1996). Almost 70% of the genome is represented by protein encoding ORFs, approximately 9% consists of RNA coding elements such as ribosomal DNA (rDNA) with the majority of the remainder (or 20%) considered 'intergenic' regions including promoters, terminators and regulatory or undefined elements (Dujon 1996). On average, there is one protein-encoding gene per 2 kb which, in comparison to the genome of the nematode worm (one per 6 kb) and the human genome (one per 30 kb), is a markedly compact arrangement of genes.

ORF orientation and proximity

The high gene density in *S.cerevisiae* is largely a result of the short intergenic regions. On average, divergent promoters are separated by 618 bp while tandem (head to tail) intergenic regions, which contain both a promoter and terminator region, average 517 bp (Dujon 1996). Intergenic regions for convergent genes are separated by an average of only 326 bp (Dujon 1996) with several examples of less than 300 bp (Barker *et al.* 1987; McGrath and Varshavsky 1989; Malavasic and Elder 1990; Peterson and Myers 1993). Such a limited intergenic sequence would allow only ~150 bp per gene for proper termination outside of the neighbouring ORF. This is less sequence than that which has so far been defined for the few characterised termination regions in *S.cerevisiae* genes (Osborne and Guarente 1989; Aranda *et al.* 1998a; Aranda *et al.* 1998b). Analysis of adjacent gene sets in *S.cerevisiae* determined that co-expressed tandem and divergent genes represent approximately equal percentages (47% each) of the total number of co-transcribed genes sets detected (Kruglyak and Tang 2000). Interestingly, less than 6% of co-expressed, adjacent genes are in a convergent orientation suggesting that there may be evolutionary pressure to select against convergently arranged, co-transcribed genes (Dujon 1996; Kruglyak and Tang 2000).

1.2 Transcription

Transcription, the process of synthesizing a RNA chain representing a portion of one strand of the DNA, is a complex process that involves different stages and factors. A transcription unit consists of the DNA between the point of initiation and termination. Synthesis, as with all nucleic acids, is in

the 5' to 3' direction. This process begins with the assembly of the polymerase and multiple initiation factors into a closed preinitiation complex (PIC) at the promoter followed by the formation of an open complex that then leads to the synthesis of the first phosphodiester bond of the nascent transcript. After transcription initiation, the polymerase complex moves away from the promoter, initiation factors disassociate, resulting in what is termed 'promoter escape'. Elongation, followed by processing and termination, completes the transcript.

In prokaryotes, all transcription occurs from a single polymerase that is responsible for RNA synthesised from all genes. *E.coli* RNA Polymerase (RNAP) consists of three core subunits ($\alpha 2/\beta/\beta'$). Binding of factors to upstream DNA and an additional subunit (σ) are required for promoter specificity (Young *et al.* 2002). However, eukaryotes utilise different polymerases for specific gene types requiring the involvement of numerous proteins at various different stages of transcription.

Different eukaryotic polymerases

Eukaryotic genes are transcribed by one of three nuclear RNA polymerases. Pol I transcribes rRNA from the multiple tandem repeats of rDNA. Pol II primarily transcribes protein-encoding genes, which are processed into messenger RNA (mRNA). Pol II is also responsible for the transcription of some snRNAs and short nucleolar RNA (snoRNA). tRNA, 5S rRNA and the remainder of the snRNAs are transcribed by RNA polymerase III (Pol III). The portion of total RNA generated varies by polymerases; on average Pol I generates approximately 60% of transcripts, Pol II 20-40% and Pol III about 10% of the total transcripts in the cell (Warner 1999).

Prokaryotic RNAP shares homology with various subunits of the eukaryotic RNA polymerases. Despite notable similarities, the structure and composition of each polymerase in *S.cerevisiae* is unique. Pol I consists of 14 subunits, Pol II contains 12 subunits and Pol III contains 13 subunits (Ishihama *et al.* 1998)(see Table A).

Pol I	Pol II	Pol III
A190	B220	C160
A135	B150	C128
A49	B44.5	C82
A43	B32	C53
AC40	ABC27	AC40
A34.5	ABC23	C37
ABC27	B16	C34
ABC23	ABC14.5	C31
AC19	B12.6	ABC27
ABC14.5	B12.5	C25
A14	ABC10α	ABC23
A12.2	ABC10β	AC19
ABC10α		ABC14.5
ABC10β		C11
		ABC10α
		ABC10β

le **A** Subunits of Pol I, Pol II and Pol III from *S.cerevisiae*. Subunits found in Pol I contain an 'A', Pol II a 'B' and Pol 'C'. If a subunit is found in more than one of the polymerases it will contain the appropriate letters. Subunits showing ology are colour coded. Subunit AC40 (identical in both Pol I and Pol II) are related to subunit B44.5 in Pol II.

structures for Pol I (Schultz *et al.* 1993; Klinger *et al.* 1996) and Pol II (Cramer *et al.* 2001; Gnatt *l.* 2001) have been elucidated by electron microscopy and x-ray crystallography. These studies anced the understanding of the mode of action for these polymerases (Buratowski 2000). For II, it was shown that the core subunits are structurally very similar to those of *E.coli* RNAP and n a cleft where the active site is located. Comparing the location of the magnesium ion catalytic with the cleft identified for DNA entry and pores for RNA exit suggest that DNA does not ow a straight path through the enzyme (Lee and Young 2000). Similarly, the main structural ure of Pol I at 30 Å resolution is a 10 nm long groove and finger like domain that form a 2.5 nm e channel demonstrating structural homology to Pol II and *E.coli* (Schultz *et al.* 1993). The ctural conservation between polymerases probably reflects the high degree of sequence ology and functional conservation between the two large subunits of these enzymes.

Chromatin

vo, DNA is wound around multi-protein complexes called nucleosomes that are further wrapped chromatin. While compacting DNA to fit in the nucleus, nucleosomes also affect the ability of A to interact with other proteins that may not be able to recognize or reach tightly bound ences. The degree of compaction is not uniform over the entire genome with the most densed regions, called heterochromatin, being generally transcriptionally inert. In addition, the

increased rate of transcription seen in naked DNA (or DNA without nucleosomes) suggests that nucleosomes are general repressors of transcription (Morales *et al.* 2001).

The eukaryotic transcription apparatus have evolved to operate in this chromatin environment. Several multi-protein complexes have been identified that modify chromosome structure, altering DNA packaging and thus DNA accessibility (Orphanides and Reinberg 2000; Svejstrup 2002). Of these, ATP-dependent remodelling complexes modify chromatin structure non-covalently. As a result of these ATPase modifications, nucleosomes are moved allowing the exposure or occlusion of DNA sequences. The two most well characterised families of ATP-dependent remodelling complexes are the SWI/SNF and the ISWI-based families of complexes that appear to function via different mechanisms as well as on different templates (for review see Narlikar *et al.* 2002).

In contrast, histone-modifying complexes add or remove covalent modifications from histone tails thus altering the chromatin density. Covalent modification such as acetylation, methylation, ubiquitination and phosphorylation are shown to be crucial for regulation in transcription (Berger 2001). Evidence suggests that hyperacetylation in promoter proximal regions of specific genes is correlated with an increase in gene expression while hypoacetylation is connected with gene repression (Kuo *et al.* 2000; Khochbin *et al.* 2001). In acetylation, the steady state level is maintained by the opposing actions of histone acetyltransferase (HAT) and histone deacetyltransferase (HDAC) complexes resulting in a complex mechanism of transcriptional control through increasing or decreasing the repression that results from chromatin (Reid *et al.* 2000; Vogelauer *et al.* 2000; Narlikar *et al.* 2002).

1.3 Transcription Phases

As this dissertation does not focus on Pol III, this introduction will describe only genes transcribed by Pol I and Pol II.

(a) Initiation

Eukaryotic transcriptional initiation depends on *cis*-acting elements and *trans*-acting factors. The *trans*-acting accessory factors recognise the *cis*-acting promoter sequences required for the proper identification of promoter regions.

Cis-acting sequence elements

Despite the fact that each polymerase recognizes a distinct gene type, the same basic principles can be seen across eukaryotic promoters. Promoter regions for Pol I and Pol II genes primarily comprise two types of sequence element (Reeder 1999). Core elements are positioned close to the transcription start site and define the site of initiation while further upstream elements bind transcription factors that control the rate of transcription. In *S.cerevisiae* both Pol I and Pol II transcribed genes have two critical *cis* domains that are involved in initiation primarily by directing assembly of the PIC.

Pol I

In *S.cerevisiae*, the Pol I promoter region contains a specific core element together with an upstream element which are separated by variable sequence but critical distance (Reeder 1999). The spacing and helical relationship between the two elements is critical as half helical insertions or deletions ($\pm$ 5 or $\pm$ 15 bp) are far more deleterious than those that retain the helical relationship suggesting that protein interactions are crucial (Pape *et al.* 1990; Xie and Rothblum 1992). The core element, which on its own is necessary and sufficient for the initiation of basal transcription, overlaps the transcription initiation site by 4-6 bp and extends to a maximum distance of approximately 50 bp upstream of the initiation site (Reeder 1999; Paule and White 2000). The core promoter region contains the ribosomal initiator (rInr), the only rRNA promoter sequence element conserved across species. The rInr is an AT-rich, TATA-like sequence surrounding the site of initiation (Paule and White 2000). Even though the rInr sequence resembles a TATA box, it is not a binding site for TATA binding protein (TBP).

The upstream promoter element (UPE) extends from about 90 bp upstream of the initiator site to a maximum distance of 150 bp upstream of the initiation site. While this element significantly stimulates transcription driven by the core promoter, in most instances it is not absolutely required for transcription initiation *in vitro* (Paule and White 2000). In *S.cerevisiae*, located just upstream of the UPE is a binding site for the Reb1p, a protein that plays an orientation specific role in rRNA gene termination. The promoter bound Reb1p has been implicated in many functions including protection from 'wandering' polymerases (Bateman and Paule 1988; Henderson *et al.* 1989), remodelling of chromatin over the promoter (Langst *et al.* 1998) and folding of the rRNA repeats in the nucleolus (Sander and Grummt 1997; Mason *et al.* 1998). However, the Reb1p binding site located upstream of the UPE is in the opposite orientation from that known to function in termination, therefore, it appears to stimulate initiation without protecting the promoter from interference (Lang and Reeder 1993; Paule and White 2000).

In *S.cerevisiae* there is a single 190 bp enhancer element that is approximately 100 bp downstream of the transcription unit, overlapping the terminator element in the rDNA intergenic sequence (IGS) (Elion and Warner 1984; Elion and Warner 1986; Elion and Warner 1986). The enhancer element can act bi-directionally and at great distances from the gene promoter increasing the level of transcription from the rRNA genes *in vitro* by 10 to 50 fold (Banditt *et al.* 1999). Chromatin studies show actively transcribed rRNA genes, which are nucleosome free, to have downstream enhancer regions that are also free of nucleosome. Nucleosome-free enhancers have also been observed in yeast strains where Pol I is non-functional, suggesting a mechanism mediated by protein binding rather than transcription (Dammann *et al.* 1995).

The rRNA gene enhancer stimulates transcription through an uncertain mechanism. It has been suggested that the enhancer functions in the first steps of transcription initiation by helping to stabilise the PIC and influence the probability that a PIC will assemble on the adjacent gene (Reeder and Lang 1994; Osheim *et al.* 1996; McStay *et al.* 1997). In this model it is assumed that once the promoter complex has been formed the enhancer is dispensable (Schultz *et al.* 1993). However, a second model proposes that the enhancer increases the rate of transcription of already active genes

rather than increasing the number of genes activated (Morrow *et al.* 1993). The enhancer, with Reb1p bound upstream of the promoter, could function by looping the IGS to bring together the termination region of one repeat with the promoter of the other (Johnson and Warner 1989; Kulkens *et al.* 1992). This model is supported by evidence suggesting that the enhancer does not play a role in the activation frequency of the genes (ie the number of repeats activated) but rather increases the rate of reinitiation on already active promoter elements (Banditt *et al.* 1999). Recent evidence suggests a third model where the enhancer allows Pol I transcription of rRNA genes not properly localized in the nucleolus to be transcribed more efficiently (Wai *et al.* 2001).

Pol II

In Pol II transcribed genes there are also two distinct promoter elements recognized by regulatory proteins directing transcriptional initiation. Gene promoters can contain both, one or neither of these elements with the sequences and spacing variable between promoters (Lee and Young 2000). The core promoter element consists of the TATA-box (25-30 bp upstream of the start site in higher eukaryotes; 40-120 bp in *S.cerevisiae*) and the transcription initiation start site that may contain an initiator motif (Inr) where regulatory proteins bind. The TATA-box consensus sequence for *S.cerevisiae* is TATAAA with derivatives of this sequence acting with diminished efficiency (Hampsey 1998). There are some reported 'TATA'-less promoters that do not have identifiable TATA-boxes. Interestingly, these promoters still require TBP to initiate, which suggests that they may bind TBP less efficiently (Hampsey 1998). The start site of transcription is defined by the Inr elements. In higher eukaryotes the spacing of the promoter elements appears to play a role in selecting the site of initiation, typically 30 bp downstream of the TATA-box (Lee and Young 2000). In *S.cerevisiae* the distance between the TATA-box and the site of initiation is more variable. The TATA-box appears to create a region between 120-40 bp downstream where Inr elements can direct initiation (Hampsey 1998). There is no distinct Inr sequence but preferred sequences have been identified (Furtergraves and Hall 1990).

Regulatory elements, or sequences bound by transcriptional regulators, are usually upstream elements consisting of gene specific sequences involved in stimulating basal transcription. These elements

include upstream activating sequences (UAS) and upstream repressor sequences (URS) (see review in Blackwood and Kadonaga 1998). UAS elements are bound by proteins that interact either directly with general transcription factors (see below) or indirectly with co-activators to increase the rate of transcription. The UAS and URS elements as well as other proximal promoter elements can be within 50 bp to a few hundred bp upstream of the start site in *S.cerevisiae* (Blackwood and Kadonaga 1998).

General Transcription Factors (GTFs)

GTFs are defined as factors required for accurate, basal-level transcription *in vitro*. While there are similarities in the different polymerase activities, overall, Pol I requires fewer transcription factors for initiation than Pol II.

Pol I

To achieve full levels of transcription initiation of rRNA genes in *S.cerevisiae* TBP, Rrn3p, Core Factor (CF), Upstream Activating Factor (UAF) and Pol I in conjunction with two critical *cis* promoter regions are required (Nomura 1998; Aprikian *et al.* 2001) (see Figure A). The minimal requirements for basal *in vitro* transcription are CF, Rrn3p and Pol I. The presence of TBP and UAF stimulate *in vitro* transcription between 10 to 50 fold (Keener *et al.* 1998). *In vivo*, lethality is seen in the absence of CF, Rrn3p, Pol I and UAF while some but not all TBP mutants are completely defective in rRNA synthesis (Keener *et al.* 1998).

RNA Polymerase I Promoter

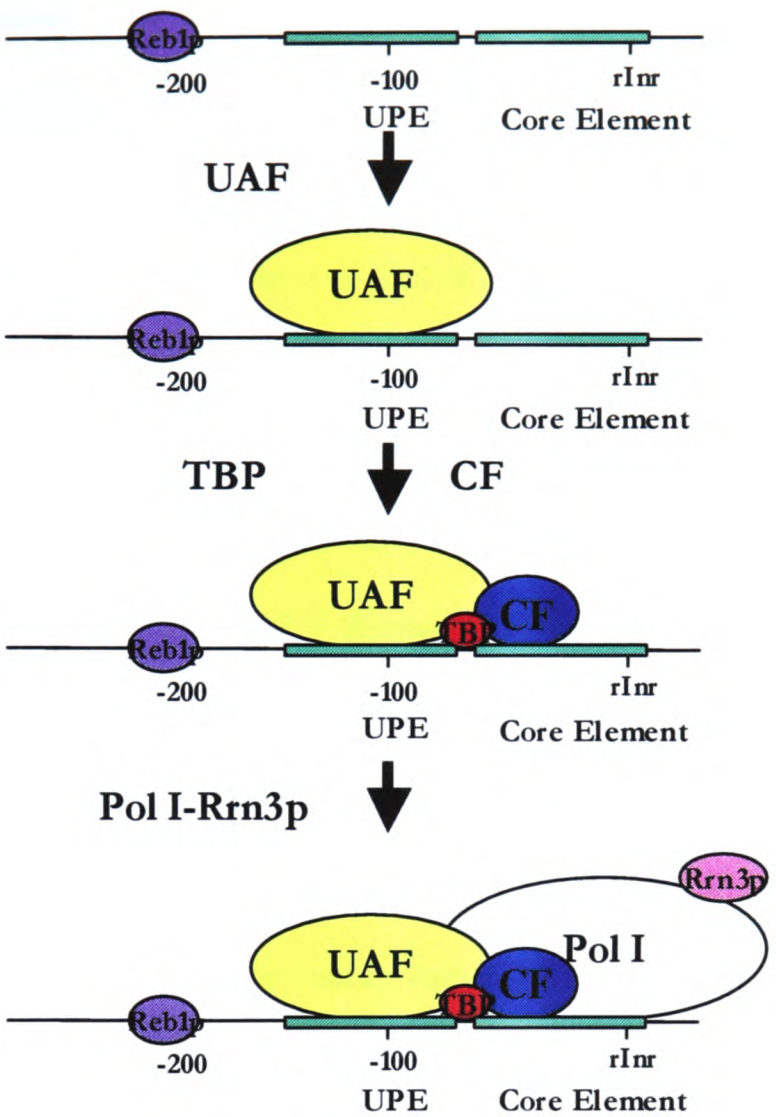


Figure A Schematic model of the assembly of a Pol I transcription complex in *Saccharomyces cerevisiae*. The upstream element (UPE) is bound by the upstream activating factor (UAF, yellow). The core factor (CF, blue) is then recruited to the promoter in a step that involves direct contact with UAF and CF but also involves TBP (red) acting as an intermediary between these two factors. After CF is in position, Pol I associated with Rrn3p (pink) can join the complex. The promoter proximal Reb1p protein (purple) is bound in the opposite orientation to induce termination. Additional uncharacterised factors may also be involved. Diagram reproduced from White (2001).

The CF interacts with the core domain of the promoter to select the site of initiation and possibly aid in the recruitment of Pol I (Paule and White 2000; White 2001). CF consists of four polypeptides (Rrn6p, Rrn7p, Rrn11p) and TBP, which interacts functionally but not stably as part of the core factor (Keys *et al.* 1994; Lalo *et al.* 1996; Lin *et al.* 1996; Aprikian *et al.* 2000). When CF is not present, Pol I can inefficiently contact any DNA sequence initiating randomly (Paule and White 2000).

The UPE is bound by UAF, which aids PIC formation and stimulates transcription (Paule and White 2000). UAF contains Rrn5p, Rrn9p, Rrn10p, histones H3 and H4 and another recently characterised protein Uaf30p (Keener *et al.* 1998; Paule and White 2000; Siddiqi *et al.* 2001). UAF, which is thought to bind the promoter first and remain after initiation, mediates the efficient recruitment of the CF,

Pol I and Rrn3p (Paule and White 2000). While unable to direct transcription on its own and not required for specific initiation, UAF is able to stimulate transcription directed by the core domain while also playing a role in stabilizing the PIC to the promoter region (Keys *et al.* 1996; Aprikian *et al.* 2001). Additionally, UAF has been implicated in suppression of transcription from cryptic Pol II promoters located in the IGS of rDNA repeats (Oakes *et al.* 1999; Siddiqi *et al.* 2001).

TBP, highly conserved among species, binds the minor grove of the TATA element in Pol II promoters, dramatically distorting DNA. The role of TBP in Pol I PIC formation is unclear, however, *in vitro* it interacts with UAF and CF through the Rrn6p and Rrn9p subunits respectively (Lin *et al.* 1996; Steffan *et al.* 1996). TBP can be recruited to the PIC in the absence of UAF presumably through interactions with CF (Aprikian *et al.* 2001). Conversely, TBP can also be recruited to the promoter in absence of CF presumably by binding UAF (Aprikian *et al.* 2000). *In vivo* TBP, as well as UAF, is necessary for the incorporation of CF into the PIC and an interaction between TBP and UAF is required for rRNA transcription stimulation (Steffan *et al.* 1996; Steffan *et al.* 1998). Accordingly, TBP probably has a different role in Pol I transcription than it has in Pol II and III (Keener *et al.* 1998).

The protein Rrn3p, which co-precipitates with Pol I, is required for rRNA transcription *in vivo* (Yamamoto *et al.* 1996; Reeder 1999). Rnr3p has been shown to interact with the Rrn6p subunit of CF and the Rpa43p (A43) subunit of Pol I (Yamamoto *et al.* 1996; Peyroche *et al.* 2000). Rrn3p is incorporated in the PIC in transcriptionally active rRNA genes and may be involved in the early steps of promoter escape and RNA elongation or regulated in response to entry into stationary phase (Milkereit and Tschochner 1998; Reeder 1999; Aprikian *et al.* 2001). Recent studies on Rrn3p, and its functionally conserved human homolog, suggest that it bridges Pol I and CF promoting productive integration of Pol I into the PIC (Peyroche *et al.* 2000; Aprikian *et al.* 2001). Interestingly, phosphorylation of the Rpa43 (A43) subunit of Pol I increases complex formation with Rrn3p and subsequently increases transcriptional initiation demonstrating that different phosphorylation states can be used to modulate rRNA gene transcription in a reversible manner (Aprikian *et al.* 2001; Fath *et al.* 2001).

Pol I formation of preinitiation complex

It has been proposed that the PIC formation is step-wise, with UAF recruiting CF, which in turn recruits Pol I-Rrn3p (Steffan *et al.* 1996). However, recent evidence that UAF cannot recruit CF in the absence of the Pol I-Rrn3p complex suggests that transcriptionally competent Pol I is required for stable association of the CF with the promoter for PIC formation (Aprikian *et al.* 2001; Bordi *et al.* 2001). UAF binding to the promoter, which occurs in the absence of active transcription, initiates transcription complex formation at rRNA promoters (Vogelauer *et al.* 1998).

Pol I Holoenzyme

Evidence in plants and mammals suggests that a Pol I 'holoenzyme', or a large pre-assembled complex containing RNA polymerases and variable sets of GTFs and co-factors capable of supporting specific initiation *in vitro*, may play a role in rRNA gene transcription (Kenzior and Folk 2001). While no yeast Pol I holoenzyme has yet been identified, recent *in vitro* evidence does not preclude the formation of a Pol I-Rrn3-CF complex capable of binding UAF loaded promoter (Aprikian *et al.* 2001). Furthermore, Pol I exists in two forms. One Pol I fraction contains the Rrn3p, and is capable of promoter specific transcription. A second, much larger fraction of polymerases, including those isolated from cells grown to stationary phase, lacks Rrn3p (Yamamoto *et al.* 1996; Milkereit and Tschochner 1998). However, whether either of these complexes fits the definition of a 'holoenzyme' is yet to be determined.

Pol II

Initiation at Pol II genes is a multi step process requiring five conserved general transcription factors (TFIIB, TFIID, TFIIE, TFIIF and TFIIH), Pol II and at least 30 additional polypeptides (Roeder 1996; Dvir *et al.* 2001). TFIIA was originally classed as a GTF but acts as a co-activator of transcription and is dispensable for accurate initiation (Hampsey 1998). Basal transcription initiation *in vivo* requires additional factors or co-activators for Pol II complex recruitment and chromatin remodelling.

Pol II formation of preinitiation complex

Initiation of transcription of protein-encoding genes *in vitro* requires the assembly of a PIC, composed of template DNA, Pol II and the five GTFs described above (Woychik and Hampsey 2002). GTFs assemble, regulated by transcriptional activators and repressors, forming a stable PIC on promoter DNA (see Figure B). There are two models for Pol II PIC assembly: i) a stepwise mechanism where each factor binds sequentially and cooperatively (Orphanides *et al.* 1996), and ii) pre-assembly of a holoenzyme containing all basal and regulatory factors and Pol II which associates with the promoter in a single step (Koleske and Young 1995). In reality, the mechanism is probably a mixture of the two models with some, but not all, basal factors associating with the polymerase before binding to the promoter (Buratowski 2000).

RNA Polymerase II Promoter

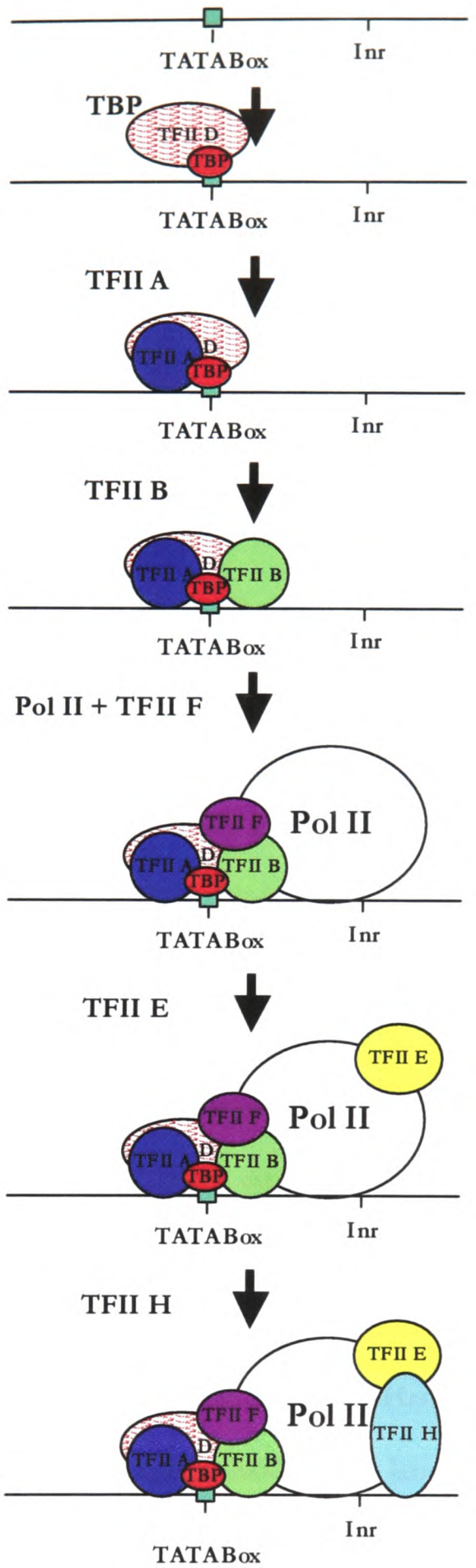


Figure B Order of assembly of the basal Pol II transcription factors according to the step-wise PIC formation model. Promoter recognition is achieved by TATA box binding by TBP (red), which is part of the TFIID (pattern red). DNA-bound TBP recruits TFIIA (dark blue) and TFIIB (green). TFIIB subsequently provides the docking site for the preformed complex of TFIIF (purple) bound to Pol II. TFII E (yellow) and TFII H (teal) are recruited. Simultaneous assembly of many of these factors may occur through the recruitment of a preformed holoenzyme complex containing Pol II. Diagram reproduced from White (2001).

Stepwise PIC formation requires TATA box (when present) recognition by TBP, which is contained in TFIID with at least eight other TBP-associated factors (TAFs) (White 2001). As a universal transcription factor (required for initiation in Pol I, Pol II and Pol III), TBP nucleates the PIC and

subsequently recruits TFIIB to the promoter. TFIIB binds to the TFIID/TATA-complex by direct interaction with TBP and the DNA on either side of the TATA element to form an intermediate complex (Dvir *et al.* 2001). TFIIA enters the complex any time after TFIID, binding upstream DNA sequences and relieving repressors associated with TBP (White 2001).

While stabilising the PIC, TFIIB also recruits the pre-formed RNA Pol II/TFIIF-complex to the promoter bound TFIID (Hampsey 1998). Pol II alone has a low affinity for the TFIID-TFIIB-promoter complex. However, TFIIF binds to RNA Pol II with high affinity prior to being recruited to the promoter complex by TFIIB (Hampsey 1998). This serves to prevent non-specific promoter binding by Pol II while also stabilizing the intermediate complex and recruiting TFIIE and TFIIH (Roeder 1996; Dvir *et al.* 2001). TFIIE, an ATP dependent helicase, has a high affinity for Pol II and may enter the PIC with Pol II and TFIIF (Hampsey 1998). TFIIH is recruited to the bound TFIIE. These two factors play an important role in DNA unwinding and promoter clearance.

Pol II Holoenzyme

An alternate view of PIC assembly is that of a pre-assembled holoenzyme which is targeted to the promoter in a single step. There have been several different holoenzymes proposed composed of a subset of GTFs (TFIIB, TFIIE, TFIIF, and TFIIH), Pol II and ancillary factors including sub-complexes of SRB (Suppressors of RNA polymerase B) and Mediator complex proteins in addition to a sub-complex that includes Gal11p, Rgr1p, Sin4p, Pgd1/Hrs1p, Rox3p, Swi/Snf and Nut1p (Myer and Young 1998; Woychik and Hampsey 2002). The large size of such a complex (>70 polypeptide) has led to the suggestion of 'transcription factories' that associate with the nuclear matrix (Cook 1994). However, the existence of a Pol II holoenzyme is contested since most proposed complexes lack the complement of GTFs required for initiation. Evidence also suggests Mediator, TFIID, TFIIA, TFIIH, TFIIE and gene specific activators remain on the promoter to facilitate reinitiation (Yudkovsky *et al.* 2000; Woychik and Hampsey 2002). Furthermore, holoenzyme factors are found in un-stoichiometric amounts, which may argue against a pre-assembled Pol II holoenzyme structure (Borggrefe *et al.* 2001; Park *et al.* 2001). Additional theories suggest that there may be multiple complexes selectively associated with Pol II under specific cellular

conditions, helping to explain evidence that a subset of genes can be transcribed with no SRB/Mediator complex (Hampsey and Reinberg 1999).

Pol II Transcription Factors (TFs)

Transcription factors (TFs) or transcription activators typically consist of two domains: one that possesses a high affinity for DNA and one that interacts with the transcription apparatus (Lee and Young 2000). Because of this dual function TFs play a crucial role in recruiting the transcription apparatus to the promoter. Once the transcription apparatus has been recruited, DNA contortion caused by the bound activators enables TBP to interact directly with DNA. Additionally, activators are involved in inducing conformational changes in the PIC components to stimulate promoter clearance and transcript elongation (Orphanides *et al.* 1996; Lee and Young 2000).

Similar to PIC formation, there is a spectrum of models that describe TF dependent recruitment of the transcription apparatus to promoters (Lee and Young 2000). At one extreme the apparatus is recruited factor-by-factor whereby the activators would affect the rate-limiting step. In this 'activation by recruitment' model, PIC assembly is directly stimulated by activators or indirectly stimulated by co-activators binding to the promoter DNA in a sequence specific manner (Kim *et al.* 1998). At the other extreme, in the holoenzyme model, bound activators only need to contact a single component of the complex. Numerous cases of physical contact between activators and transcription machinery components have been observed, especially with TFIIA, TFIIB and TFIID (Orphanides *et al.* 1996) and it is likely that different promoters utilize different mechanisms along this spectrum (Lee and Young 2000).

Pol II Transcription Co-activators

Co-activators are transcription factors that mediate activation. These are distinct from GTFs and TFs because they are dispensable for basal-level transcription *in vitro* and do not bind DNA respectively. The co-activator Mediator is a large modular complex that transduces positive and negative regulatory information between gene-specific TFs and repressors to the core transcription machinery (Myer and Young 1998; Hampsey and Reinberg 1999). Mediator appears to be required for transcription of all genes *in vivo* and only associates at the promoter in the initiation complex

(Yudkovsky *et al.* 2000; Pokholok *et al.* 2002; Svejstrup 2002). It has been suggested that Mediator be included with GTFs and Pol II in the minimal components required for regulated transcriptional initiation (Liu *et al.* 2001; Woychik and Hampsey 2002).

Pol II transcription repressors

Transcriptional repressors are either general, such as chromatin structure, or gene specific. These can function by binding to activators or by competing for activator binding sites (Lee and Young 2000). In some cases, the loss of function of an activator can be compensated for by the loss of function of one of several repressors demonstrating the balance that exists between the positive and negative regulatory factors (Lee and Young 2000).

(b) Open complex formation

During initiation, RNA polymerase recognizes the DNA template, a stable transcription initiation complex is formed, local unwinding generates a transcription bubble and the first nucleotide bond is synthesized.

Pol I

Once a Pol I promoter is open, the rate of Pol I loading is controlled by a secondary mechanism that may involve the levels of Rrn3p (Aprikian *et al.* 2001). Promoter melting in rDNA occurs spontaneously upon binding of Pol I in a temperature-dependent manner and is not dependent on ATP hydrolysis (Paule and White 2000). Following promoter clearance, fully functional UAF complex is left on the promoter, enabling highly efficient repetitive polymerase recruitment and a high density of transcribing polymerase on the active repeats. (Paule and White 2000; Aprikian *et al.* 2001).

Pol II

After full assembly of the PIC on Pol II transcribed genes, an ATP-dependent DNA helicase activity associated with TFIIH melts the DNA at the transcriptional start site to create an open complex (Roeder 1996). Additionally, TFIIE is also required for ATP-dependent formation of the open

promoter complex and may help to stabilise the single stranded DNA while the open complex initiates RNA transcript synthesis (Conaway and Conaway 1993; Dvir *et al.* 2001; Woychik and Hampsey 2002).

The precise mechanism that controls the switch from initiation to elongation is unknown. TFIIH kinase activity is partially responsible for phosphorylating the CTD of the large Pol II subunit from the hypophosphorylated form to the hyperphosphorylated form enabling promoter clearance and facilitates the transition to the elongation phase of transcription (Lee and Young 2000). Formation of the first phosphodiester bond allows the open complex to be extended and continued elongation results in promoter clearance (Orphanides *et al.* 1996; Roeder 1996). ATP hydrolysis and TFIIH are necessary for the disassociation of the PIC following initiation (Yudkovsky *et al.* 2000). By the time the transcription complex reaches +30, Pol II has made the crucial transition to a fully competent elongation form. During early elongation TFIIB and TFIIF disassociate from the promoter, leaving TFIID, TFIIA, TFIIH, Mediator and gene specific activators on the promoter to stimulate re-initiation (Zawel *et al.* 1995; Yudkovsky *et al.* 2000). As Pol II begins the elongation phase of transcription factors involved in processing of the pre-mRNA enter the complex (discussed in more detail later).

(c) Elongation

Elongation is loosely defined as the phase following the formation of the first phosphodiester bond and encompasses numerous other changes that occur to complete promoter escape and the formation of elongation competent polymerases (Lee and Young 2000). Elongation involves the catalysis of successive nucleoside monophosphates into an oligoribonucleotide transcript by extension of the RNA chain in a 5' to 3' direction.

The elongating polymerase must be processive enough to transcribe genes several thousand base pairs in length, agile enough to surmount transcriptional impediments resulting from such things as chromatin structure or DNA polymerase coming from both directions and responsive enough to recognise specific signals directing the polymerase to pause or terminate transcription (Uptain *et al.*

1997; Lee and Young 2000). Intrinsic sequence elements and *trans*-acting factors can influence the elongating polymerase complex either through direct interaction with the elongating complex or by creating impediments to processive elongation (Reines *et al.* 1999; Conaway and Conaway 2000; Conaway *et al.* 2000).

Mechanism of elongation

There are several models for the mechanism of transcriptional elongation by *E.coli* RNAP. In the 'monotonic' model, the polymerase elongates by a single nucleotide maintaining a constant number of nucleotide contacts, until termination results from the partial disruption of the RNA-DNA hybrid by a hairpin formation (Yager and Vonhippel 1991). Also proposed is the 'discontinuous elongation' or 'inchworm' model where alternate locking and sliding of front and back DNA and RNA contacts assure that the enzyme complex always maintains at least one tight contact with the template (Uptain *et al.* 1997). Combining these is the 'sliding-clamp' or 'double-clamp' model where the polymerase encircles the duplex DNA ahead of the transcription complex enabling backsliding to increase the stability of the RNA-DNA hybrid (Landick 1997; Nudler *et al.* 1997; Nudler *et al.* 1998). Evidence confirms the 3' proximal RNA-DNA hybrid plays a role in elongation complex stability with specific sequences influencing the extent of backsliding that results from paused or arrested complexes (Nudler 1999). While these models are based on prokaryotic RNAP, similarities between the structure and function of transcription allow for speculation that similar mechanisms are employed in eukaryotes.

Promoter escape and early elongation

Two distinct classes of early elongating complexes exist: one that undergoes abortive elongation giving rise to short transcripts and a second class that carries out productive elongation (Shilatifard 1998). Very early elongation complexes of less than approximately 9 nt are unstable and susceptible to abortive transcription as each step of nucleotide addition is a competition between disassociation of the elongation complex and formation of the next phosphodiester bond (Shilatifard 1998).

Pol I

Following initiation, UAF remains bound to the promoter while Pol I, Rrn3p and CF and TBP escape the promoter (Aprikian *et al.* 2001). Although almost all of the Rrn3p is released, only a small fraction of CF and TBP are released suggesting a transition to transcriptionally inactive PICs that do not contain Rrn3p (Aprikian *et al.* 2001).

Pol II

After initiation, a scaffolding remains at the Pol II promoter to stimulate further rounds of initiation. This scaffolding has been suggested to consist of the gene specific activator, TFIID, TFIIA, TFIIH and Mediator (Yudkovsky *et al.* 2000). However, recent evidence suggests that the factors involved in initiation all disassociate before elongation with only Spt5p and Paf1p/Cdc73p associated with the elongating transcription complexes *in vivo* (Pokholok *et al.* 2002).

Phosphorylation of the CTD

Over 15 years ago it was discovered that the phosphorylation of the heptapeptide repeats of the CTD of Pol II played a crucial role in the transition from initiation to elongation (Dahmus 1996). Reversible phosphorylation of the CTD allows the conversion from a Pol II that is includable in the pre-initiation complex (hypophosphorylated) to the form that is seen in elongating complexes (hyperphosphorylated). Additionally, the hyperphosphorylated form of the CTD plays a role in recruiting the mRNA capping enzyme to the nascent transcript with capping occurring shortly after promoter clearance (Cho *et al.* 1997). The switch in phosphorylation status that occurs between initiation and elongation appears to cause Pol II to switch co-factors allowing differentiation between initiation competent and elongation competent Pol II complexes (Lee and Young 2000).

Elongation factors

A number of elongation factors are thought to affect the rate of elongation (Shilatifard 1998; Conaway *et al.* 2000; Orphanides and Reinberg 2000). Two models have been proposed to explain how elongation factors function to alleviate barriers imposed by chromatin (Orphanides and Reinberg 2000). In the first model, the elongation factors promote decompaction over the whole

gene thus allowing the elongating polymerase to encounter partially decompacted nucleosomes as it transcribes. In the second model, the elongation factors promote complete decompaction over the promoter region and partial decompaction over the remainder of the gene leaving the elongating polymerase to face a partially compacted template. In this latter model, the first polymerase to transcribe a gene is a 'pioneer' and subsequent polymerases are somehow different. This is supported by evidence that, following initiation, an intermediate containing several initiation factors is left at the promoter suggesting that the loading of subsequent polymerases on to the promoter is different from the first round of initiation (Orphanides and Reinberg 2000; Yudkovsky *et al.* 2000; Liu *et al.* 2001).

Pol I

In Pol I, efficient elongation requires additional constitutive activities. For example, elongation through chromosomal rDNA requires the function of either topoisomerase I or topoisomerase II (Kim and Wang 1989). If there is a shortage of topoisomerase activity the rDNA repeats become extrachromosomal suggesting that for proper rRNA gene transcription there is a need to remove torsional strain created by transcriptional supercoiling changes (Kim and Wang 1989).

Pol II

Unlike the absolute requirement of GTFs for initiation, transcript elongation can occur in the absence of accessory factors *in vitro*. However, many factors have been isolated through their ability to stimulate transcript elongation (Shilatifard 1998). These elongation factors are classified in three broad categories: those that reactivate arrested elongation complexes, those that prevent transient pausing by the enzyme and increase the rate of elongation or those that regulate the rate of elongation through chromatin (Shilatifard 1998; Conaway and Conaway 2000).

The first category of elongation factors reactivate arrested elongation complexes thought to be the result of slippages that generate 3' overhangs of the RNA. Members of the TFIIS family have been demonstrated to restart an arrested polymerase by activating an endonuclease intrinsic in Pol II (Uptain *et al.* 1997; Wind and Reines 2000). This cleaves the nascent transcript at the 3' end and so

creates a newly aligned 3' end with respect to the catalytic site (Conaway and Conaway 2000). It has been proposed that TFIIS binds to an arrested Pol II, the nascent transcript is hydrolysed at an internal phosphodiester bond, a conformation change facilitates cleavage and nucleotide incorporation continues (Conaway *et al.* 2000; Wind and Reines 2000). This set of ordered events is consistent with recent studies indicating that TFIIS associates with elongating Pol II only when the cell is under stress (Pokholok *et al.* 2002).

A second category of elongation factors prevents transcriptional arrest. These proteins are thought to increase the catalytic rate of transcription elongation by decreasing the fraction of time spent in the inactive paused conformation (Conaway *et al.* 2000). Elongation factors such as TFIIF, Elongin, CSB protein, Tat-Sf1, CA150 proteins and the ELL complex have been identified as functioning this way (Conaway *et al.* 2000). The discrepancy of elongation rates of *in vitro* and *in vivo* transcription suggest that elongation factors which suppress pausing may play a crucial role in facilitating mRNA synthesis (Conaway *et al.* 2000).

The third class of elongation factors are loosely defined as factors that affect Pol II transcript elongation through chromatin but do not effect initiation or elongation on a naked DNA template (Svejstrup 2002). ATP-dependent chromatin remodelling and HAT activities probably cooperate to overcome the transcriptional repression that is imposed by chromatin packing (Orphanides and Reinberg 2000). As discussed earlier, the chromatin remodelling complex Swi/Snf promotes elongation through nucleosomes (Brown *et al.* 1996). Additionally, Elongator has also been shown to have HAT activity and can assist in chromatin modification in advance of elongating polymerase (Otero *et al.* 1999; Svejstrup 2002).

DNA dependent pause, arrest and transcription blocks

Pausing is an important regulatory mechanism in both transcriptional elongation and termination. During a transcriptional pause, the polymerase temporarily stops but remains catalytically active while bound to the template (White 2001). Pause signals are multipartite and can consist of either intrinsic RNA or DNA sequence elements or extrinsic *trans*-factors. Intrinsic signals such as RNA

secondary structure, DNA signals downstream of the pause site, the coding strand and the availability of nucleotide substrates have all been shown to effect pausing (Uptain *et al.* 1997). Extrinsic factors that can physically block polymerase transcription elongation have been described in both prokaryotes and eukaryotes (Lang *et al.* 1994; Uptain *et al.* 1997). In these cases, there is a correlation between the rate of disassociation of the DNA-binding protein and the ability to block elongation with a slow disassociation rate being more effective in blocking (Pavco and Steege 1991).

(d) Processing

Following Pol I and Pol II transcription, processing events mature the primary transcript into a functional product. In many cases, processing signals have been shown to play an important role in termination.

Pol I

Maturation of the 35S pre-rRNA requires numerous *trans*-acting factors in a multistep pathway that act on 10 known processing sites (Tollervey and Kiss 1997; Kressler *et al.* 1999). Interestingly, pre-rRNA transcripts do not appear to process at the nascent level (Allmang and Tollervey 1998). It has been speculated that a signal present in the 3' external transcribed sequence (ETS) could be required to initiate rRNA processing. However, deletion analysis of the 3' ETS did not reveal a sequence required to trigger rRNA gene maturation (Allmang and Tollervey 1998).

Shortly after formation of the pre-rRNA transcript covalent modification at many specific nucleotides occurs (Kressler *et al.* 1999). The pre-rRNA transcript is generated by Rnt1p mediated endonucleolytic cleavage at a position between +15 to +50 from the mature 3' end of the 25S gene (Kufel *et al.* 1999) (discussed in more detail in Part II). Cleavage is followed by exonucleolytic digestion to the +7 position and many additional processing events (Kempersveenstra *et al.* 1986; Elela and Ares 1998; Venema and Tollervey 1999). These modifications and other processing steps occur in the presence of ribosomal proteins forming pre-ribosomal particles in the nucleolus (for review see Venema and Tollervey 1999).

Pol II

To create a fully translatable transcript that is capable of export out of the nucleus, three processing events need to occur to the nascent Pol II transcript (for review see Zhao *et al.* 1999). These include the acquisition of a cap structure at the 5' terminus, splicing of any introns in the body of the transcript and generation of the 3' end usually by the addition of a poly(A) tail. While each step is a biochemically distinct event, it is now believed that most mRNA processing events occur co-transcriptionally and that they are all interlinked, effecting changes in specificity and efficiency to each process (for review see Proudfoot *et al.* 2002).

The first processing event is a three-step reaction that adds a cap structure to the 5' end of all mRNAs (Proudfoot *et al.* 2002). This takes place after 20-30 nt have been synthesised, beginning with the hydrolysis of the first triphosphate into a diphosphate by RNA 5' triphosphatase (RTP) (Cramer *et al.* 2001). Next, through an unusual 5'-5' triphosphates linkage, a guanylyltransferase (GT) fuses a GMP moiety from GTP to the first nucleotide of the pre-mRNA (Shatkin and Manley 2000). The process is completed by a methyltransferase (MT) methylating the N7 position of the transferred GMP moiety (Proudfoot *et al.* 2002). In yeast, unlike mammals, separate enzymes catalyse each of the three reactions creating an initial cap structure that is recognized by subsequent proteins in the cap-binding complex (CBC) (Cramer *et al.* 2001). Binding of the CBC is thought to play a major role in stabilization of the mRNA by obstructing the 5'-3' exonucleases (Beelman and Parker 1995). Furthermore, in the cytoplasm the cap is bound by translation initiation factors that promote the engagement of ribosomes and, therefore, enhance translation (Proudfoot *et al.* 2002).

Splicing, to remove introns from the pre-RNA, is another processing event that is involved in generating a functional mRNA (Proudfoot *et al.* 2002). While few yeast genes contain introns, the process is similar to that which occurs in mammals. Chemically, splicing takes place in two steps. First, the 2' OH of the branchpoint adenosine nucleophilically attacks the 5' exon-intron border followed by *trans*-esterification freeing the 5' exon and the lariat shaped intron with the 3' exon (Proudfoot *et al.* 2002). Second, the intron 3' exon border is attacked by the freed 3' OH. This is followed by *trans*-esterification that fuses the two exon sequences and releases the lariat-shaped

intron (Moore and Sharp 1993). Small nuclear RNAs (snRNAs) and several other proteins assemble over the intron to form the spliceosome, which facilitates splicing (Burge *et al.* 1998). Splicing requires that the transcribed pre-mRNA contain consensus elements in *cis* that mark the junctions or splice sites as well as the upstream branchpoint (Burge *et al.* 1998). The conserved splice site attracts the U1 snRNA, which initiates spliceosome assembly where complex RNA-RNA interactions between snRNAs and the pre-mRNA form the catalytic core (Madhani and Guthrie 1992; Sun and Manley 1995).

In addition to the processes described above, almost all protein encoding mRNAs require a uniform 3' end consisting of around 200 adenosine residues (Proudfoot *et al.* 2002). Cleavage of the pre-mRNA transcript must occur before the poly(A) tail formation can be directed by sequences within the pre-mRNA and by the polyadenylation machinery (Zhao *et al.* 1999). In most pre-mRNAs the cleavage site lies between the highly conserved AAUAAA hexamer and a downstream sequence element (DSE) rich in U or GU residues (Barabino and Keller 1999). The polyadenylation signals in budding yeast genes are less sequence specific than in mammalian systems. In mammals, an almost invariant AAUAAA hexonucleotide is located 15-20 nt upstream of the cleavage site. This works in conjunction with a less well-defined DSE to promote proper 3' end formation (Zhao *et al.* 1999). In *S.cerevisiae*, a more degenerate poly(A) signal is assisted by an upstream efficiency element (EE; consensus UAUAUA) to promote 3' end formation. Additionally, at 15-30 nt upstream of the cleavage site a positioning element (PE; consensus AAAAAA or AAUAAA) appears to direct cleavage at the preferred site Y(A)n (Irniger and Braus 1994; Guo *et al.* 1995; Guo and Sherman 1995; Guo and Sherman 1996; Guo and Sherman 1996).

In mammals, two multi-subunit protein complexes recognize the *cis* elements. Cleavage and polyadenylation specificity factor (CPSF) interacts with the AAUAAA hexamer and contributes to the stabilization and specificity (Shatkin and Manley 2000). The DSE interacts with the cleavage stimulation factor (CstF) and, by cooperatively binding, enhances the affinity of CPSF for the hexamer (Zhao *et al.* 1999). Cleavage factor I (CF I) and cleavage factor II (CF II) are also essential for cleavage of pre-mRNA. Additionally, poly(A) polymerase (PAP) is usually required for the

cleavage reaction and, with the help of CPSF, directs the addition of the adenosine residues (Dichtl and Keller 2001). The emerging poly(A) tail is bound by poly(A) binding protein (PABII) which enhances the processivity of the poly(A) polymerase (Wahle and Ruegsegger 1999). In yeast, despite the differences from mammalian polyadenylation sequence elements, notable conservation exists amongst the protein factors required for efficient cleavage and polyadenylation (Colgan and Manley 1997). In *S.cerevisiae* there are four *trans*-acting factors (or complexes) involved in polyadenylation: poly(A) polymerase (Pap1p), cleavage factor I (CF I which consists of CF IA and CF IB), cleavage factor II (CF II) and polyadenylation factor I (PF I) (Wahle and Keller 1996; Preker *et al.* 1997; Shatkin and Manley 2000) (See Figure C).

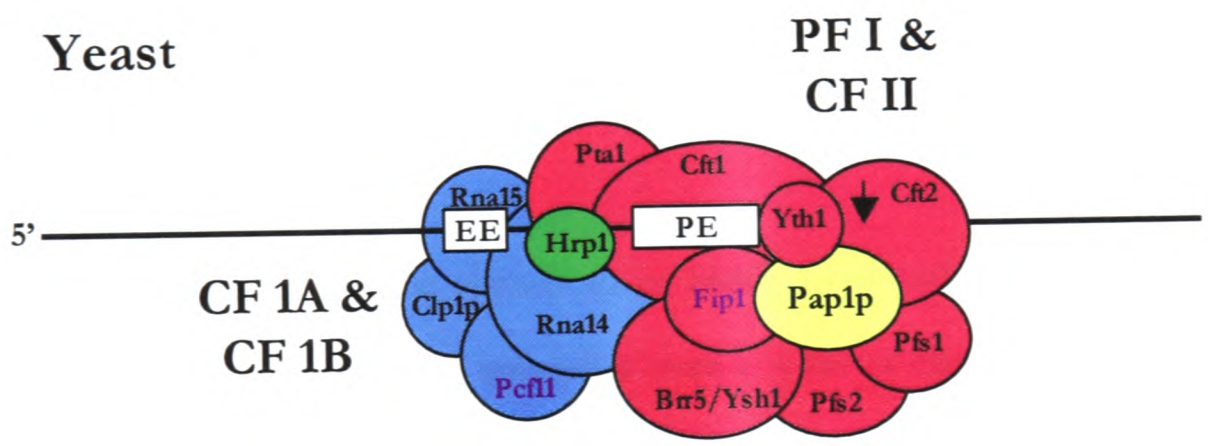


Figure C Diagrammed comparison of yeast and mammalian polyadenylation proteins. The known components are listed for both yeast (top) and mammals (bottom). The binding specifics for the yeast complexes are uncertain but are predictions from the mammalian system. Colours are used to show homologous protein complexes with yeast CF 1A/mammalian CstF shown in light blue; yeast PF II and CFII complex/ mammalian CPSF in red; yeast Pap1p/ mammalian PAP in yellow. Yeast Pcf11 and Fip1 lack mammalian homologs and denoted in purple letters. Symplekin, shown in darker blue, is associated with both CstF and CPSF but shares similarity with yeast Pta1. Other similarities are: CstF-77=Rna14; CstF-64=Rna15; CPSF-160=Cft1; CPSF-100= Cft2; CPSF-73=Brr5/Ysh1; and CPSF-30=Yth-1. Mammalian CF1 and yeast Hrp1, both otherwise unrelated proteins, are coloured in light and dark green respectively. The unpurified mammalian CF II is in grey. RNA signal sequences are boxed. Arrow indicates site of endonucleolytic cleavage. Diagram largely reproduced from Shatkin and Manley 2000.

(e) Termination and Release

Termination of transcription occurs when the ternary elongation complex, consisting of the template DNA, nascent RNA and polymerase complex, disassociates into its constituent parts, disengaging from the DNA template. In addition to properly ending the transcript, termination plays an important role in preventing transcriptional interference (Greger and Proudfoot 1998). In prokaryotes there are generally two types of termination events; factor dependent and factor independent with the 3' end of the nascent transcript representing the termination site (Henkin 2000). In eukaryotes, the process of termination differs between polymerases and the mature 3' end is not generally the site of termination. The exact sites of termination are difficult to define due to the instability of nascent transcripts.

Pol I

The mechanism for Pol I termination appears to be highly conserved from yeast to mammals (Reeder and Lang 1997). Transcription terminates over a series of terminator sequences found within the first 1000 bp downstream of the rRNA coding sequence. These terminators consist of at least two DNA sequences; an orientation specific element that binds a terminator protein and an upstream T-rich (on the non-coding strand) element that encodes the last 10-12 nucleotide of the terminated transcript (Reeder and Lang 1997). Termination of rRNA genes is thought to occur in a two-step process: pause and release where the paused transcripts are precursors to released transcripts (Lang *et al.* 1994; Reeder and Lang 1997).

The use and function of a DNA-binding protein in Pol I termination is conserved between species (White 2001). In all organisms studied so far, Pol I termination occurs 12-20 bp upstream of the binding site of the terminator protein (Reeder and Lang 1998). The terminator proteins, Reb1p in yeast and TTF-1p in vertebrates, have Myb homology DNA binding domains and appear to be able to induce termination in rRNA genes from any species (Evers and Grummt 1995). Reb1p, a transcription activator in certain Pol II promoters, has also been shown to rearrange nucleosomes in regions surrounding its binding site (Reeder 1999). Deletion analysis has shown that the C-terminal domain of the terminator protein is required for termination while the N-terminal end is involved in

processes unrelated to termination (Evers and Grummt 1995; Reeder and Lang 1997). Additionally, the proteins only function as a terminator when bound in an orientation specific manner (White 2001). There is evidence that pausing occurs as a result of a roadblock to Pol I, however, studies in yeast using a lac protein in place of Reb1p suggest that there might be a more complicated mechanism. In these experiments, a lac repressor protein was bound to the exact region of Reb1p inducing pausing of Pol I but not as efficiently as Reb1p (Jeong *et al.* 1995). While the precise mechanism for the 'pause' is still unresolved, the affinity of the DNA-binding protein is the primary regulator of the pausing (Reeder and Lang 1998). In *S.cerevisiae*, most mutations of the protein-binding site eliminate protein binding and pausing with the exception of a single mutation that increased the affinity of the protein binding and subsequently increased the pausing (Lang and Reeder 1995).

In rRNA genes the upstream termination element consists of a DNA sequence encoding the last 10–15 nt of the terminated transcript and is located 25 nt upstream of the Reb1p binding site (Reeder and Lang 1998). In *S.cerevisiae*, this T-rich element (in the non-template strand) has been shown to eliminate termination when mutated (Lang and Reeder 1995). The overall base composition of the upstream element, rather than a specific sequence appears to regulate the efficiency of transcript release (Reeder and Lang 1997). Taken together with the length of the upstream element this suggests a mechanism for disassociation of the transcript involving instability of the nascent RNA-DNA hybrid in the active site of the polymerase (Reeder and Lang 1997). It has been proposed that Pol I slippage may play a role in termination. After the polymerase backtracks from the pause site, instability increases due to base pair mismatching in the hybrid and the transcript is released (Paule and White 2000). In rRNA genes where the T-rich region has been replaced with a homopolymeric T element, Pol I undergoes severe iterative slippage and release is impaired (Jeong *et al.* 1996). Iterative slippage occurs when a RNA polymerase pauses near a homopolymer region in the template and the paused polymerase, catalytic site and nascent transcript slide backwards along the template for an unspecified number of nucleotides (see Figure D). The transcript rehybridizes and polymerisation proceeds forward again (Jeong *et al.* 1996). Interestingly, if there is as little as one

base pair mismatch at a critical region the RNA-DNA hybrid formed by the iterative slippage is compromised so the ternary complex becomes unstable resulting in release (Jeong *et al.* 1996).

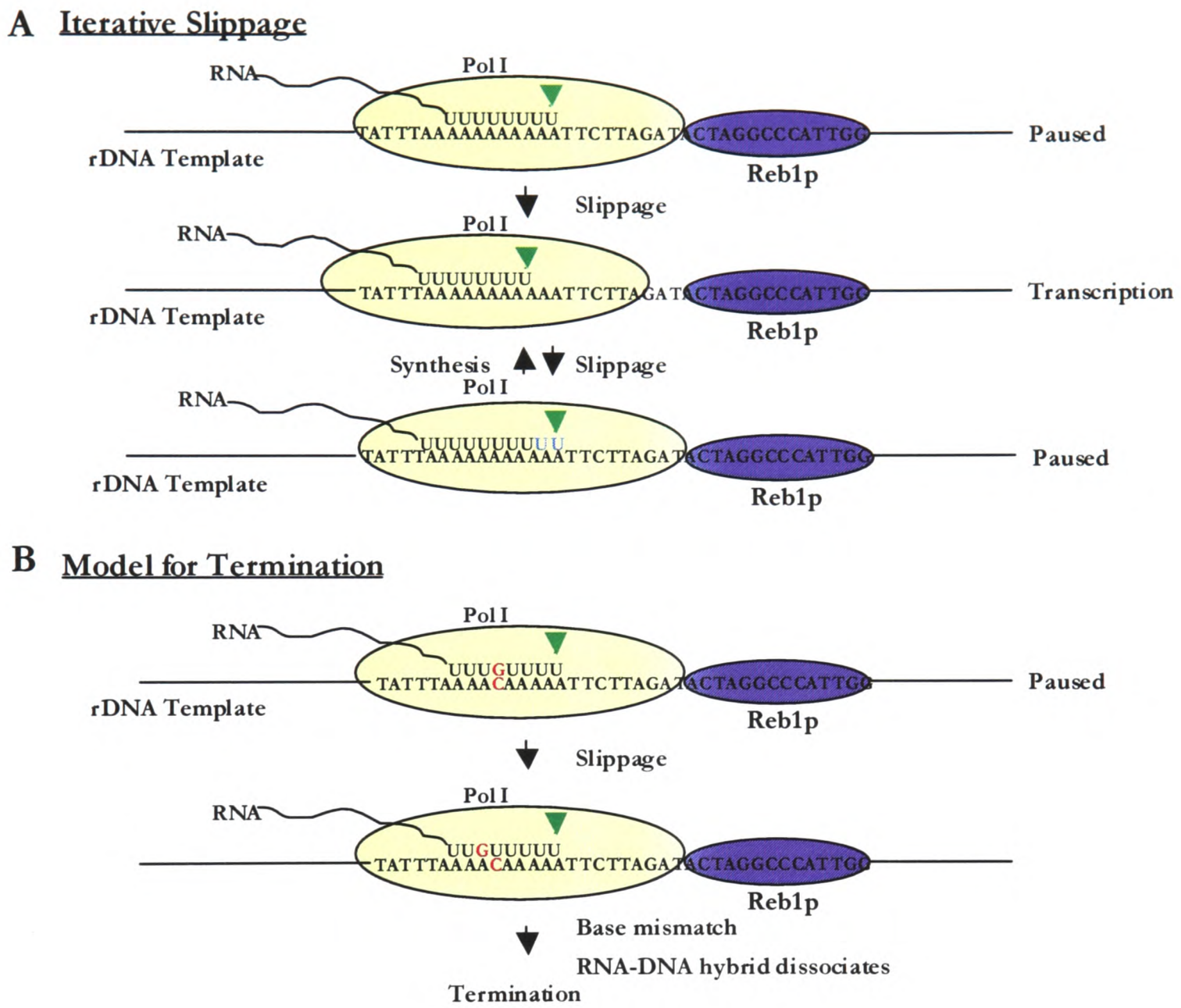


Figure D Model for transcription termination at the yeast Pol I terminator. (A) Iterative slippage. A Pol I ternary complex (beige) is paused by the Reb1p (purple). The wavy line represents the primary RNA transcript; the small vertical green arrows indicate the catalytic site of Pol I. The complex contains an 8 bp U/A RNA-DNA hybrid and is paused over the upstream element (top picture). Weak base pairing results in slippage, whereby the catalytic centre of Pol I slips 2 nt backwards together with the nascent RNA (middle picture). The hybrid reforms due to the homopolymeric sequence (bottom picture), allowing re-synthesis (blue nucleotides) until Reb1p is re-encountered. The cycle repeats itself and can result in the addition of thousands of U residues to 3' end of the nascent RNA. (B) Termination. Pol I paused after the encounter with Reb1p (top picture). The polymerase slips backwards but a stable RNA-DNA hybrid cannot form in the backtracked position due to the mismatch (red nucleotides in bottom picture). Re-synthesis cannot occur and the ternary complex falls apart with results in termination. Reproduced from Reeder and Lang 1997.

In vivo, Pol I has been observed to elongate at a rate of about 20 nt per second (Reeder and Lang 1997). Since a density of approximately 1 polymerase every 100 bp has been observed on electron micrographs of active rRNA genes this suggests that Pol I initiation and termination occur every 5 seconds (Reeder and Lang 1997). These *in vivo* kinetics are quicker than *in vitro* results where a paused polymerase has a half-life of 2 minutes before release in purified reactions using Reb1p and

up to 9 minutes before release using the lac repressor in place of Reb1p (Jeong *et al.* 1995; Lang and Reeder 1995; Reeder and Lang 1997). In whole cell extracts, the half-life of paused polymerases decreases to 30 seconds suggesting that the purified system is suboptimal or missing a component (Reeder and Lang 1997). Recently evidence suggests that an additional protein, Pol I Transcript Release Factor (PTRF), may increase the rate of transcript release by a paused polymerase through interacting with Pol I and the terminator protein (Reb1p or TTF-1) in the mouse system (Jansa and Grummt 1999). Work in yeast systems has also shown that a crude extract increases release suggesting the role of a PTRF functional homolog (Mason *et al.* 1997; Tschochne and Milkereit 1997).

In vitro analysis of the yeast rRNA genes identified several termination sites. The first and primary location of termination is +93 bp downstream of the end of the 25S rRNA coding sequence with additional termination sites mapped to about +250 bp and +350 bp with the former labelled the 'failsafe' terminator (Reeder *et al.* 1999). The primary termination site, located at 93 bp downstream of the 25S rRNA coding sequence, occurs just upstream of the Reb1p binding site. However, these *in vitro* data have not been adequately confirmed using *in vivo* analysis. Part II of this thesis resolves this omission by showing that Pol I transcription run-on analysis is a way of measuring *in vivo* Pol I termination.

Pol II

As a result of the complex co-transcriptional processing of Pol II transcripts, the mechanism for termination and coupled mRNA 3' end processing is less well understood than for other polymerases. However, it is known that following transcript cleavage at the poly(A) site, the upstream product is polyadenylated to become mRNA (Zhao *et al.* 1999). The downstream product, destabilised by cleavage, results in a short-lived transcript that cannot normally be visualised by standard RNA mapping techniques. Despite the complexity of Pol II termination, it appears to be primarily regulated by mRNA 3' end processing which is modulated by signals located upstream and downstream of the poly(A) site (Zhao *et al.* 1999).

Transcriptional termination is dependent on the presence of a functional poly(A) signal (Proudfoot 1989). This processing signal has been suggested to function in two ways: as an 'antiterminator' in which a conformational change occurs in the transcriptional complex upon recognition of the poly(A) signal creating a termination competent complex (Logan *et al.* 1987), or as in the 'torpedo' model where cleavage of the transcript at the poly(A) site initiates termination by promoting rapid degradation of the 3' end product while it is still attached to the elongating polymerase (Connelly and Manley 1988; Proudfoot 1989). These models are not mutually exclusive and it is likely that elements of both of are actually involved in Pol II termination.

In *S.cerevisiae* it has been shown that termination is entirely dependent of cleavage/polyadenylation factors. Properly regulated termination is prevented when these factors are mutated, highlighting the interdependence of these functions (Birse *et al.* 1998). Additional evidence suggests that RNA processing occurs some distance downstream of where the poly(A) site is transcribed. Nascent transcription has been shown to extend several hundred nucleotides beyond the site of cleavage and poly(A) addition (Proudfoot *et al.* 2002). This fact, taken with the CTD association of cleavage and poly(A) factors suggest a model for post poly(A) site polymerase in which the nascent RNA loops out from the CTD-associated poly(A)/cleavage factors and the Pol II RNA exit channel (Proudfoot *et al.* 2002).

Processing of the mRNA 3' end is not the only requirement for transcriptional termination. Termination specific factors and sequence elements located downstream of the poly(A) site also play a role (Proudfoot *et al.* 2002). In yeast, several proteins such as the Sub1p and Mbf1p are required for efficient transcriptional termination (Aranda and Proudfoot 2001; Calvo and Manley 2001). Detailed studies of nascent transcription in downstream flanking regions of genes have identified sequences that may play a vital role in termination. Additionally, transcriptional events can influence RNA processing. Downstream transcriptional pause sites are thought to increase 3' end processing by transiently pausing the polymerase so as to enhance poly(A) site recognition. This can be seen in *S.pombe* by the presence of a pause inducing downstream sequence element that is required for proper poly(A) site recognition and termination (Birse *et al.* 1997; Aranda and Proudfoot 1999).

Also, recent evidence in mammals demonstrates that specific sequences located downstream of the poly(A) site can direct cotranscriptional cleavage (CoTC), which may also play an important role in termination (Dye and Proudfoot 2001).

Both CoTC and 3' end processing are required for transcriptional termination in higher eukaryotes although they occur independently of each other (Dye and Proudfoot 2001). This has lead to a two-step model for Pol II termination in higher eukaryotes (Dye and Proudfoot 2001). In the first step, the nascent transcript is cleaved in a sequence dependent manner by a CoTC event while the transcript continues to interact with the CTD. This interaction and subsequent 3' end processing at the poly(A) site then plays a role in the second step of the termination process: polymerase release. While CoTC operates independently of mRNA 3' end processing, polymerase release is entirely dependent on it suggesting a scenario where mRNA 3' end processing effectively triggers polymerase release (for review see Dye and Proudfoot 2001).

1.4 rDNA repeat: Model Pol I system

Transcription of the rRNA genes in *S.cerevisiae* by Pol I takes place in the nucleolus yielding the large polycistronic 35S precursor RNA that is subsequently processed into the 18S, 5.8S and 25S mature rRNAs. As the sole prolific transcripts produced by Pol I, rRNA genes were a strong candidate for TRO analysis to study nascent transcription termination in *S.cerevisiae* Pol I.

Sequence specifics

Multiple copies of rDNA repeats are found in clusters arranged head to tail on chromosome XII. In *S.cerevisiae*, there are ~150 copies (variable by strain) of the rDNA 9.1 kb repeat each containing the 35S pre-rRNA (Goffeau *et al.* 1996)(See Figure E).

Saccharomyces cerevisiae rDNA repeat

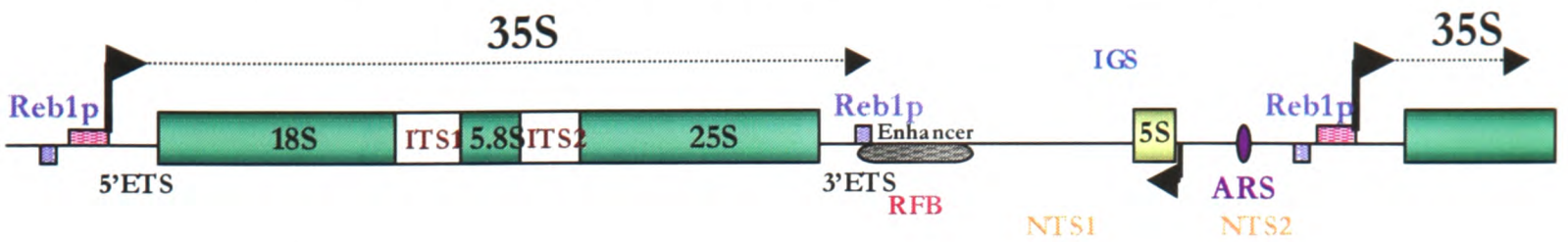


Figure E Diagram of rDNA repeat. Direction of the 35S polycistronic transcript is indicated by arrow and dotted line. The 18S, 5.8S and 25S genic regions are indicated in green. The intergenic sequence (IGS) is the sequence between the rDNA repeats and labelled in blue. The 5' and 3' external transcribed spacers (ETS) are indicated. The internal transcribed spacers (ITS1 and ITS2) are in brown. The non-transcribed spacers (NTS1 and NTS2) are the sequences between the transcription units and indicated in orange. Pol III transcribed 5S gene is identified by the yellow box with direction of transcription indicated by the arrow. The purple box identifies the location of binding of Reb1p. Promoter region of 35S transcript is identified by pink box. The purple oval identifies the location of the autonomously replicating sequence (ARS). The approximate region of the replication fork barrier (RFB) is indicated in red. The enhancer region is identified by grey oblong.

Each polycistronic transcript is comprised of two internal transcribed spacers sequences (ITS1 and ITS2) and flanked by two external transcribed spacer sequences (5'ETS and 3'ETS) (Kressler *et al.* 1999). A Pol III transcription unit for 5S is located in the intergenic spacer (IGS) region. Two non-transcribed spacers (NTS1 and NTS2) are located in the IGS. These regions contain the rRNA gene promoters, termination sites, an autonomously replicating sequence (ARS) element, a replication fork barrier (RFB) and a region that has all of the attributes of the single Pol I transcriptional enhancer (Banditt *et al.* 1999). At any given time, about half of the rDNA repeats contain the appropriate factors to stimulate transcriptional activation with the inactive fraction found to be packaged into nucleosomes (Conconi *et al.* 1989; Dammann *et al.* 1993). Transcriptionally active genes are randomly distributed along the yeast rRNA gene locus (Dammann *et al.* 1995).

Proximal to the rRNA gene promoter there is a terminator protein-binding site (Rep1p in yeast and TTF-1 in mice) that has been shown to augment transcription (Henderson *et al.* 1989; McStay and Reeder 1990). While the mechanism is unclear, it appears that in mice enhancement occurs irrespective of location and has been suggested to prevent promoter interference (Henderson *et al.* 1989). When bound by TTF-1p, the promoter proximal terminator has been shown to function in chromatin remodelling by repositioning a nucleosome over the core promoter (Langst *et al.* 1998). Additionally, the terminator protein Reb1p is capable of oligomerization *in vitro* suggesting that it may link the end of one repeat to the beginning of the next (Sander and Grummt 1997). In yeast, the Reb1p binding site, located 220 bp upstream of the promoter, is in the wrong orientation for

functional termination but deletion of the binding site results in a twofold decrease in the level of transcription *in vivo* (Kulkens *et al.* 1992). As there is functional homology between terminator proteins in eukaryotes, the yeast promoter proximal terminator may play a role in chromatin remodelling of the promoter.

While the rDNA gene repeats are usually located in tandem on the chromosome, it has been shown that yeast are viable when all the repeats exist as extra-chromosomal plasmids (Reeder 1999). In respiratory-deficient (petite) strains of *S.cerevisiae*, cryptic Pol II promoters in the IGS of episomal rDNA repeats generate functional rRNA transcripts suggesting a mechanism for selective Pol II activation of rRNA gene transcription (Conradwebb and Butow 1995). Recently the UAF complex was identified as being involved in the suppression of cryptic Pol II promoters located in the IGS of the rRNA genes (Oakes *et al.* 1999; Vu *et al.* 1999).

Regulation

While rRNA gene products are produced in all viable eukaryotic cells, regulation of rRNA gene synthesis occurs. For example, transcription of rRNA is highly active during the logarithmic growth phase and almost completely abolished in stationary phase (Reeder 1999). Additionally, there appears to be a set point on the number of rRNA genes that can be activated at any given time. Experiments with chicken rDNA indicated that an increase in the number of rDNA repeats did not result in an increase in the level of rRNA (Muscarella *et al.* 1987). Evidence also suggests that rRNA genes are either completely loaded with Pol I or transcriptionally silent (Reeder 1999).

Independent regulation of transcription of rRNA genes occurs by influencing Pol I or PIC activity (Reeder 1999). Short-term down regulation of rRNA synthesis requires decreasing the activity of Pol I, allowing for the possibility of rapid up-regulation if the situation improves. Some evidence suggests Pol I activity can be regulated through the association of the Pol I with Rrn3p, which has been shown to distinguish transcriptionally active polymerase complexes (Milkereit and Tschochner 1998). Indeed, formation and disruption of the Pol I-Rrn3p complex, affected by phosphorylation of Pol I, might constitute a molecular switch for rRNA synthesis (Fath *et al.* 2001). Alternately,

regulation at the level of PIC formation is likely to occur for long-term changes in growth rate or nutrition (Reeder 1999). Transcriptional regulation at the level of PIC formation appears to be mediated by the enhancer. An 'open' enhancer structure (the absence of nucleosomes), is found in spacers where the adjacent gene is activated and also in strains of yeast where Pol I is inactivated (Lucchini and Sogo 1998; Reeder 1999). This is consistent with the formation of the PIC as the limiting step, but does not eliminate regulation at the level of Pol I function (Reeder 1999).

1.5 Galactose metabolism genes: Model system for Pol II

The regulatory circuit that controls the metabolism of galactose via the *GAL* genes is among the most intensively studied and best understood in yeast. These highly regulated and quickly inducible genes can be used to study transcriptional changes in the TRO assay. Previously, interference has been demonstrated in the Pol II transcribed *GAL* genes located in a cluster on chromosome II (Greger and Proudfoot 1998). This well understood system was employed to investigate the implications of convergent transcription in Pol II transcribed genes. Subsequently, a brief account of the regulation involved in transcription of the *GAL* gene cluster is provided (discussed further in Part I).

Genes involved in the galactose metabolism

In *S.cerevisiae,* the metabolism of galactose involves a highly regulated set of genes induced by galactose and repressed by glucose. The structural *GAL* genes (*GAL1*, *GAL2*, *GAL7*, *GAL10* and *MEL1*) are controlled by the regulatory *GAL* genes (*GAL3*, *GAL4* and *GAL80*) at the level of transcription. Gal4p, a transcription activator that binds to the promoters of all *GAL* genes and is activated in the presence of galactose, is a key player in this process (Johnston 1987; Johnston and Carlson 1992; Lohr *et al.* 1995; Greger and Proudfoot 1998). The *GAL* operon is conserved across species. The build-up of the intermediate galactose-1-phosphate, resulting from the lack of function of the Gal7p homolog in each species, has proven toxic in all organisms studied so far (Johnston 1987).

GAL gene repression

Glucose, the preferred carbon source of *S.cerevisiae*, tightly represses the galactose metabolic gene pathway (Johnston and Carlson 1992). Additionally, glucose inhibits numerous other genes such as those involved in the metabolism of other sugars, mitochondrial respiratory chain proteins and enzymes involved in gluconeogenesis. However, the repression in the *GAL* genes is exceptionally high at 1000-fold compared to the 15-fold repression seen in *MAL* genes, involved in the metabolism of maltose. The *GAL* genes are also inactive, but not tightly repressed, in other carbon sources such as glycerol, ethanol or raffinose. While the activator is bound to the UAS and poised for induction, Gal80p binds to Gal4p and inhibits transcriptional activation.

Conversion of galactose to glucose-6-phosphate

Galactose, through a four-step pathway first described by Leloir (Leloir 1951), is converted into glucose-6-phosphate that can then be funnelled into the glycolytic pathway. After galactose is taken into the cell through the *GAL2* encoded permease, the structural gene product of the *GAL1* gene (galactokinase) converts the substrate into galactose-1-phosphate. It is this intermediate that is toxic if allowed to accumulate in the cell. This product is then uridylated by the Gal7p, galactose-1-phosphate uridyl transferase into UDP-galactose with the uridyl group derived from uridine di-phosphate glucose. The third step involves the inversion of the –OH group at C-4 by Gal10p (UDP-glactose-4-epimerase) resulting in glucose-1-phosphate. Gal5p, a constitutive expressed phosphoglucomutase, carries out the last step by isomerisation to glucose-6-phosphate that can freely enter the glycolytic pathway (see review in Johnston 1987).

1.6 Aims of this study

These studies set out to investigate nascent transcriptional termination in Pol I and Pol II transcribed genes. Nascent transcriptional run-on (TRO) analysis of the galactose metabolising genes, as the model system for Pol II, allowed an investigation into the effects of convergent transcription. Manipulation of the wild type orientation of the *GAL10* and *GAL7* genes was carried out to model the effect of convergent, co-transcribed genes. Eliminating the intergenic region of the convergently arranged genes demonstrated the effect of improper termination on convergent transcription. This

result graphically illustrates the need for efficient Pol II termination. In its absence inhibition of both genes occurs by a mechanism referred to as transcriptional collision. In Part II, the site of termination for Pol I transcription of rRNA genes was identified. Deletion of factors involved in the expression of Pol I genes (transcriptional machinery and processing of the pre-rRNA transcript) demonstrate a key role played by specific factors in proper termination.

Materials

Reagents

Most chemicals and reagents were purchased from BDH and were of 'AnalaR' grade. Chemical reagents obtained from other sources are listed below:

Acrylamide (40%, acrylamide:bis 19:1)	Camlab
Agar n°1	Oxoid
Agar, bacterial	Difco
Agarose	Flowgen
Ampicillin	Sigma
β-mercaptoethanol	Sigma
Boric acid	BDH
Bovine Serum Albumin (BSA)	Sigma
Bromophenol Blue	Sigma
Chloroform	Sigma
Deoxyribonucleotide triphosphates (dNTPs)	Roche
Dimethyl sulphate, 99+%	Roche
Dimethyl sulphoxide	Aldrich
Dithiothreitol (DTT)	Roche
Ethanol	Hayman Ltd
Ethidium Bromide	Sigma
Formaldehyde	Sigma
Formamide	Sigma
Galactose, 99% (<0.01% glucose)	Sigma
HEPES	Life Technologies Inc.
Herring sperm DNA	Sigma
Hybond NX	Amersham-Pharmacia
IPTG	Sigma
Orange G	Sigma
PEG 4000, 8000	Sigma
Peptone, bacterial	Oxoid
Phenol	Sigma
Polyvinylpyrrolidone	Sigma
Raffinose	Sigma
Ribonucleoside triphosphates	Roche
Sucrose	Rose Chemicals Ltd

TEMED	Sigma
Tetracycline	Sigma
Tris	ICN Biochemicals
Xylene Cyanol	Sigma
Yeast Extract	Oxoid
Yeast Nitrogen Base (+ammonium sulphate, w/o amino acids)	Difco

All essential amino acids contained in SC dropout media were obtained from Sigma

Enzymes

Restriction endonucleases were obtained from Amersham-Pharmacia, Roche, Life Technologies Inc. and New England Biolabs. Other enzymes were obtained from the following suppliers:

Calf Intestinal Phosphatase	Roche
DNase I, RNase Free	Roche
DNA Polymerase I, (Klenow)	New England Biolabs
Lysing Enzymes	Sigma
Pfu DNA Polymerase	Stratagene
RNase A	Roche
RNasin	Promega
SP6 RNA Polymerase	Roche
SuperScript II	Gibco-BRL
T3 RNA Polymerase	Roche
T4 DNA Ligase	New England Biolabs
T4 DNA Polymerase	New England Biolabs
T4 DNA Polynucleotide Kinase	New England Biolabs
T7 RNA Polymerase	Promega
T7 Sequenase	Amersham-Pharmacia
Taq DNA Polymerase	Promega

Radioisotopes

[α^{32}P] dATP	3000 Ci/mmol; 10 mCi/ml
[α^{32}P] rUTP	3000 Ci/mmol; 10 mCi/ml
[α^{35}S] dATP	1000 Ci/mmol; 10 mCi/ml

All radioisotopes were obtained from Amersham-Pharmacia.

Kits

Sequenase version II Kit (Amersham-Pharmacia/United States Biochemicals)

Qiagen Plasmid Purification Kit (Qiagen, Hilden, Germany)

Qiaquick Gel Purification Kit (Qiagen, Hilden, Germany)

Organic Reagents

Phenol for DNA work was dissolved in TE at 65°C, equilibrated twice with fresh TE, and stored at 4°C. Phenol for RNA work was equilibrated with water and stored at 4 °C. Phenol:chloroform was prepared in the ratio of 1:1.

Molecular Weight Markers

Agarose gel markers: λ DNA digested with *Hind* III mixed with pβE [Proudfoot, 1992] digested with *Eco*R I + *Hind* III, *Bam*H I + *Sst* I, and *Hinf* I. Radioactive markers for acrylamide gels were an *Eco*R I + *Hinf* I digest of pβE Klenow labelled with [α^{32}P] dATP.

Media

E.coli

LB (Luria-Bertani) (per litre)

Bactotryptone	10 g	1% w/v
Sodium chloride	10 g	85 mM
Yeast extract	5 g	1% w/v

2xTY (per litre)

Tryptone T	20 g	1.6% w/v
Sodium chloride	10 g	85 mM
Yeast extract	5 g	1% w/v

LB-agar: 15 g/litre of LB

Ampicillin plates

20 g agar technical (Difco) was dissolved in 1 litre of deionised H_2O, autoclaved and ampicillin was added to final concentration of 50 µg/ml.

Tetracycline plates

10 mM $MgSO_4$ and 20 µg/ml tetracycline were added to autoclaved LB-agar.

S.cerevisiae

YP-medium (per litre)

Bacterial peptone	20 g
Yeast Extract	10 g
Carbon Source (Galactose, Glucose, Raffinose) was added to	2%

SC-medium (per litre)

Yeast nitrogen base, w/amino acids	6.7 g
Dropout mix (0.2%)	1.5 g
Carbon source (Galactose, Glucose, Raffinose) was added to	2%

YP-plates (per litre)

Bacterial peptone	20 g
Yeast Extract	10 g
Agar	20 g
Carbon Source (Galactose, Glucose, Raffinose) was added to	2%

SC-plates (per litre)

Yeast nitrogen base, w/amino acids	6.7 g
Dropout mix (0.2%)	1.5 g
Agar	20 g
Carbon source (Galactose, Glucose, Raffinose) was added to	2%

Drop-out Mix

Adenine	0.5 g
Arginine	2.0 g

Asparagine	2.0 g
Histidine	2.0 g
Isoleucine	2.0 g
Leucine	10.0 g
Lysine	2.0 g
Methionine	2.0 g
Phenylalanine	2.0 g
Threonine	2.0 g
Tryptophan	2.0 g
Uracil	2.0 g
Valine	2.0 g

The amino acids listed above (minus the appropriate 'drop-out') were mixed thoroughly by rotating at 4°C for a minimum of 4 hours.

Solutions and Buffers

Gel Electrophoresis Buffers

50x TAE-Buffer (agarose gels)

Sodium acetate	250 mM
EDTA	50 mM
Tris	2 M
Glacial acetic acid	91 ml/litre

10x Orange G dye

Ficoll	5% w/v
Orange G	2.5 mg/ml
EDTA	10 mM

5x TBE (polyacrylamide gels)

Tris	50 mM
EDTA	10 mM
Boric acid	50 mM

1x Formamide loading dye (for urea and formaldehyde gels)

Bromophenol blue	10 mg/ml
Xylene cyanol	10 mg/ml
EDTA	10 mM
Formamide	80% v/v

5x SDS loading dye (for native gels)

Bromophenol blue	10 mg/ml
Xylene cyanol	10 mg/ml
Ficoll	10 mg/ml
SDS	1%

Solutions for DNA analysis

Mini-preparation of plasmid DNA

STET

Tris-HCl, pH 8.0	10 mM
EDTA	50 mM
Sucrose	10%
TritonX-100	0.5%

TE

Tris-HCl, pH 8.0	10 mM
EDTA	1 mM

DNA-preparation Buffer for yeast cells

Sorbitol solution

Sorbitol	900 mM
Tris-HCl, pH 8.0	100 mM
EDTA	100 mM

Breaking Buffer

TritonX-100	2%
SDS	1%
NaCl	100 mM
Tris-HCl, pH 8.0	10 mM
EDTA	1 mM

Solutions for RNA analysis

AE Buffer

NaOAc	50mM
EDTA	10mM

Adjust to pH 5 with acetic acid

10x MOPS

MOPS	200 mM
NaOAc	50 mM
EDTA	10 mM

Adjust to pH 7 with NaOH

TES

Tris-HCl, pH 7.5	10 mM
EDTA	10 mM
SDS	0.5%

MMF

Formamide	50%
Formaldehyde	16.8%
10X MOPS	1X

TRO 2.5x Transcription Buffer

Tris-HCl, pH 7.4	50 mM
$MgCl_2$	25 mM
NaCl	100 mM

RNA Elution Buffer

NH_4OAc	0.5 mM
SDS	0.1%
EDTA	1 mM

Mix for Random Prime Labelling

DTM:

Tris-HCl, pH 8.0	250 mM
$MgCl_2$	25 mM
β-mercaptoethanol	50 mM
dNTPs minus dATP	100 µM

OL:

Random hexamers (Amersham-Pharmacia) dissolved in TE to final concentration of 90 OD units/ml.

OLB-A (final mix):

1M HEPES, pH 7.6, DTM and OL were mixed 25:25:7.

Blotting Solutions

50x Denhardt's

Ficoll	1%
BSA	1%
Polyvinylpyrrolidone	1%

20x SSPE

NaCl	3 M
NaH_2PO_4	200 mM
EDTA	20 mM

Adjusted to pH 7.4 with NaOH

20x SSC

NaCl	3 M
Sodium citrate	300 mM
Adjusted to pH 7 with HCL	

Denaturing Solution

NaCl	1.5 M
NaOH	0.5 M

Neutralizing Solution

NaCl	1.5 M
Tris-HCl, pH 7.2	0.5 M
EDTA	1 mM

Hybridisation Solution

SSPE	6x
Formamide	50%
Denhardt's Solution	5x
Herring Sperm DNA	50 µg/ml
SDS	0.1%

Bacterial Strains

NM554: F-*ara*D139, Δ(*ara-leu*)7696, *galE*15, *galK*16, Δ(*lac*)X74, *rpsL*(Str^r), *hsd*R2, (r_k^- m_k^+), *mcr*A, *mcr*B1, *rec*A13

DH5α: F' *endA*1, *hsd*R17 ($r_k^-m_k^+$), *supE*44, *thi*-1, *recA*1, *gyrA*96, *relA*1, Δ(*lacIZYA-argF)*U169, *deo*R, Φ80d*lac*Δ(*lacZ*)M15

XL1-Blue: F'::*Tn*10 (*tet^r*), *proAB*, *lacI^q*, Δ(*lacZ)*M15, *recA*1, *endA*1, *gyrA*96, *thi*-1 *hsd*R17, ($r_k^-m_k^+$), *supE*44, *recA*1 *lac*

Yeast Strains

BY4732: MATα; *his3Δ200; met15Δ0; trp1Δ63; ura3Δ0*

BY4733 or Δ*GAL*	BY4732 derivative with the *GAL10-7*::*URA3*
MWK028	MATa; *trp*1-1; ura2-52; can1-100; leu-3,112; his3-11,15; ade2-1; met2-Δ1; lys2-Δ2; LEU2 ::snr6, YCp50 (URA3, CEN4, ARS1, SNR6) reference: (Kaiser and Brow 1995). Supplied by R. Reeder.
RLY10	MWK028 with *HIS*3::rpa12. Supplied by R. Reeder.
RLY11	RLY10 with pRS424 (*TRP*1, 2micron, *RPA*190). Supplied by R. Reeder.
*RTN*1	MATa; *ura*3-1; Δ*trp*1; *ade*2-1; *leu*2-3, 112; *his*3-11, 15. Supplied by D. Tollervey.
rnt1	*RNT*1 strain above with *RNT*1:: *TRP*1. Supplied by D. Tollervey.

Plasmid Constructs

All plasmid constructs were verified by sequence analysis.

pYC10-7 (WT) was based on Ycplac22 and is described previously (Greger and Proudfoot 1998). A 4814 bp *GAL10-7* fragment was inserted into the *Bam*HI site of Ycplac22.

pYC10-7Con (Con) was created by complete digestion of pYC10-7 with *Xba1* followed by partial digestion with *EcoRV*. The 7658bp band was isolated and filled in with Klenow. A *GAL7* insert fragment was generated by PCR with blunt ends and a unique *XhoI* site 103 bp after the poly(A) site (see primers).

pYC10-7Fus (Fus) was created by ligating long range PCR products generated from Con using *Pfu* DNA polymerase (Stratagene) and kinased primers Fuse7 and Fuse10 located directly inside the open reading frames for both genes (see primers). The 209bp *GAL7* poly(A) site fragment (*GAL7*:

1118 to 1327) in the forward orientation for *GAL7* was inserted into the *SacI* site of the vector backbone upstream of the *GAL10* promoter (-442).

pYC10-7FusΔTATA*GAL7* (Fus-Δ7P) had a 9 bp deletion of the TATA box 5'-ATATAAAA-3'(*GAL7*: -179 to -188) in the Fus construct generated by long range PCR with primer ΔTATAa and ΔTATAb (see primers).

pYC10-7FusΔUAS*GAL10* (Fus-Δ10P) had the *EcoRI* fragment of Fus replaced with the *EcoRI* fragment from pΔ-UAS_{G10} plasmid described previously (Greger and Proudfoot 1998). This created the Fus with a deletion that eliminates –136 to –414 of *GAL10*, removing all UAS sites.

pYC10-7FusΔTATA*GAL7*/ΔUAS*GAL10* (Fus-Δ7P/Δ10P) generated by replacing the *EcoRI* fragment from Fus-Δ7P plasmid with the *EcoRI* fragment from Fus-Δ10P.

pYC10-7FusΔTATA*GAL7*&ΔUAS*GAL10*(Fus-Δ7P&Δ10P) was a sequential transformation where the yeast carrying Fus-Δ7P was transformed with Fus-Δ10P and selected for on SC- plates (lacking uracil and tryptophan) with 2% galactose as the carbon source.

M13 Probe Constructs

Single stranded DNA probes were made (see Table B) by cloning PCR generated fragments (using *Pfu* DNA polymerase; Stratagene) from Con into the double digested *Xba*I and *Eco*RI sites of M13mp18 (RF) and M13mp19 (RF). Primers used to generate fragments had either a *Xba*I or *Eco*RI restriction site so that there were site specific sticky ends. This gave probes for both strands: *GAL7* reading frame and *GAL10* reading frame. The M13 control probe (m) contained no insert. The positive control actin probe 'a' had 567 bp fragment containing 277-844 bp 3' of the *ACT1* start codon inserted into the *Hinc*II site of M13mp18 (RF).

Ribosomal M13 probes were made in the same manner, cloning PCR (using *Pfu* DNA Polymerase) fragments generated from chromosomal DNA from *S.cerevisiae* strain BY4733 6-1/13 into M13mp18 (RF). A negative control was an M13 probe with a *ura4* inset from *S. pombe* as listed.

Table B TRO M13 probes

Probes	Location (size in bp)	U Content
	S.cerevisiae	
A	*GAL7*: -401/-212 (189)	51
A		74
B	*GAL7*: 73/252 (179)	58
B		47
C	*GAL7*: 638/793 (155)	45
C		47
D	*GAL7*: 800/944 (144)	45
D		44
E	*GAL7*: 938/1119 (181)	55
E		60
F	*GAL7*: 1110/1294 (184)	62
F		77
G	*GAL7*: 1184/1354 (170)	67
G		61
H	*GAL7*: 1266/1362 (96)	63
	GAL10: 2406/2325 (81) [177]	
H		54
I	*GAL10:* 2245/2055 (190)	72
I		56
J	*GAL10:* 2070/1884 (186)	58
J		60
K	*GAL10:* 1885/1727 (158)	46
K		51
L	*GAL10:* 1720/1572 (148)	38
L		56
M	*GAL10:* 1537/1379 (158)	46
M		51
N	*GAL10:* 330/152 (178)	55
N		53
O	*GAL10:* -187/-371 (184)	46
O		43

Pro	rDNA: -85/-310 (225)	61
18S	rDNA: 2070/2284 (214)	54
25S	rDNA: 5953/6152 (199)	62
r2	rDNA: 6592/6730 (138)	29
r3	rDNA: 6701/6847 (146)	47
r4	rDNA: 6804/6972 (168)	40
r5	rDNA: 6902/7075 (173)	49
r6	rDNA: 6948/7128 (180)	57
r7	rDNA: 6999/7198 (199)	66
r8	rDNA: 7052/7255 (203)	70
r9	rDNA: 7223/7390 (167)	56
	S.pombe	
neg	*ura4*: 1227/1612 (385)	114

Primers

CC *GAL7* Insert

GAL7+100	CTCGAGACTTACAAGCTGCATTGTATTC
Up*GAL7*	GGAGTTCATTTCGTTACTTTTG

FF Construct

Fuse7	CCACTTTCTTTTTACAGTCTTTGTAGATAATG
Fuse10	GCTGGCAAATCAGGAAAATCTGTAGAC

Δ7P Construct

Δ7TATAa	TTTGCTAGCCAAACTAACTGAACATAG
Δ7TATAb	GCAGGTCGGAAATATTTATGGGCATTA

Northern Probes

The northern probe was a 1.5 kb fragment generated by double restriction digest with *Bgl* II and *Sal* I of the WT *GAL10-7* locus which contained 427 bp of 5' *GAL7* and 403 bp of 3' of *GAL10*. The *ACT1* probe was a 567 bp fragment (277 –844 bp 3' of the *ACT1* start codon).

Methods

Phenol extraction/ethanol precipitation

An equal volume of phenol (~pH 7.9 for DNA, acidic phenol for RNA) was added to solution to be extracted, mixed and microfuged. The aqueous phase was removed and precipitated with 0.3M NaAc and 100% ethanol (2.5 volumes) for 30 minutes at 4°C. The sample was pelletted for 15 minutes, ethanol removed, washed with 70% ethanol and spun again. The resulting pellet was resuspended in an appropriate amount of dH_2O.

Restriction enzyme digests

Plasmid DNA (1-10 µg) from mini- or maxi-prep was routinely digested in a 25 µl mix [1.5 µl restriction enzyme (10 U/µl), 2.5 µl 10x restriction enzyme buffer, 18 µl dH_2O]. Restriction enzyme buffers (New England Biolabs) were optimal buffers for the restriction enzyme in a single digest or the best compromise buffer in the case of multiple enzyme digests. Incubation at the appropriate temperature was carried out for 1-3 hours. Digestion products were analysed by agarose gel electrophoresis.

Phosphatase treatment

Restriction enzyme digests were directly supplemented with 1 µl Calf Intestinal Phosphatase (1 U/µl). The concentration of buffer was adjusted to a final volume of 30 µl (with 10x AP-buffer). The reaction was incubated at 37°C for 1 hour.

Flushing ends

3' overhang

3' overhang ends were made blunt with T4 DNA Polymerase. Crude PCR reactions or DNA digests were phenol extracted and ethanol precipitated. The pellet was resuspended in 20 µl dH_2O and supplemented with 2 µl T4 DNA Polymerase (3 U/µl), 3 µl 10x T4 DNA Polymerase Buffer and 5 µl dNTP (10 mM). Reaction was carried out at 12°C for 20 minutes.

5' overhang

5' overhang ends were made blunt with Klenow DNA Polymerase. DNA digests were phenol extracted and ethanol precipitated, and resuspended in 20 µl dH_2O were supplemented with 2 µl Klenow DNA Polymerase (5 U/µl), 3 µl 10x Klenow DNA Polymerase Buffer and 5 µl dNTP (10 mM). Reaction was carried out at room temperature for 15 minutes.

Polymerase Chain Reactions

PCR reactions using *Taq* DNA Polymerase

PCR reactions were carried out in a final volume of 50 µl containing ~ 1 ng DNA template, 50 pmol each primer, 0.8 mM dNTP, 1x Mg Free Thermophilic DNA Polymerase Buffer (supplied by Promega), 2 mM $MgCl_2$, 1.5 U *Taq* DNA Polymerase and the remainder of the volume with dH_2O. The PCR enhancer Betain was also sometimes used at the final concentration of 1 M.

PCR reactions using *Pfu* DNA Polymerase

Reactions were carried out in final volume of 50 µl containing 1 µl DNA template (approximately 10pm), 5 µl of each primer (10 pM), 4 µl dNTP (10 mM), 5 µl 10x Cloned *Pfu* Buffer (supplied by Stratagene), 1 µl Cloned *Pfu* DNA Polymerase (2.5 U/µl), 29 µl of dH_2O. If optimising, would sometimes use 10 µl Betain (5M) in place of 10 µl of dH_2O. For PCR products greater than 3 kb Turbo *Pfu* DNA Polymerase was used.

Cycling

PCRs were cycled in a Biometra Trio thermocycler using either a heated lid unit or a 100 µl over layer of mineral oil. The initial denaturing step was carried out at 93°C for 2-3 minutes. The following cycle was repeated for 25-35 cycles: 93°C for 1 minute, 'annealing temperature' for 50 seconds and 72°C for extension time. Extension time was determined by allowing 30 seconds for each 500 bp to be amplified. The approximate annealing temperature was calculated using the formula (4°C per G or C, 2°C per A or T, minus 5°C). A final extension time of 3 – 5 minutes was added after the final cycle to allow complete extension. This protocol was optimised for different primer and template combinations.

Colony screening

For screening *E.coli* colonies, a single colony was picked with sterile tip, added to *Taq* PCR mix and boiled at 95°C for 4 minutes before proceeding with the cycling as usual.

Plaque screening

For screening M13 plaques prior to preparing DNA, 0.2 µl of the M13 phage stock was used in place of DNA template.

Phosphorylation of primers

For cloning of PCR products, oligonucleotides were kinased prior to performing the reaction. Oligos were phosphorylated in 20 µl reaction-mix [2 µl oligo (100 pmol/µl), 2 µl rATP (10mM), 2 µl 10x T4 PNK Buffer, 2 µl T4 Polynucleotide Kinase (10 U/µl), 12 µl dH_2O]. The reaction was incubated at 37°C for 1 hour then used directly in the PCR reaction.

Ligation

A typical ligation reaction was performed in a final volume of 20 µl containing 50 ng linearised vector, at least 4 fold Molar excess of insert, 2 µl 10x T4 DNA Ligase Buffer and 14 µl of dH_2O. Ligations were performed at 16°C for between 4 to 12 hours.

Plasmid preparation

Mini-preps

Plasmids were isolated from transformants using the boiled STET-lysis method (Sambrook *et al.* 1989). In brief, 1.5 ml of an overnight culture [grown in 2.5 ml of 2xTY, with 5 µl ampicillin (50 mg/ml)] was harvested, the cells resuspended in 200 µl of STET solution and boiled for 90 seconds at 95°C. Lysed cells were centrifuged for 10 minutes, the pellet removed, isopropanol was added to the supernatant and spun for 15 minutes. The plasmid DNA pellet was then resuspended in 50 –100 µl of dH_2O or TE.

Maxi-preps

Large-scale preparations of plasmid DNA were prepared using Qiagen Maxi Prep Kits employing the alkaline lysis method (Sambrook *et al.* 1989).

Electrophoresis

Agarose gel electrophoresis

Samples (in 1x Orange G) were run on 1-3% agarose/1x TAE Buffer (gels containing 1 µg/ml ethidium bromide) with current limiting at 200 mA (minigels) or 250 mA (midigels). Gels were visualized on UV transilluminator (Stratagene model 2040EV), and photographed using a UVP digital camera system. When a fragment was to be isolated for cloning, the band was visualized under 254nm UV light from a UVP Mineralight UVS 54 hand lamp and excised from the gel using a clean razor blade. DNA was isolated using Qiaquick Gel Extraction Kit (Qiagen).

Polyacrylamide gel electrophoresis

50 ml native and denaturing (7 M urea) gels were prepared in 1x TBE using 40% acrylamide:bis-acrylamide (19:1) to approximately 6.5%. Polymerisation was initiated with the addition of 50 µl TEMED and 500 µl 10% ammonium persulphate and the gel was cast between two 20 cm x 38.5 cm glass plates. Gels were run at 30 Watts with 1x TBE. Gels were washed with dH_2O to remove urea, transferred to Whatman 3MM paper, covered with Saran Wrap and dried for 30 minutes in BioRad model 583 gel dryer at 80°C. Dried gels were given a light dusting of talcum powder before being exposed to film (Biomax MR film) at room temperature or –80°C if magnification screen was used.

Isolation of fragments from denaturing polyacrylamide gels

After separating the plates, the gel was covered in Saran Wrap and exposed to film. Marker pen was used to mark the edges of the gel following exposure. Bands were excised from the gel using a fresh razor blade. RNA was extracted by incubation in RNA Elution Buffer for 2-4 hours at 37°C with shaking. After extraction, the polyacrylamide was removed and the RNA precipitated by the addition of three volumes of cold 100% ethanol.

Sequencing

Recombinant DNA clones were sequenced using the automated sequencing service. Manual sequencing was occasionally performed using T7 Sequenase™ Version 2.0 dGTP reagent Kit as described below:

Plasmid (approximately 1 µg) was denatured with 4 µl of Denaturing Mix (2M NaOH, 1 mM EDTA) with 1 µl 10mM primer to a final volume of 20 µl with dH_2O. This was incubated at 37°C for 30 minutes. To this 4 µl 2M ammonium acetate and 55 µl 100% ethanol (at 4°C) were added, spun down and air-dried for 10 minutes. The pellet was re-dissolved in 2 µl 5x Sequenase Reaction Buffer and 8 µl dH_2O. Sequence mix [1 µl 0.1 M DTT, 0.4 µl 5x Labelling Mix, 0.25 µl Sequenase Enzyme, 0.5 µl [^{35}S] dATP, 1.75 µl Enzyme Dilution Buffer and 1.6 µl dH_2O] was added to denatured DNA for 2-5 minutes at room temperature, 3.5 µl of this reaction was added to 2.5 µl termination mix of each of four nucleotides at 37°C for 5 minutes. Formamide loading dye was added (4 µl), sample was denatured at 95°C for 5 minutes and 6 µl was loaded onto a 6% denaturing polyacrylamide gel. This was run at 30 Watts for 1-2 hours.

Transformation and Growth of Bacteria

E.coli strains

Transformation

Competent *E.coli* cells were prepared by Joan Monks using the $RbCl_2$ method (Sambrook *et al.* 1989) and stored at -80°C. For transformations, cells were thawed on ice, 10 µl of ligation (or 0.1 µg of DNA if transforming a plasmid) was added to 100 µl of thawed competent cells on ice for 20 minutes. Heat shock was carried out at 42°C for 2 minutes followed by the addition of 200 µl of 2xTY media and incubation at 37°C for 30 minutes. Cells were then collected by brief centrifugation and 100 µl was plated onto the appropriate selective agar plate.

Growth

Cell cultures were grown in 2xTY with appropriate antibiotics at 37°C with shaking.

M13 phage

Transformation

Transformation was performed as above using XL1-Blue competent cells and an incubation period of 15 minutes following the addition of 200 µl of 2X TY. This mixture was added to 100 µl of an 8-hour culture of XL1-Blue plating bacteria, mixed with 3 ml of pre-melted top agar containing X-Gal (133 µg/ml) and IPTG (266 µg/ml) and poured onto pre warmed nutrient agar plates. Plates were incubated at 37°C overnight, then stored at 4°C to intensify blue colour of wild type plaques.

Growth

Mini preps

M13 plaques were picked with sterile yellow tip, added to 2.5 ml 2X TY containing 50 µl of an 8-hour culture of XL1-Blue plating bacteria. This was grown for 5-6 hours at 37°C with shaking. Cultures were harvested, collecting the supernatant as phage stock and the pellet for RF M13 DNA.

Maxi preps

2 ml of XL1-Blue culture (8 hour growth) was added to 50 – 200 µl of M13 phage stock for 5 to 10 minutes at room temperature. This was added to 200 ml of 2X TY and shaken for 6 hours at 37°C. The culture was harvested at by centrifuging for 15 minutes at 10,000 RPM in a Beckman J2-21 centrifuge. To the supernatant, 50 ml of 20% PEG/2.5M NaCl was added and chilled at 4°C overnight. This was spun down at 13,000 RPM for 30 minutes, the supernatant removed and the pellet resuspended in 750 µl of TE Buffer. The sample was spun at maximum speed for 1 minute, supernatant removed and extracted once with phenol, twice with phenol/chloroform and once with chloroform. The resulting aqueous layer was precipitated with 100% ethanol overnight. After spinning in microfuged at maximum speed for 20 minutes the pellet was rinsed with 70% ethanol, dried briefly and resuspended in 100 µl dH_2O. ssDNA concentration and integrity was determined spectrophotometrically and visually on a native agarose gel.

S.cerevisiae Strains

Transformation

Transformation of plasmids into *S.cerevisiae* 6-1/13 *(ΔGAL)* was according to (Gietz and Schiestl 1995). Pre-cultures were grown in 10 ml volume overnight at 30°C, expanded to 50 ml and grown for 3 to 5 hours or approximately OD_{600}= 0.5. The culture was harvested in a tabletop centrifuge for 5 minutes at 3000 xg, rinsed with 25 ml dH_2O and resuspended in 1 ml 100 mM LiAc. The cells were pelletted at top speed for 15 seconds. The pellet was resuspended in 400 µl 100 mM LiAc, vortexed, separated into 50 µl aliquots and the supernatant was removed. To this the "transformation mix" was added in the following order: 240 µl PEG (50% w/v), 36 µl 1.0 M LiAc, 25 µl herring sperm DNA (2.0 mg/ml), 50 µl of dH_2O and plasmid DNA (0.1-10 µg). Each tube was vigorously vortexed until cell pellet was completely mixed, incubated for 30 minutes at 30°C and heat shocked at 42°C for 20 minutes. The cells were then microfuged for 15 seconds at 6000 rpm, the transformation mix was removed and resuspended in 1 ml dH_2O. From this 200 µl was plated onto SC- plates lacking uracil and tryptophan and with 2% glucose. Incubation occurred at 30°C for 2-5 days.

To achieve **pYC10-7FusΔTATA*GAL7*&ΔUAS*GAL10* (Fus-Δ7P&Δ10P),** sequential transformation into the *S.cerevisiae* 6-1/13 *(ΔGAL)* was performed. The *S.cerevisiae ΔGAL* carrying Fus-Δ7P was transformed with Fus-Δ10P and selected for on SC- plates lacking uracil and tryptophan and with 2% galactose as the carbon source. In addition, 20 µg/ml ethidium bromide was added to the plates to impede respiration.

Growth

S.cerevisiae strain 6-1/13 *(ΔGAL)* was grown on SC media lacking uracil and with 2% raffinose at 30°C.

S.cerevisiae 6-1/13 *(ΔGAL)* carrying plasmids were grown in SC- media lacking uracil and tryptophan and with 2% raffinose at 30°C.

S.cerevisiae 6-1/13 *(ΔGAL)* carrying Fus-Δ7P&Δ10P was grown on SC- media lacking uracil and tryptophan with 2% galactose at 30°C.

S.cerevisiae MWK0128 was grown on SC- media lacking leucine and uracil and with 2% glucose at 21°C.

S.cerevisiae RLY10 was grown on SC- media lacking leucine, uracil and histidine with 2% glucose at 21°C.

S.cerevisiae RLY11 was grown on SC- media lacking leucine, uracil, histidine and tryptophan with 2% glucose at 21°C.

Prior to TRO analysis, *S.cerevisiae* 6-1/13 *(ΔGAL)* carrying plasmids (except for pYC10-7 and Fus-Δ7P&Δ10P) were grown on SC- media lacking uracil and tryptophan and with 2% raffinose until early log phase then induced with 2% galactose for 30-60 minutes. *ΔGAL* carrying pYC10-7 (WT) plasmid was grown in SC- media lacking uracil and with 2% raffinose until early log phase then induced with 2% galactose for 30-60 minutes. *ΔGAL* carrying Fus-Δ7P&Δ10P plasmids was grown on 2% galactose.

MWK0128, RLY10 and RYL11 were grown to early log phase at 21°C prior to performing the run on.

Blotting

Labelled Ends

3' recessed ends

Digests of 2 µg of DNA were incubated with 35 µCi (800 Ci/mmol) of [α-^{32}P] dATP, 5 mM of the respective 'cold' dNTP (dGTP, dCTP, dTTP) and 1 U Klenow fragment. Reactions were incubated at room temperature for at least 2 hours. Sample was then passed through a Sephadex G-50 column

equilibrated with TE-buffer as described by the manufacturer. The elute was heated at 95°C for 10 minutes, chilled on ice and added to the hybridisation solution.

Northern Blotting

Total RNA extraction

Total RNA was extracted from 10 ml culture grown to mid-exponential phase (OD_{600} = 1.0) using acidic phenol method (Current Protocols in Molecular Biology, 1995). Cultures were harvested and washed with cold water. The pellet was resuspended in 600 µl TES followed by 600 µl acid phenol, vortexed and incubated at 65°C for 60 minutes with periodic mixing. After incubation period, samples were chilled on ice for 5 minutes and spun at top speed for 5 minutes. Aqueous top layer was removed, added to 500 µl acidic phenol/chloroform, vortexed vigorously and spun down. This was repeated with 400 µl chloroform. After separating the phases, 40 µl of 3 M sodium acetate (pH 5.3) was added to the aqueous layer followed by 1 ml of 100% ethanol, precipitated for 30 minutes followed by a 70% ethanol wash. The RNA pellet was resuspended in 50 µl dH_2O. The RNA concentration was determined spectrophotometrically and the integrity determined on a native agarose gel.

Gel Preparation

A 1.5% denaturing agarose gel was made with 6 g RNase Free Agarose, 280 ml dH_2O, 40 ml 10X MOPS and 68ml formaldehyde. This was melted, cooled, mixed, poured in the hood and pre-run in 2L of 1X MOPS for 30 minutes at 80 volts.

Sample Preparation

Approximately 5 µg of total RNA was added to 23 µl of MMF denaturing solution, mixed and heated for 20 minutes at 60°C. To this 4 µl of 10x formamide loading dye was added and the sample was loaded onto the gel. This was run at 100 volts for 20 minutes then lowered to 35 volts for 8-9 hours or overnight depending on the size of the gel.

Membrane Transfer

The gel was rinsed with dH_2O and irrelevant parts of the gel cut away. RNA was blotted overnight onto Hybond-NX (Amersham-Pharmacia) by capillary transfer using 20X SSC. The membrane was washed with 2X SSC, air-dried, auto cross linked in a UV Stratalinker 1800. The filter was denatured with Denaturing Solution for 5 minutes followed by neutralization with Neutralization solution for 2 minutes prior to prehybridized overnight at 42°C.

Probing/Rinses

The appropriate random primed ^{32}P labelled probe was added and hybridised overnight at 42°C. The membrane was washed twice for 20 minutes each in 2x SSC, 1% SDS and placed in a Molecular Dynamics Phosphor Storage cassette overnight at room temperature. If necessary, more stringent washes were performed in 0.2x SSC, 0.1% SDS between 37-42°C for 20 minutes. Blots were exposed to Kodak X-omat film at –70°C or Molecular Dynamics Phosphor Storage screens. Quantification was performed using Image Quant software.

In Vitro Transcription

Synthetic Ribo-Probes Template preparation

DNA corresponding to the region desired was cloned into vector pGem4. About 5 µg of each plasmid was linearised with *Pvu*II in a volume of 30 µl. These reactions were ethanol precipitated, taken up in filter sterilized TE and added to *in vitro* transcription reactions

Transcription Reaction

The table below give the composition of the basic transcription reaction used to produce U-labelled synthetic RNA.

Component	Stock Concentration	Volume (µl)	Final Concentration
Water	--	3.0	--
Transcription Buffer	5x (Promega)	4.0	1x
DTT	100mM	1.0	10mM
RNasin	40 units/µl	0.5	1 unit/µl
rATP, rCTP, rGTP	2.5mM	4.0	500 µM each
rUTP	0.25mM	2.5	25µM
DNA template	0.5 – 1.0 µg/µl	1.0	
[α^{32}P] rUTP	800 Ci/mmol; 20 mCi/ml	2.0	40 µCi/reaction
RNA Polymerase	20 units/µl	1.0	1 unit/µl
		20.0	

Components, added in the order above, were incubated at 37°C for 2-3 hours in a PCR machine. Full-length ribo-probes were isolated from a denaturing polyacrylamide gel as described previously.

Transcription Run-on Analysis (TRO)

Transcription Run On analysis was performed as described in (Birse *et al.* 1997) and is summarised below.

Harvesting Cells

Cultures were grown to an OD_{600} = 0.5 –1.5 in 100 ml volume and if required were induced with 2% galactose for between 30 and 60 minutes (*S. cerevisiae* 6-1/13 *ΔGAL*). Cells were harvested at 4°C in a tabletop centrifuge for 4 minutes at 3000 rpm, washed with 5 ml ice-cold dH_2O. Cells were resuspended in 1.2 ml ice-cold dH_2O, separated into two aliquots: 250 µl for steady state analysis and 950 µl for TRO analysis. Cells for steady state analysis were frozen on dry ice immediately. The remaining cells were permeabilised with a final concentration of 0.05% sodium lauroyl sarcosinate

(Sarkosyl) on ice for 20 minutes. Cells were spun down at 5000rpm for 1 minute, supernatant removed and spun again.

Transcription

After removing all traces of the supernatant, cells were resuspended in 60 µl 2.5X TRO Buffer, 8 µl rNTPs (10 mM each ACG) and 3 µl 100 mM DTT. To this 5 µl of [α^{32}P] rUTP was added and incubated at 30°C for 5 minutes. Cells were washed briefly with 900 µl AE Buffer, spun down at full speed and the supernatant discarded.

RNA Extraction

The pellet was resuspended in 400 µl TES Buffer, transferred to a tube containing 400 µl acidic phenol pre-heated to 65°C and vortexed vigorously. This was incubated at 65°C for 1 hour (mixing 3-4 times), transferred to ice for 5 minutes and centrifuged at full speed. From this, 350 µl of the aqueous phase was re-extracted with 350 µl acidic phenol/chloroform. 300 µl of the aqueous phase was removed, 30 µl of 5 M LiCl was added followed by 825 µl 100% ice cold ethanol and precipitated overnight at -20°C. RNA was pelleted at top speed for 30 minutes, rinsed with 70% ethanol and re-pelleted.

Hydrolysis

RNA was resuspended in 40 µl dH_2O followed by partial hydrolysis of sample with 10 µl of 1.0M NaOH for 5 minutes on ice. This was then neutralized with 20 µl 0.5M Tris/0.5M HCl.

Hybridisation

To each sample 10 µg of single stranded M13 (with no insert) was added with 650 µl Hybridisation Mix and incubated at 42°C for 90 minutes with light shaking. Subsequently, the RNA was added to pre-hybridised filters at 42°C and hybridised over night.

Filter Preparation

Filters were made with a GIBCO-BRL slot blot and Hybond-NX membrane (Amersham-Pharmacia) as per manufacturer's instructions. The membrane was washed with 5X SSPE, 5 µg of each ssM13 probe (diluted in 200 µl of 5X SSPE) was added to each slot and rinsed with 200 µl 5X SSPE. The filter was denatured with Denaturing Solution for 5 minutes followed by neutralization with Neutralization solution for 2 minutes. The filter was air-dried, auto cross-linked with UV Stratalinker 1800, added to pre hybridisation solution in appropriate size tube and incubated at 42°C in mini-Hybaid oven overnight.

Washing Filters

Filters were washed 4 times for 10 minutes in 2X SSC; 0.1%SDS in 1 litre volumes at room temperature. This was left overnight to expose a Molecular Dynamics Phosphor storage cassette. If needed, more stringent washes were performed in 0.2x SSC, 0.1% SDS between 37-42°C for 20 minutes. Filters were exposed to Kodak X-omat film at –70°C with intensifying screens or Molecular Dynamics Phosphor Storage screens. Quantification was performed using ImageQuant software.

Part I: Convergent Transcription

Preface

The aim of this study was to investigate what effect convergent orientation has on nascent transcription of highly expressed Pol II genes. In Section 1, nascent transcription of the highly expressed, co-induced *GAL10-7* gene cluster was investigated in its wild type tandem arrangement. These genes were then rearranged to create a set of convergently transcribing genes. In Section 2, the intergenic region of the convergent orientation was deleted and nascent transcription of the fused construct was studied.

Introduction

As previously described, the genome of *Saccharomyces cerevisiae* is highly compact with over 70% represented by open reading frames (ORFs) (Feldmann *et al.* 1994). This high gene density is largely a result of short intergenic regions between adjacent gene ORFs. These average 618 bp between diverging promoters and 517 bp between tandem (head to tail) genes having promoter/terminator combinations (Dujon 1996). Convergent genes average only 326 bp apart (Dujon 1996) with several examples of less than 300 bp (Barker *et al.* 1987; McGrath and Varshavsky 1989; Malavasic and Elder 1990; Peterson and Myers 1993). Such a limited intergenic sequence would allow only ~150 bp per gene for proper termination outside of the neighbouring ORF. This is less sequence than the few termination regions so far defined in *S.cerevisiae* genes (Osborne and Guarente 1989; Aranda *et al.* 1998a; Aranda *et al.* 1998b). Analysis of adjacent gene sets in *S.cerevisiae* determined that less than 6% of co-expressed, adjacent genes are in a convergent orientation, with a bias to divergent transcription (Kruglyak and Tang 2000). These statistics suggest that there may be evolutionary pressure to select against convergently arranged, co-transcribed genes (Dujon 1996).

With such a densely packed genome it is not surprising that interference has been detected between transcribing genes (Springer *et al.* 1997; Greger and Proudfoot 1998; Puig *et al.* 1999). Transcriptional Interference (TI) occurs when an elongating polymerase fails to terminate properly and so reads into a co-transcribed downstream transcription unit, thereby reducing promoter activity.

TI has also been documented to affect the proper use of both ARS and CEN elements and in so doing inhibits either replication or chromosomal segregation in mitosis (Hill and Bloom 1987; Snyder *et al.* 1988; Tanaka *et al.* 1994). Sequence elements that mediate enhanced meiotic gene conversion are also susceptible to TI (Rocco *et al.* 1992). At the nascent level, the *S.pombe* genes *nmt1* and *nmt2* were shown to transcribe into two poorly defined convergent ORFs with little detectable interference (Hansen *et al.* 1998). While the *nmt* genes are both transcribed at high levels upon induction, the convergent genes *avn2* and *gut2* are ubiquitously expressed at much lower levels. The possibility that *avn2* and *gut2* are expressed separately remains to be determined. In contrast, steady state analysis in *S.cerevisiae* has shown that the elimination or constitutive activation of one promoter in a convergent pair of genes affects the steady state levels of mRNA generated from the opposing gene (Peterson and Myers 1993; Puig *et al.* 1999).

In the tightly regulated metabolic pathway of galactose fermentation, the tandem *GAL10* and *GAL7* genes are affected by TI (Greger and Proudfoot 1998). A fraction of polymerases elongate beyond *GAL10* and have an inhibitory effect on transcription initiation from the downstream *GAL7* promoter. Furthermore mutations in the *GAL10* poly(A) signal which prevent Pol II termination have a severe effect on the levels of *GAL7* gene expression. Interestingly, *GAL7* expression is recoverable by over-expression of transcription factor Gal4p (Greger *et al.* 2000). Presumably read-through polymerases compete with Gal4p for access to the *GAL7* promoter and in so doing inhibits transcriptional activation, a mechanism termed promoter occlusion.

Efficient gene expression requires 3' end processing and termination to function properly. Transcription initiation and elongation of *S.cerevisiae* Pol II has been extensively studied but the mechanism for termination and coupled mRNA 3' end processing is less well understood. However, mutations in cleavage/polyadenylation factors also affect termination, highlighting the interdependence of these functions (Birse *et al.* 1998). The polyadenylation signals in budding yeast genes, described in more detail in a previous section, are generally complex and degenerate. In the *GAL10* and *GAL7* gene cluster the poly(A) signals have been characterised and shown to play a crucial role in mitigating TI. Deletions of as little as 40 bp in the *GAL10* 3'-UTR result in a

significant decrease in *GAL10* poly(A) site usage with transcripts reading through the 3' end processing signals and terminating at the *GAL7* poly(A) site located about 1850 bp further downstream (Greger and Proudfoot 1998). In *GAL7*, mutational analysis of the poly(A) region has demonstrated that a 26 bp sequence element containing a stretch of alternating TA residues is located 7 nt upstream of the poly(A) site and appears to be necessary and sufficient for 3' end formation (Abe *et al.* 1990). In absence of efficient 3' end processing signals, transcription continues beyond where it would normally terminate and can influence downstream gene signals and functions.

Part I - Section 1

Convergent nascent transcription in *GAL10* and *GAL7*

In these studies, transcriptional termination and interference in the wild type, tandemly arranged *GAL10-7* cluster was analysed by TRO analysis. Subsequently, the *GAL10-7* construct was manipulated to create co-transcribed convergent genes. Comparing the transcription profiles of the tandem and convergent orientation of the same gene set elucidates information about the effects of orientation on transcriptional interference.

1.1 TRO analysis of nascent transcription

TRO analysis of *S.cerevisiae* was performed as described in Birse *et al.* (1997)(Figure I 1.1). In this protocol, a yeast culture is grown to an optimal stage of growth (OD_{600}=0.05-0.1) and harvested by centrifugation. The cells are then permeabilised to allow uptake of radioactive nucleotides by treatment with 10% Sarkosyl. After a short incubation, the cells are spun down and resuspended in optimised TRO buffer, ribonucleotides (lacking rUTP) and radiolabeled [α^{32}P] rUTP. Transcription is allowed to proceed for approximately 5 minutes before the cells are harvested and RNA isolated. Hydrolysis and neutralisation results in short fragments of radiolabeled RNA representing the nascent, polymerase-attached transcripts that were transcribed during the brief incubation with radiolabeled rUTP. The relative abundance of radiolabeled RNA at different positions on the template DNA is detected by hybridisation with a bank of single stranded DNA probes. Thus, the radioactive signal intensity detected over a given probe should be directly related to transcribing polymerase density over that region of DNA.

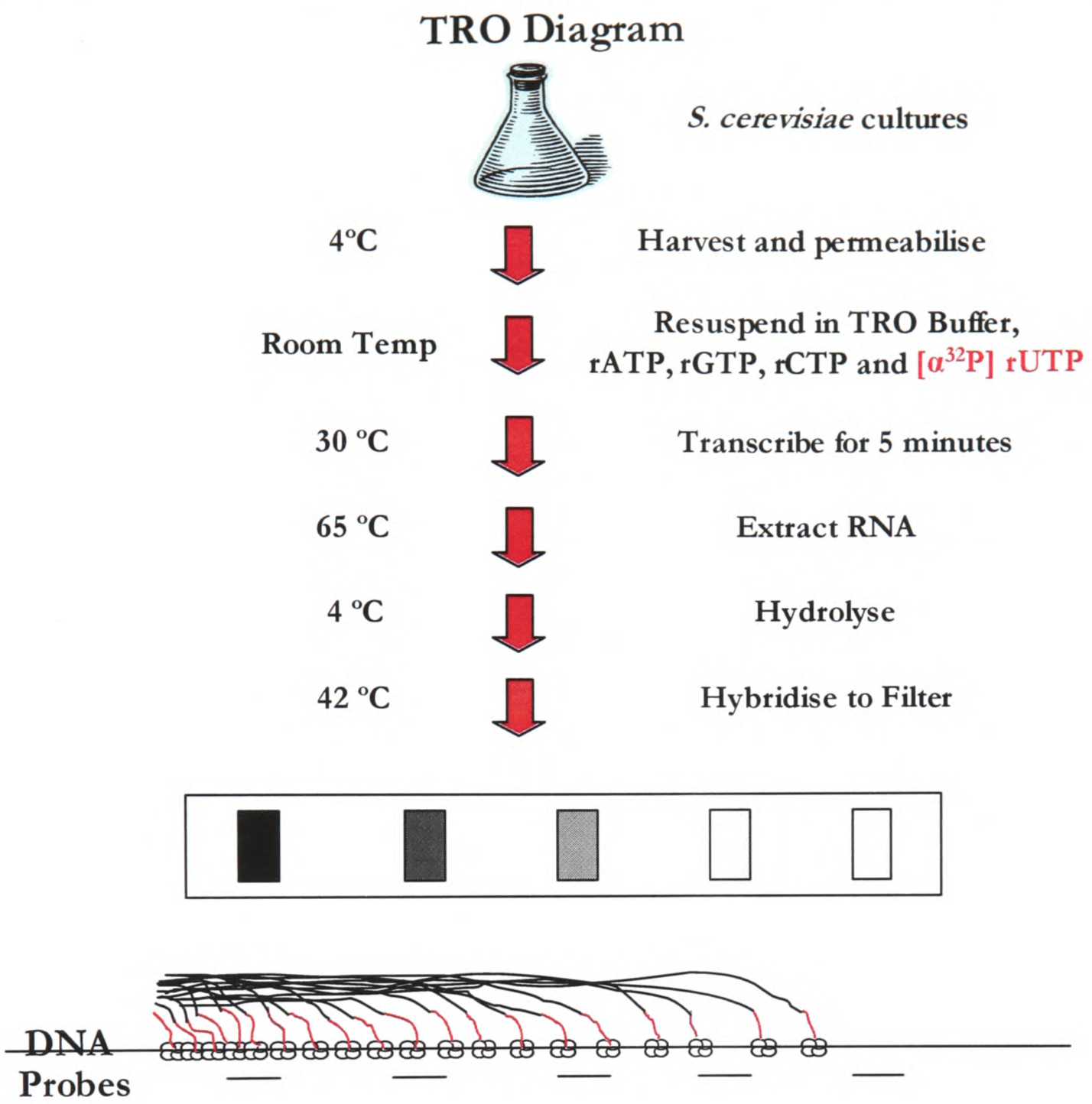


Figure I 1.1 Diagram of the transcriptional run-on (TRO) protocol. 100 ml cultures were grown for each strain with the appropriate energy source and at the desired temperature. More detail of the steps involved is given in the Methods Section. The region of transcribed RNA that is radiolabeled is shown in red.

1.2 Transcription of WT *GAL10* and *GAL7* with redesigned probes

A previous study determined the transcription profile of *GAL10* and *GAL7* in the wild type tandem conformation. This was achieved by TRO analysis of a centromeric plasmid containing the *GAL10-GAL7* genes (pYC10-7) transformed into a *GAL10-GAL7* deletion strain (Δ*GAL*) (Greger and Proudfoot 1998). In this original study, single stranded probes were designed to detect transcription in the sense orientation with respect to wild type *GAL10* and *GAL7*. However, to perform TRO analysis on the *GAL* gene convergent orientation, single stranded M13 probes were needed for both strands of the DNA. As these had not been made in the previous analysis, all of the probes were remade with additional probes designed to detect nascent transcription in the areas where the genes

converged. The TRO of the plasmid pYC10-7 in the strain Δ*GAL* was repeated to validate the new probes.

A diagram of pYC10-7 indicating the location of the redesigned probes is given in Figure I 1.2. In the figures and text the probes have been colour coded to enable differentiation of the *GAL7* sense (blue) and *GAL10* sense (yellow) strand probes for the convergent orientation. The probes cover the same region of the double stranded DNA template but, as they are single stranded DNA probes, only hybridise with transcripts from one strand of the duplex. To generate the probes, PCR products for each region indicated were synthesized with primers that contained *Xba*I and *Eco*RI restriction sites at either end of the product. The PCR products were digested with *Xba*I and *Eco*RI and purified to generate direction specific 'sticky' ends. These fragments were ligated into both M13mp18 and M13mp19 vectors that had been digested with *Xba*I and *Eco*RI. The use of both vectors, which differ only in the orientation of the poly linker, enabled efficient generation of single stranded probes for both orientations from the same PCR fragment.

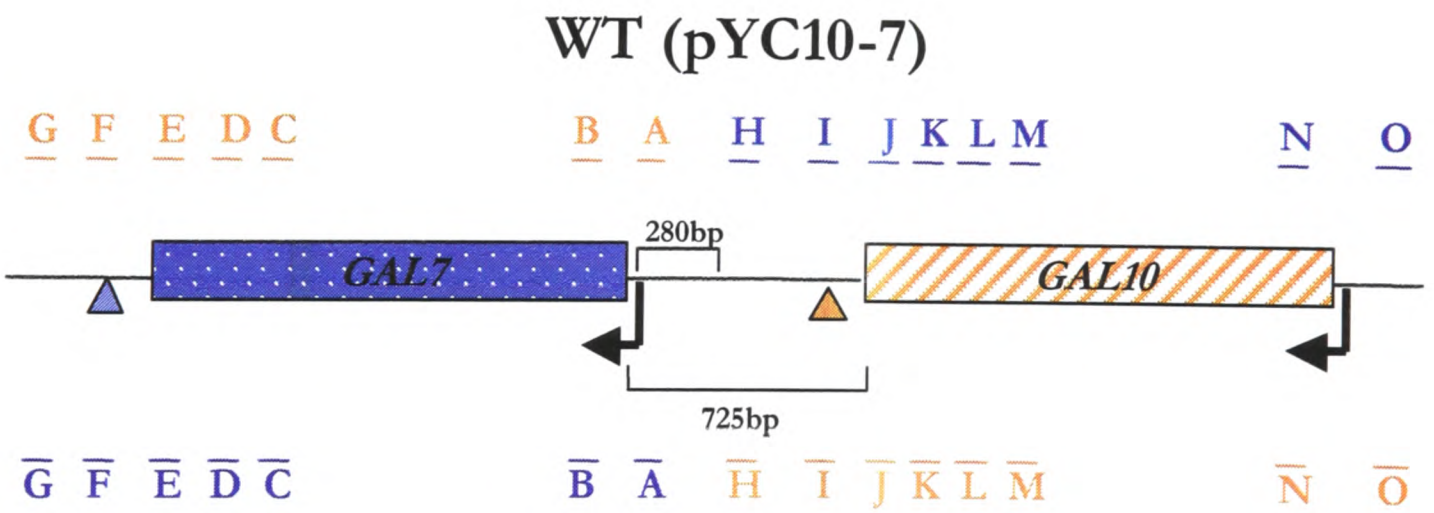


Figure I 1.2 pYC10-7 (WT) tandemly arranged *GAL10* and *GAL7* genes (Greger and Proudfoot 1998). The direction of transcription for each gene is indicated with an arrow. Location of single stranded M13 probes are indicated (exact sequences identified in the Materials Section). The WT 725 bp intergenic region and 280 bp *GAL7* promoter region are shown. Blue triangle indicates *GAL7* poly(A) site. Yellow triangle indicates *GAL10* poly(A) site.

In addition to probes containing *GAL10-7* sequences, two additional probes were used as controls in these experiments. Probe 'm' is an M13 probe with no insert that serves as a negative control for background hybridisation and probe 'a' contains a fragment of the *ACT1* gene from *S.cerevisiae* that serves as a positive transcriptional control.

A typical transcription profile for pYC10-7 using the new probes is shown in Figure I 1.3. Quantitation of transcription levels for this membrane is shown in Figure I 1.4. To quantitate the extent of radiolabeled RNA hybridised to each slot, a PhosphoImage of the membrane was taken, scanned and analysed with ImageQuant software package. Each probe was corrected for background, adjusted for the number of U's in that probe and, when compared across filters, to the actin positive control probe (a).

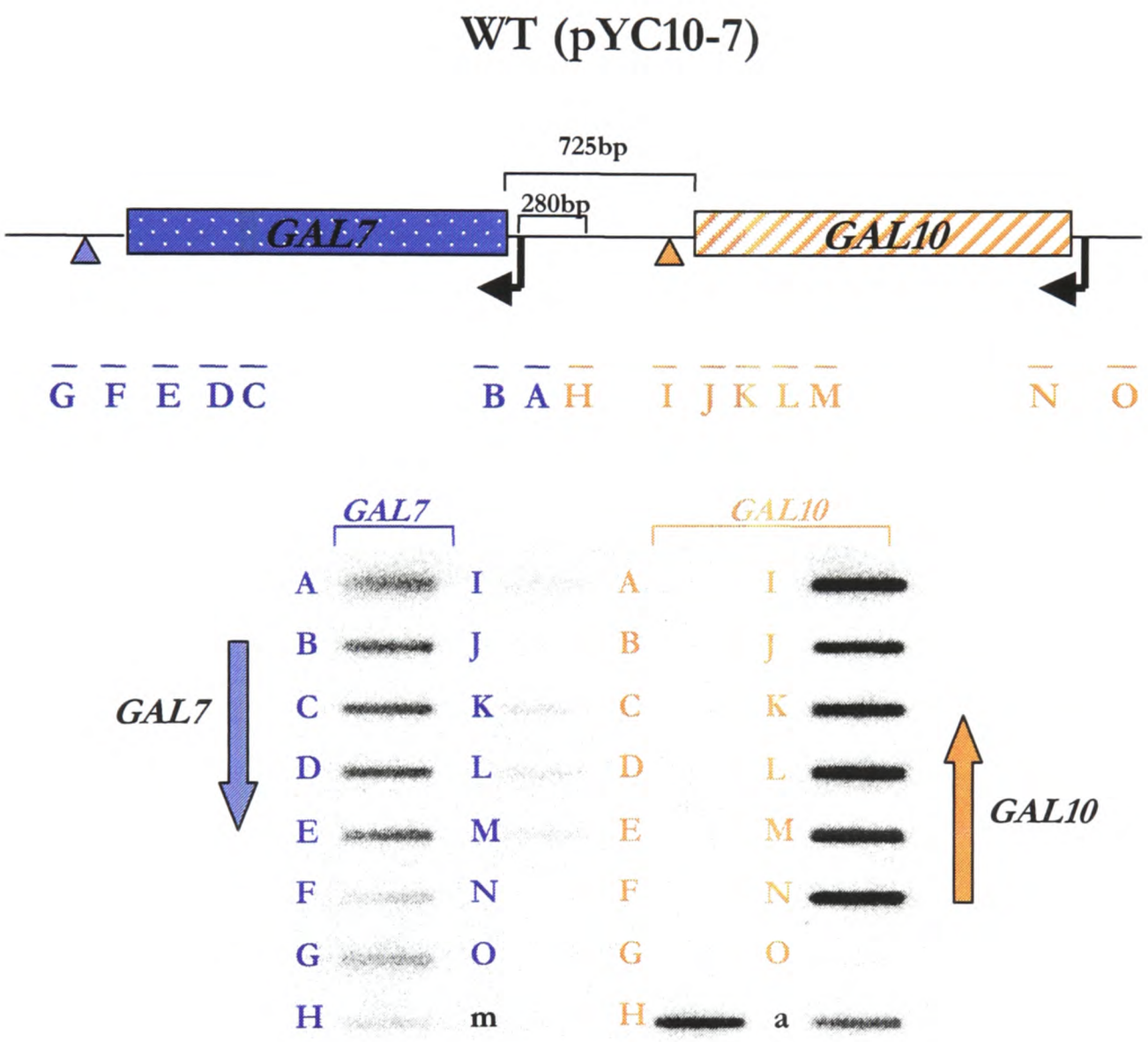


Figure I 1.3 TRO of pYC10-7 (WT) transformed cells. Diagram shows the tandem arrangement of the *GAL10* and *GAL7* genes with direction of transcription indicated with an arrow. The 725 bp intergenic region and 280 bp *GAL7* promoter region is identified. Blue triangle indicates *GAL7* poly(A) site. Yellow triangle indicates *GAL10* poly(A) site. The position of single stranded M13 probes used are shown and described in the Materials Section. Probe 'm' is an M13 probe without an insert serving as a negative hybridisation control. Probe 'a' is an M13 probe with a fragment of the *S.cerevisiae* actin gene serving as a positive transcription control.

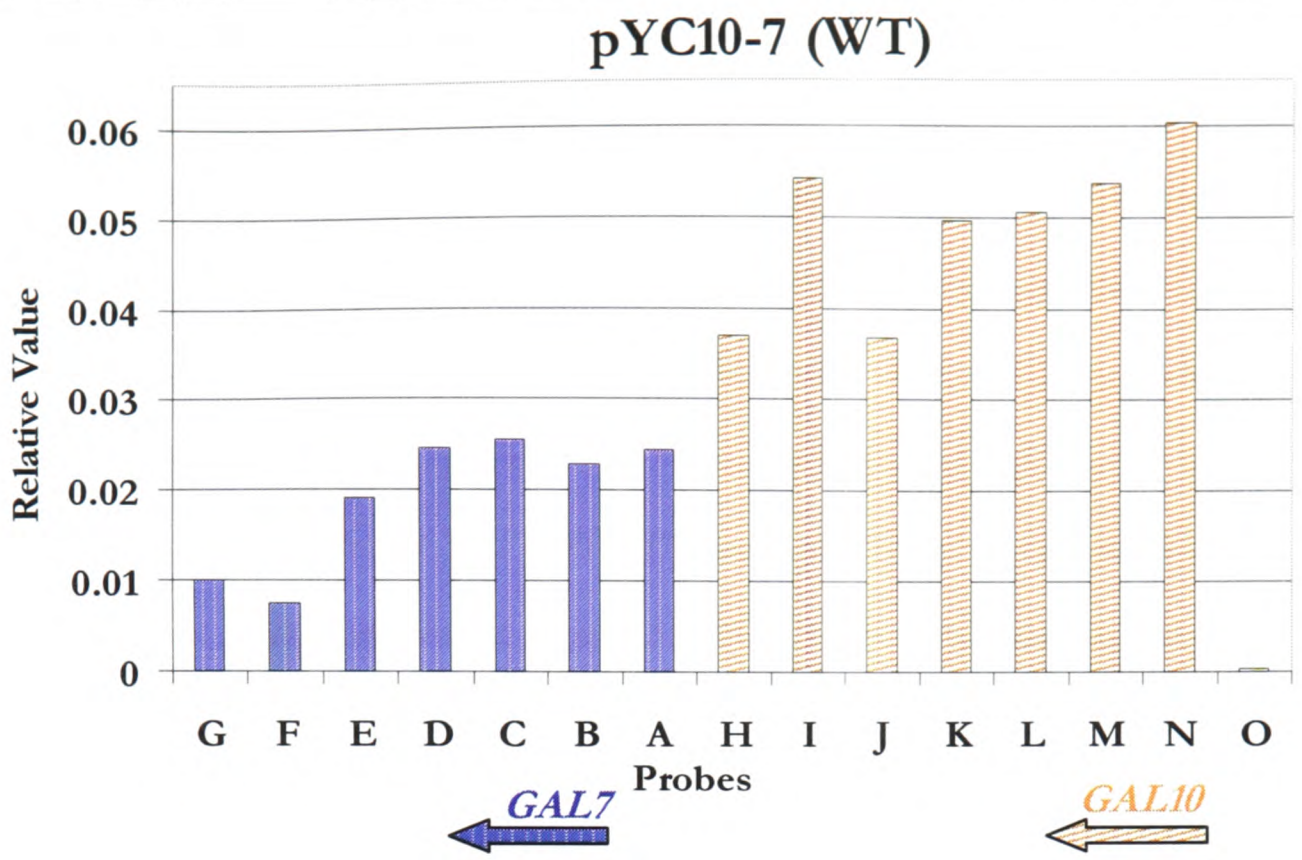


Figure I 1.4 Quantitation of data shown in Figure I 1.3 - TRO of WT plasmid. Arrows indicate direction of transcription for each gene. Signals were quantified using a PhosphoImager (Molecular Dynamics). The values were corrected for background hybridisation and 'U' content in each probe. Signals were normalized to the actin positive transcription probe 'a' on each filter to allow direct comparison of transcription levels.

As would be expected, the lack of signal over probe O indicates that there is no transcription occurring before the *GAL10* promoter. A higher level of nascent transcription is seen over the *GAL10* gene than the *GAL7* gene. Most of the transcription detected over *GAL10* fails to terminate beyond the *GAL10* poly(A) signals, reads through the intergenic region and accumulates over the *GAL7* promoter. In the previous study (Greger and Proudfoot 1998) analysis of a plasmid containing the *GAL10-GAL7* genes with the *GAL7* promoter inactivated revealed that most *GAL10* transcription terminates upstream of the *GAL7* transcription start site. Consequently, run-on signals over the *GAL7* gene shown in Figure I 1.3 are likely to derive from transcription initiation on the *GAL7* promoter. These results are consistent with the data presented in Greger and Proudfoot (1998) and serve to validate the newly designed probes.

1.3 Transcription of *GAL10* and *GAL7* genes in a convergent orientation

To investigate the effect of concurrent, convergent transcription, the *GAL7* gene was inverted to generate pYC10-7Con (Con) (Figure I 1.5). This was achieved by complete digestion of pYC10-7 with *Xba*I followed by partial digestion with *Eco*RV to create a vector with the *GAL7* gene deleted. This 7658 bp band was isolated and the ends filled in. A *GAL7* insert fragment was generated by PCR with blunt ends and a unique *XhoI* site 103 bp after the poly(A) site with primers *GAL7*+100

and Up*GAL7* (restriction sites and primers are labelled in Figure I 1.6). This fragment was ligated to generate junctions as indicated in Figure I 1.7. The natural polyadenylation sites and a portion of the *GAL10-7* intergenic region were maintained. The WT intergenic region between *GAL10* and *GAL7* (from 3' end of *GAL10* ORF to the 5' end of the *GAL7* ORF) is 725 bp including the 280 bp promoter region of *GAL7* (see Figure I 1.2). In Con, the intergenic region is 533 bp from the 3' end of *GAL10* ORF to the 3' end *GAL7* ORF. This includes 240 bp beyond the 3' end of the *GAL7* ORF and 293 bp downstream of the 3' end of *GAL10* ORF. The Con plasmid, as with all future plasmids discussed, was verified by sequencing.

Con (pYC10-7Con)

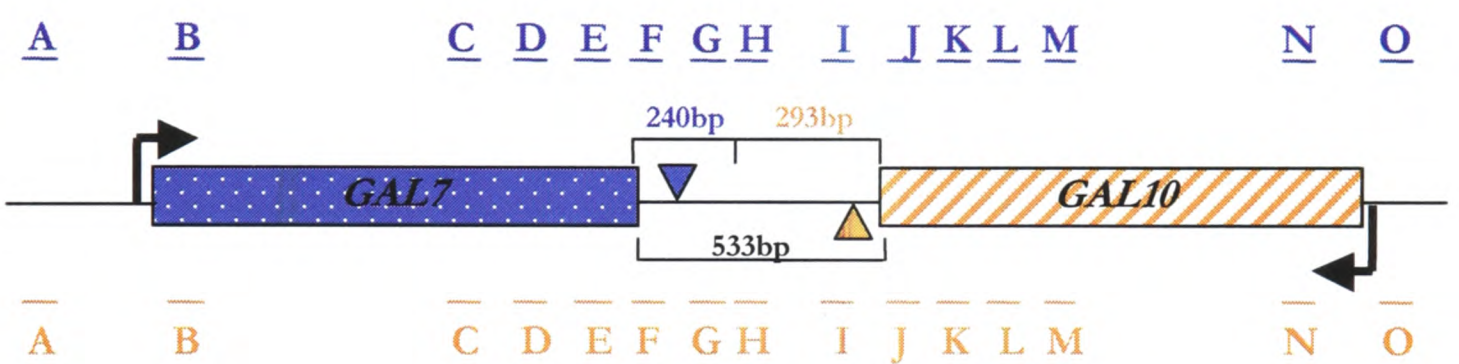


Figure I 1.5 pYC10-7con (Con) tandemly arranged *GAL10* and *GAL7* genes. The direction of transcription for each gene is indicated with an arrow. Location of single stranded M13 probes are indicated (exact sequenced identified in the Materials Section). This construct includes 240 bp downstream of *GAL7* and 293 bp downstream of *GAL10* to make a unique intergenic region of 533 bp. Blue triangle indicates *GAL7* poly(A) site. Yellow triangle indicates *GAL10* poly(A) site.

Generation of Con

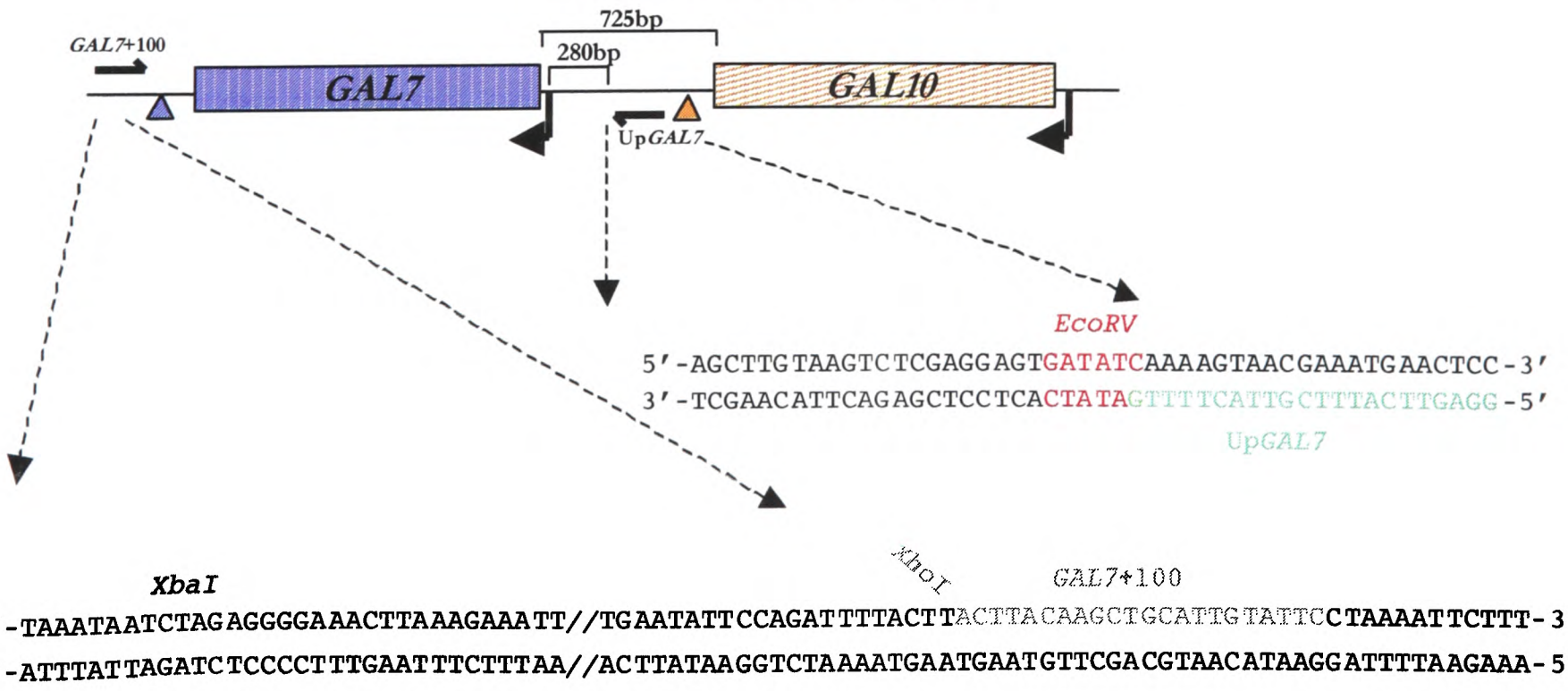


Figure I 1.6 Diagram indicates primers used to generate the convergent construct. Primers used in PCR generation of the *GAL7* insert are shown in green. Restriction sites are identified in red.

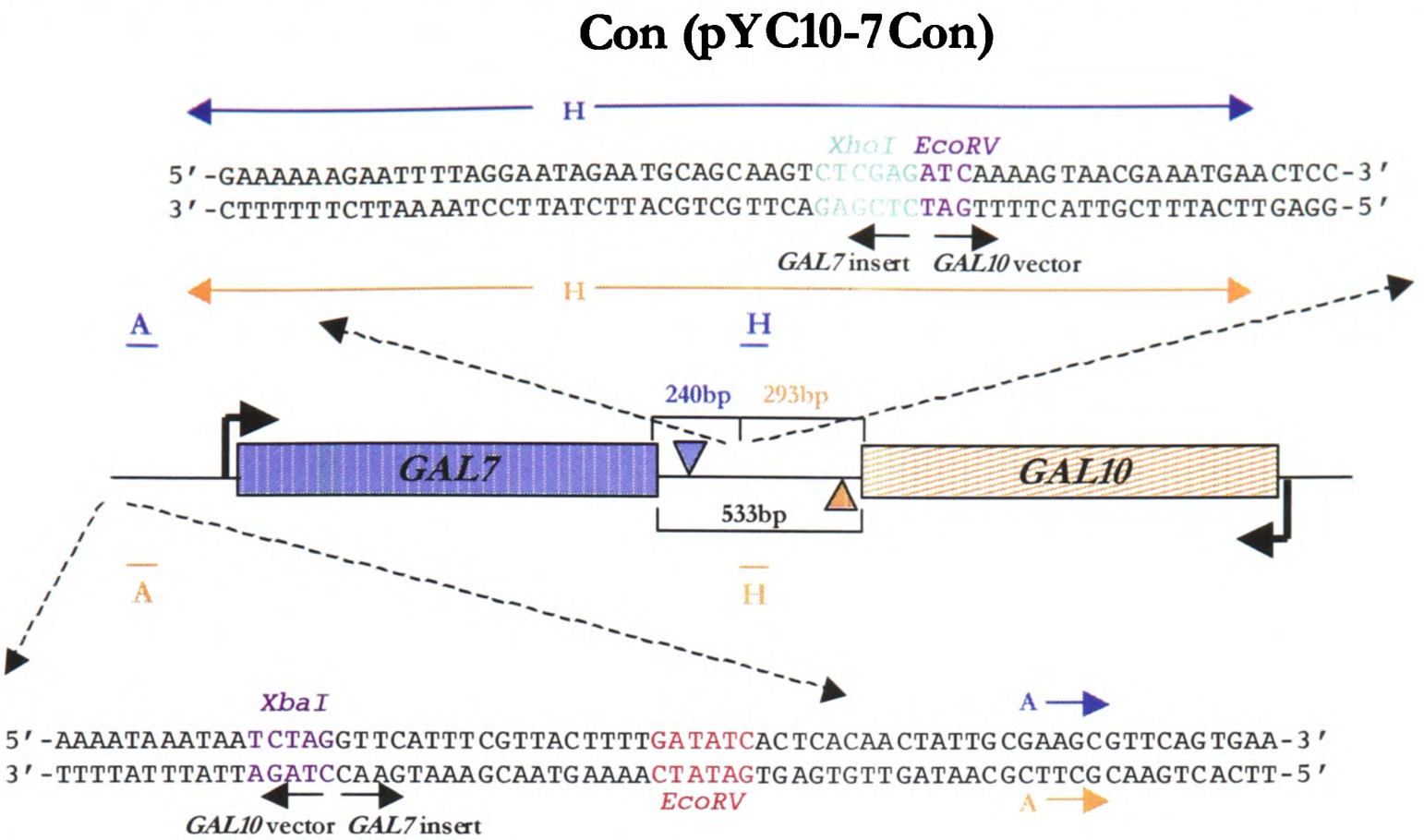


Figure I 1.7 Diagram of the convergent construct (Con). Sequence is given for the junctions. *EcoRV* site is shown in red. An *XhoI* restriction site was inserted at the end of the primer used to generate the *GAL7* insert and is indicated in green. The restriction sites in purple were used for cloning and are no longer functional in this construct. The M13 single stranded probes that detect RNA generated at the junctions are labelled and the sense strand of that the probe recognises is colour coded.

A diagram of probes and typical slot blot membrane for TRO of Con is shown in Figure I 1.8. Quantitation of slot blot membrane as described previously is shown in Figure I 1.9. TRO analysis of Δ*GAL* yeast strain transformed with Con yields a nascent termination profile similar to the WT (pYC10-7) tandem situation with a few notable exceptions. Transcription in the *GAL10* and *GAL7* genes seems to terminate near the same probes regardless of orientation (Figure I 1.3 and Figure I 1.8). However, the level of read-through transcription from *GAL10* decreased in Con [compare *GAL10* probe G of Con with *GAL7* probe A of WT]. *GAL7* transcription terminated shortly after the wild type poly(A) site in both orientations. Additionally, the levels of transcription appeared higher for both *GAL10* and *GAL7* genes in Con (compare the levels of *GAL10* and *GAL7* nascent transcripts to the endogenous actin probe (a) for WT and Con plasmids). In particular, there was a two-fold increase in the level of nascent transcription from *GAL7*. This may be the result of alleviation of promoter occlusion occurring from read-through transcription beyond the *GAL10* poly(A) site into the *GAL7* promoter. The relative increase in *GAL7* gene expression is reflected by the steady state mRNA analysis (Northern Blot) shown in the Figure I 1.10 (lanes 1 and 2) where *GAL7* mRNA levels increase 2.3 fold with the Con construct. However, we also note that *GAL10* mRNA levels increase 1.7 fold in the Con construct as compared to WT. In this case TI is unlikely

to be important since no TRO signal is detectable over the *GAL10* promoter (probe O) in WT. It is therefore possible that other factors such as supercoiling effects may be involved.

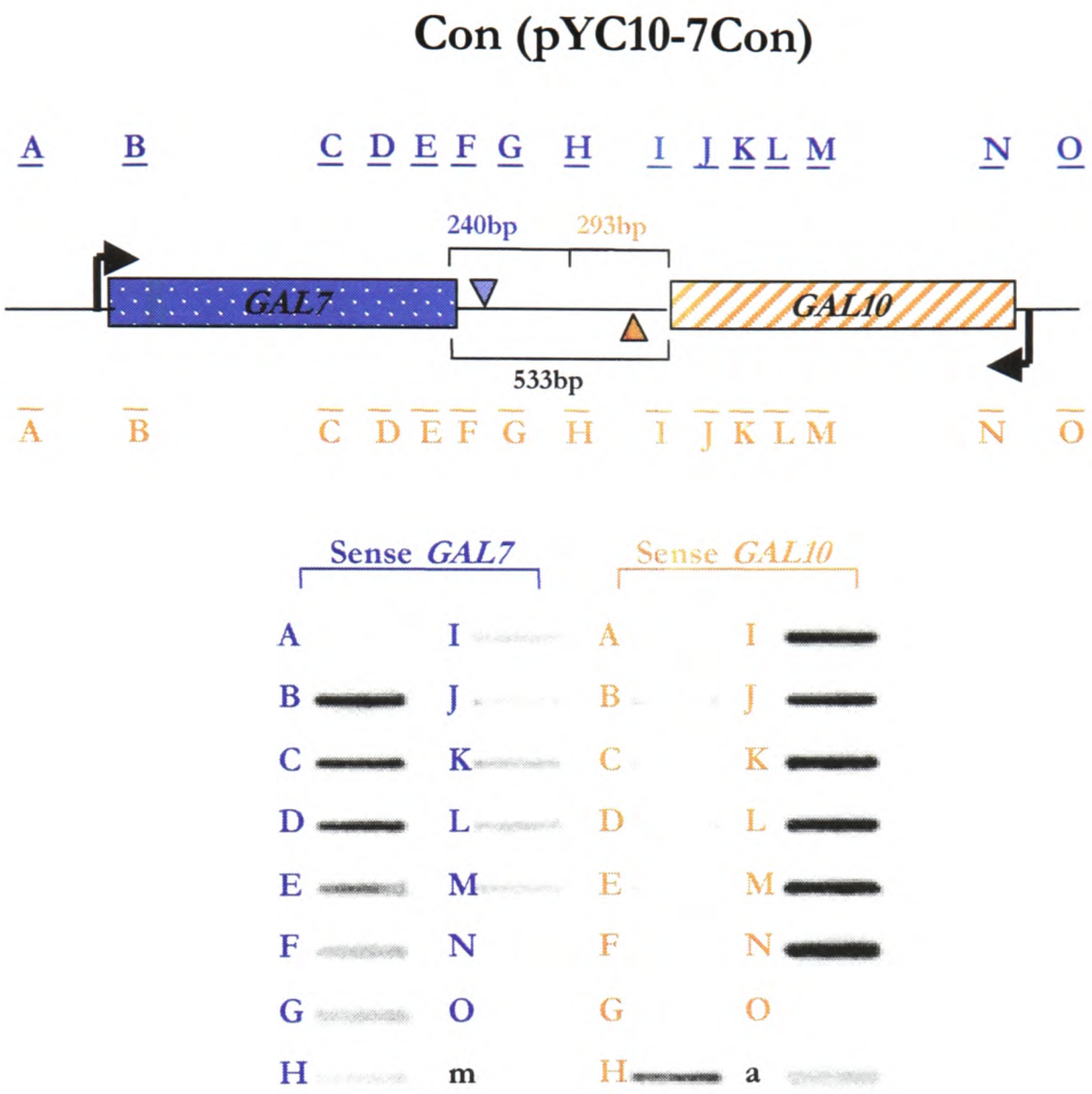


Figure I 1.8 TRO of pYC10-7con (Con) transformed cells. Diagram shows the tandem arrangement of the *GAL10* and *GAL7* genes with direction of transcription indicated with an arrow. Construct includes 240 bp downstream of *GAL7* and 293 bp downstream of *GAL10* to make a unique intergenic region of 533 bp. Blue triangle indicates *GAL7* poly(A) site. Yellow triangle indicates *GAL10* poly(A) site. The position of single stranded M13 probes used are shown and described in the Materials Section. Probe 'm' is an M13 probe without an insert serving as a negative hybridisation control. Probe 'a' is an M13 probe with a fragment of the *S.cerevisiae* actin gene serving as a positive transcription control.

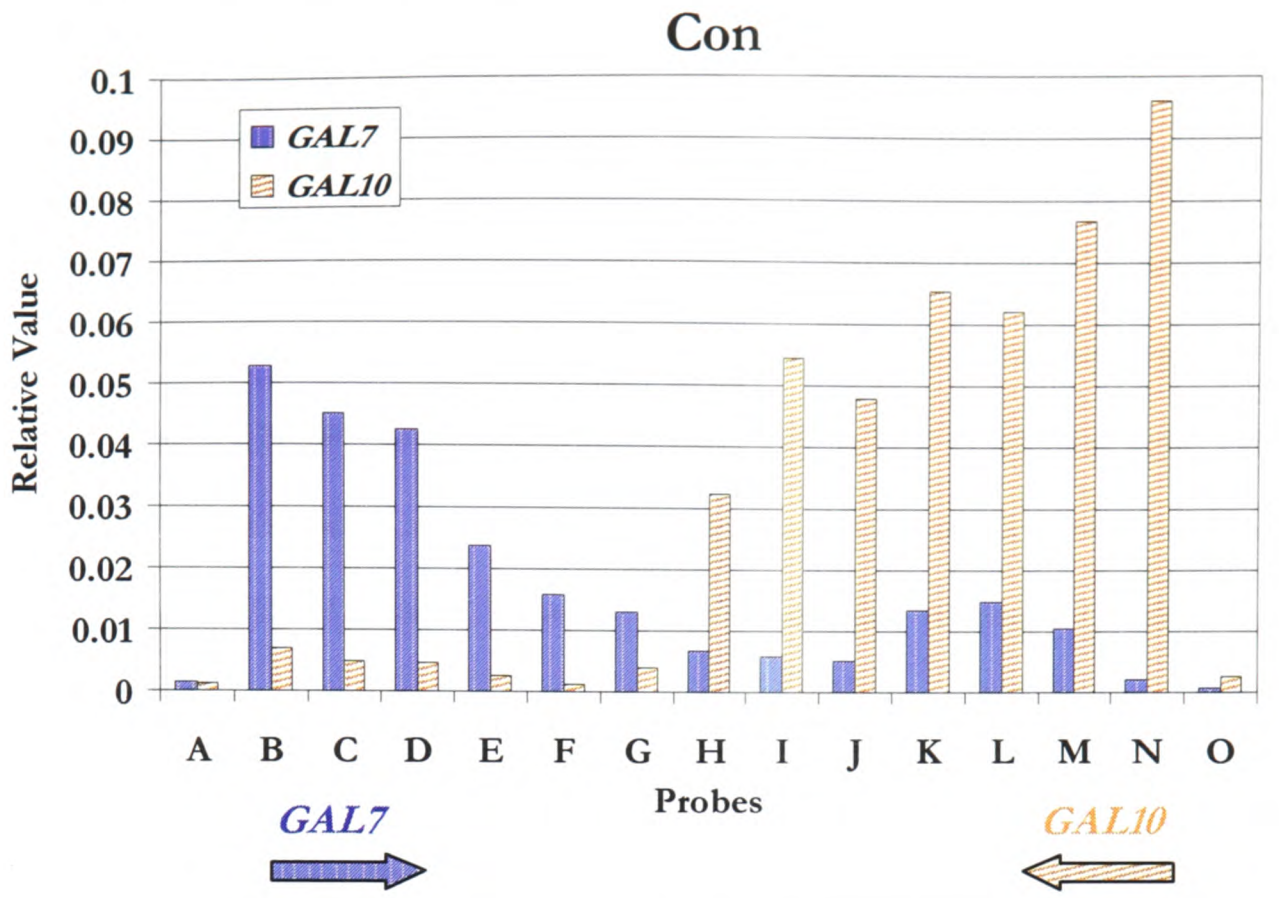
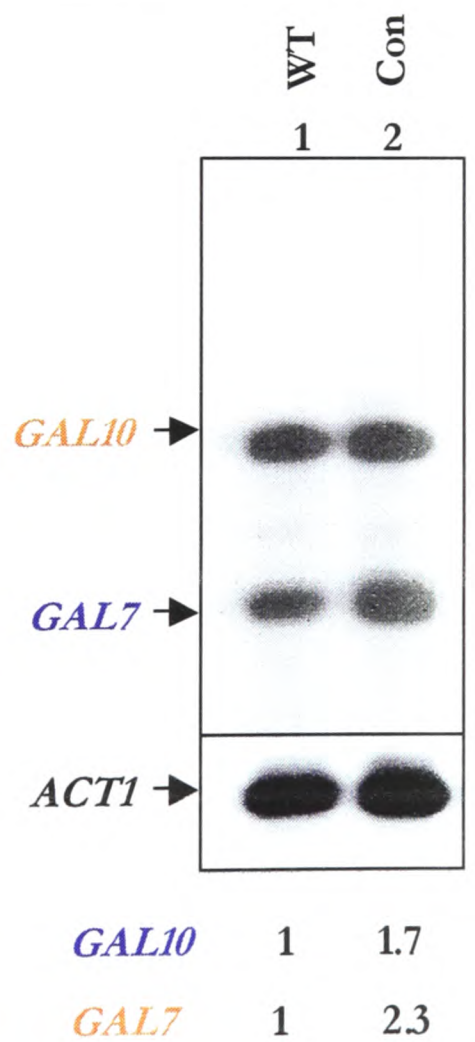


Figure I 1.9 Quantitation of data shown in Figure I 1.8 - TRO of Con plasmid. Arrows indicate direction of transcription for each gene. Signals were quantified in a Phosphoimager (Molecular Dynamics). The values were corrected for background hybridisation and 'U' content in each probe. Signals were normalized to the actin positive transcription probe 'a' on each filter to allow direct comparison of transcription levels. Blue arrow and bars indicate *GAL7* transcription. Yellow arrow and bars indicate *GAL10* transcription.



GAL10	1	1.7
GAL7	1	2.3

Figure I 1.10 Northern blot with total RNA from pYC10-7 (WT) (lane 1), pYC10-7Con (Con) (lane 2). Probe was a 1.5 kb fragment generated by double restriction digest with *BglII* and *SalI* of the WT *GAL10-7* locus which contained 427 bp of 5' *GAL7* and 403 bp of 3' of *GAL10*. Filter was stripped and re-probed with *ACT1* probe as a loading control. The *ACT1* probe was a 567 bp fragment (277 –844 bp 3' of the *ACT1* start codon). Numbers shown beneath gel indicate quantitated levels of steady state *GAL10* and *GAL7* mRNA, corrected for actin and compared using steady state level of *GAL10* and *GAL7* mRNA from WT as 1.

Discussion

This study provides further information about the capacity of *GAL10* and *GAL7* to interact in their WT arrangement. The level of transcription before the *GAL7* promoter (*GAL7* probe A) in WT clearly illustrates the extent of read-through transcription into the *GAL7* promoter region. This is in contrast to the same region in Con (Figure I 1.8), which showed no detectable signal before the promoter of *GAL7*. Presumably, elimination of transcribing polymerases over the promoter for *GAL7* in Con (Figure I 1.8, probe A) permits the unimpeded binding of Gal4p and other transcription factors (Greger *et al.* 2000) and thus increases transcription of *GAL7*.

The termination profile of Con demonstrates the impact of co-transcribed convergent genes given intergenic termination and processing signals. The steady state transcripts generated from Con are indistinguishable in size from those generated in the WT conformation. Additionally, the approximate region of termination for both genes at the nascent level is comparable. This suggests that convergent, co-transcribed genes that possess termination and processing signals, as well as an adequate separation, efficiently express mRNA.

While efficient mRNA synthesis occurs in both the WT and Con, the level of nascent transcription does appear to vary. The quantified 2 fold increase in *GAL7* nascent transcripts from WT to Con is lower than previously reported 3 fold increase when promoter occlusion is eliminated (Greger and Proudfoot 1998). However, the 2-fold increase reported here could under-represent the true increase in transcripts that initiated at the *GAL7* promoter as the WT level of transcription detected in the *GAL7* region is a combination of transcripts that initiated at *GAL7* and the read-through from the *GAL10* promoter. This observed two-fold increase is consistent with the increase at steady state levels (Figure I 1.10, lane 2). Thus we observe an increase of 2.3 fold in *GAL7* mRNA for Con as compared with WT. These data further support the existence of transcriptional interference in the WT orientation of *GAL10* and *GAL7*.

The increased level of *GAL10* initiation is less easily understood. It is unlikely that TI is important in this case as no TRO signal is detectable over the *GAL10* promoter (probe O) in WT. It has been

reported that negative supercoils, generated behind a transcribing polymerase, can stimulate promoter initiation (Wu *et al.* 1988). It is possible that this supercoiling effect would also stimulate initiation over the *GAL7* if it were significant enough to increase initiation of *GAL10* by 1.7 fold. However, we do not see a large enough increase in *GAL7* to suggest the additive effects of both relief of promoter occlusion and increased promotion due to supercoiling effects. In fact, in this system there is only a 2.3 fold increase between WT and Con as compared to the >3 fold increase reported previously from relief of promoter occlusion (Greger and Proudfoot 1998). These two results do not necessarily conflict, as it is possible that the *GAL7* promoter in both of these constructs has reached a maximum rate of initiation and is therefore not able to generate any more transcripts despite other stimulatory influences.

Part I - Section 2

Transcriptional Collision

As demonstrated in the previous section, co-transcribed convergent genes can function in a similar manner to tandemly arranged co-transcribed genes. The nascent and steady state data demonstrate that the convergent *GAL10-7* construct (Con) generates transcripts in positions consistent with the WT processing and termination signals. However, what would happen if these genes are prevented from terminating properly before encountering the transcription unit of the converging gene? In the following chapter, the entire intergenic region of Con was deleted to investigate the effect of termination deficiencies on convergently transcribed genes. The results graphically illustrate that, in the absence of normal Pol II termination, inhibition of both genes occurs by a mechanism we refer to as transcriptional collision.

2.1 Transcriptional collision in non-terminating convergent transcription

To test the effect of transcriptional interference in convergently arranged genes that do not terminate, the *GAL10* and *GAL7* genes were fused together. This was achieved by the deletion of the entire intergenic region from Con to make the plasmid pYC10-7Fus (Fus). The fused construct (Fus) was generated by self-ligating long-range PCR products from primers Fuse7 and Fuse10 (Figure I 2.1). This created a construct with intact ORF's for each gene but lacking any termination or processing signals beyond the 3' end of either gene (Figure I 2.2). In should be noted that in so doing, sequences complimentary to the probes F, G, H & I were deleted. Additionally, the 209 bp *GAL7* poly(A) site fragment (*GAL7*: 1118 to 1327) in the forward orientation for *GAL7* was inserted into the *SacI* site of the vector backbone upstream of the *GAL10* promoter (-442) (see blue triangle in Figure I 2.2A). It was thought that this poly(A) site could function as a 3' end signal to generate *GAL7* transcripts if necessary.

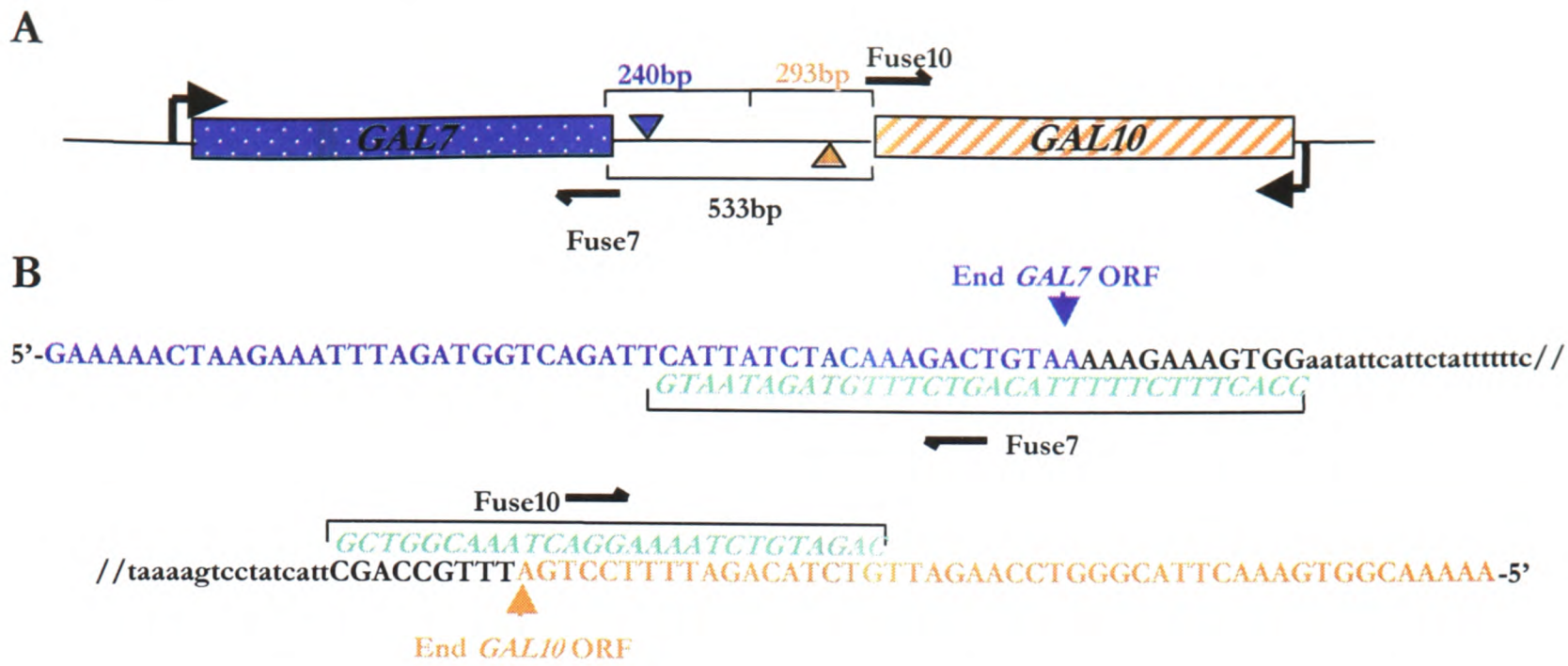


Figure I 2.1 (A) pYC10-7Con (Con) convergently arranged *GAL10* and *GAL7* genes. Primers used in generating the fused construct ('Fuse7' and 'Fuse10') are labelled. The direction of transcription for each gene is indicated with an arrow. Location of single stranded M13 probes are indicated. This construct includes 240 bp downstream of *GAL7* and 293 bp downstream of *GAL10* to make a unique intergenic region of 533 bp. Blue triangle indicates *GAL7* poly(A) site. Yellow triangle indicates *GAL10* poly(A) site. (B) Sequence specifics for primers used in long range PCR to generate the fused construct Fus (pYC10-7Fus). The blue sequence indicates coding strand of the *GAL7* ORF. Yellow sequence indicates the coding strand of the *GAL10* ORF. Sequence in lower case letters indicates what was deleted in the Fus construct. Green italicised letters indicate sequence, direction and location of primers used in long

pYC10-7Fus (Fus)

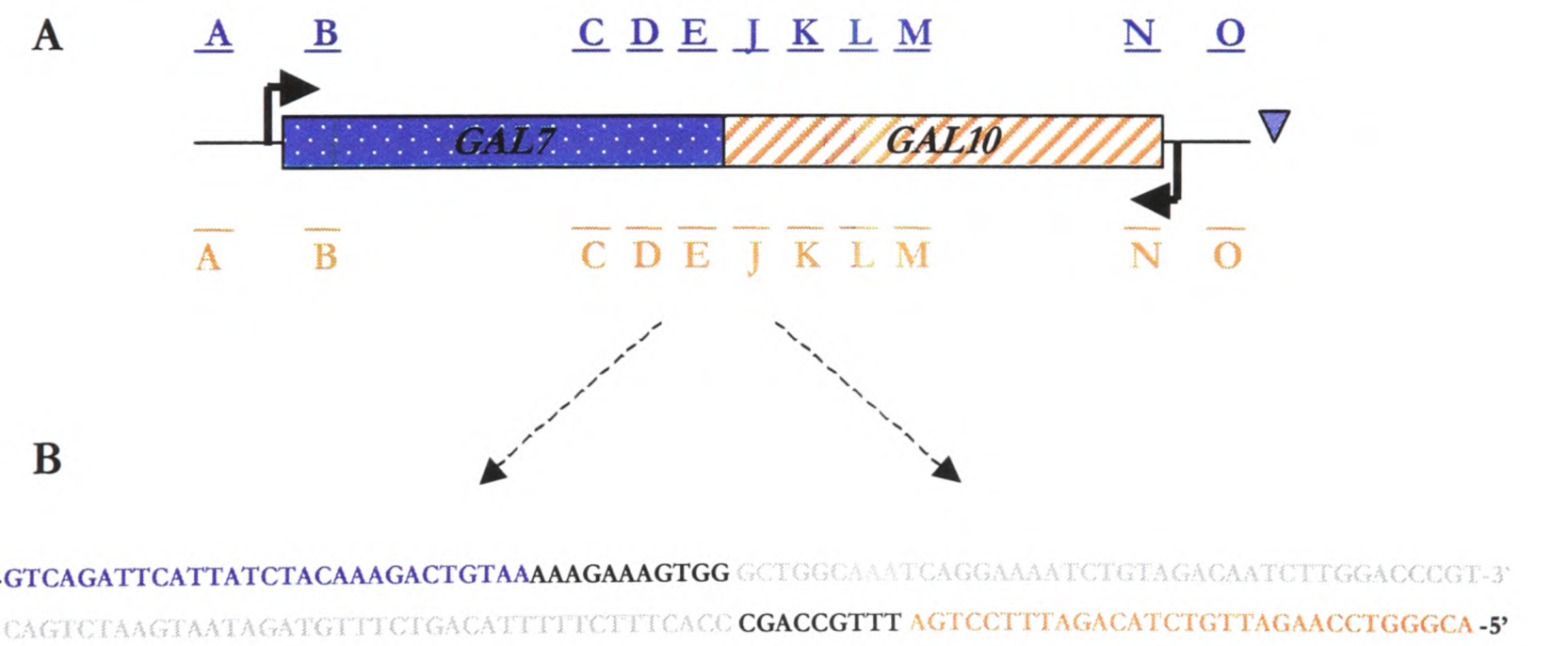


Figure I 2.2 (A) pYC10-7Fus(Fus). The direction of transcription for each gene is indicated with an arrow. Location of single stranded M13 probes are indicated (exact sequences identified in the Materials Section). Blue triangle indicates 190 bp *GAL7* poly(A) site placed upstream of *GAL10* in the vector backbone. (B) Sequence specifics for the junction of Fus. Blue sequence indicates sense strand of *GAL7* ORF. Yellow sequence indicates sense strand of *GAL10* ORF.

2.1.2 Nascent transcription from Fus construct

TRO analysis of Δ*GAL* yeast transformed with Fus showed dramatic transcriptional interference when these genes were simultaneously transcribed (Figure I 2.3). Quantitation, performed as before,

is seen in Figure I 2.4. For *GAL10*, transcription signals in Fus showed full levels at initiation but then dropped over probe M and more gradually to background levels from probe L to B. Similarly, the initiation level of *GAL7* transcription was similar to the level in Con but signals dropped at probe E. However, in *GAL7* the level of transcription continued at a significant level leading to a more gradual decline through the *GAL7* ORF tailing off beyond the region that would normally contain a poly(A) site in the WT construct. Therefore, it appears that transcriptional interference of convergent transcripts, or transcriptional collision, acts during transcriptional elongation rather than at initiation.

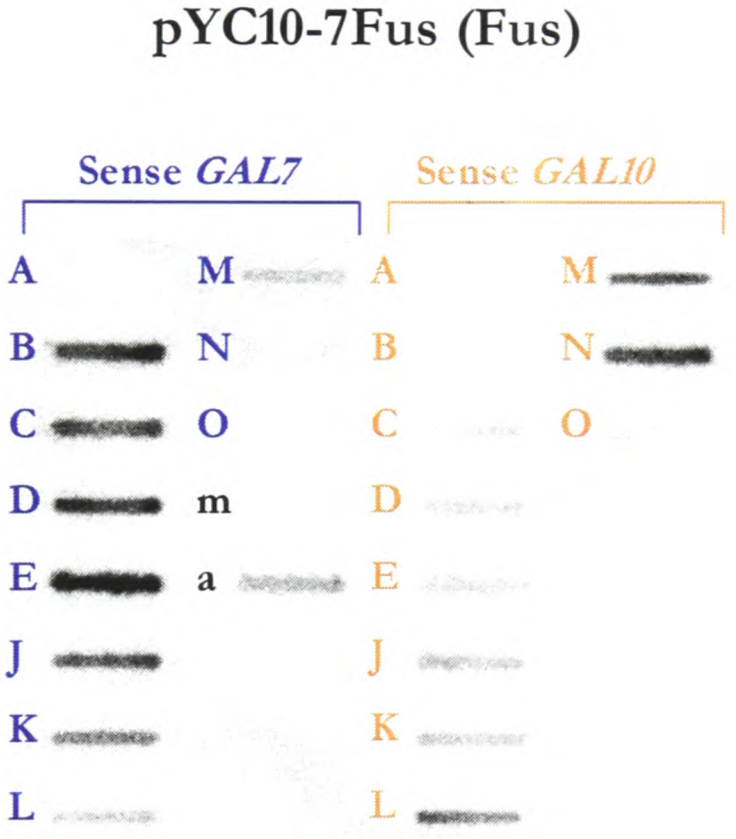


Figure I 2.3 TRO of pYC10-7Fus (Fus) transformed cells. Location of single stranded M13 probes are indicated in Figure I 2.2A. Probe 'm' is an M13 probe without an insert serving as a negative hybridisation control. Probe 'a' is an M13 probe with a fragment of the *S.cerevisiae* actin gene serving as a positive transcription control.

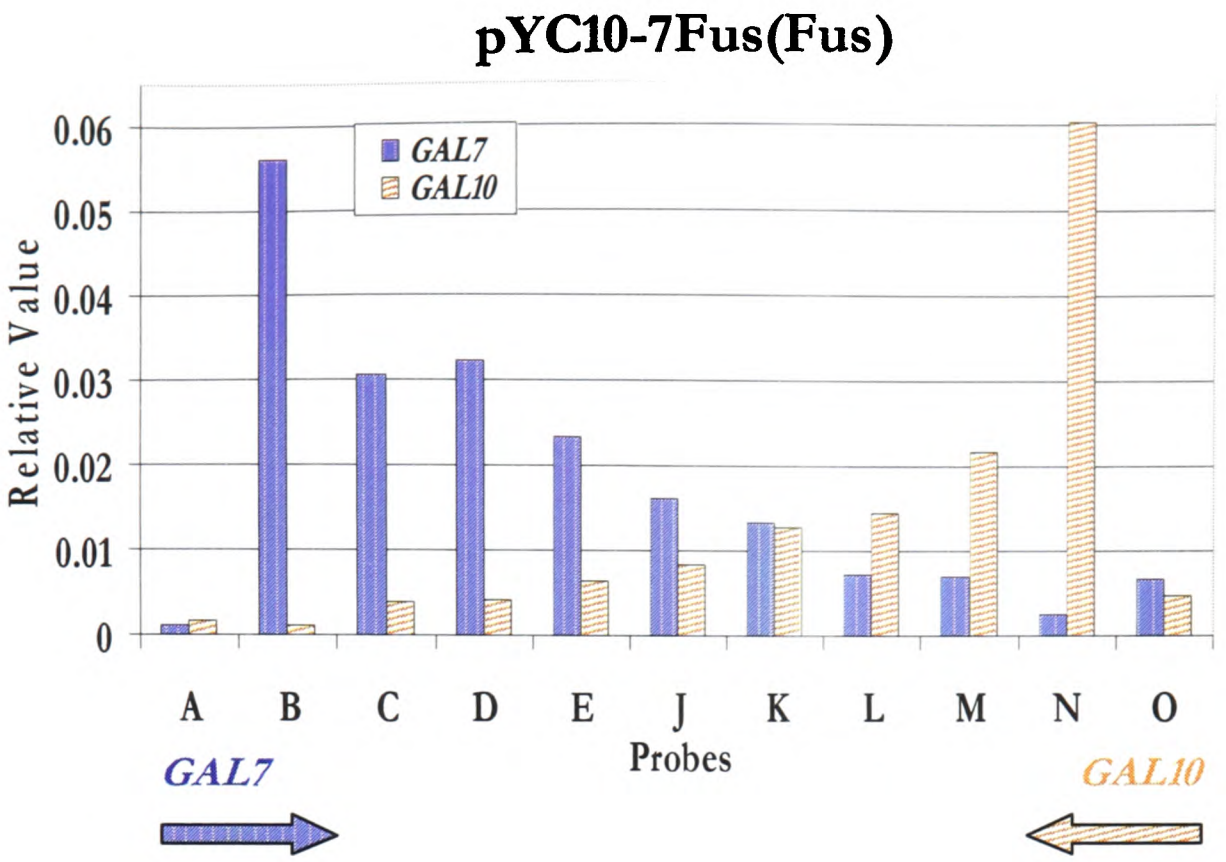


Figure I 2.4 Quantitation of data shown in Figure I 2.3 - TRO of Fus plasmid. Arrows indicate direction of transcription for each gene. Signals were quantified in a Phosphoimager (Molecular Dynamics). The values were corrected for background hybridisation and 'U' content in each probe. Signals were normalized to the actin positive transcription probe 'a' on each filter to allow direct comparison of transcription levels.

Steady state analysis (Northern blot) of *GAL10* and *GAL7* mRNA produced by Fus is entirely consistent with nascent TRO analysis. Comparing the level of steady state *GAL10* from Fus to the level in the Fus construct with *GAL7* transcription abolished (Fus-Δ7P, described in more detail below) show that mRNA levels are reduced by 20 fold (Figure I 2.5, compare *GAL10* in lane 3 and lane 4). Additionally, *GAL7* in Fus appears to decrease by 4 fold (Figure I 2.5, compare *GAL7* in lane 3 and lane 5) from a Fus construct with *GAL10* transcription eliminated (Fus-Δ10P; described in more detail below). From this it is clear that transcriptional collision of convergent transcripts results in a drastic reduction in gene expression.

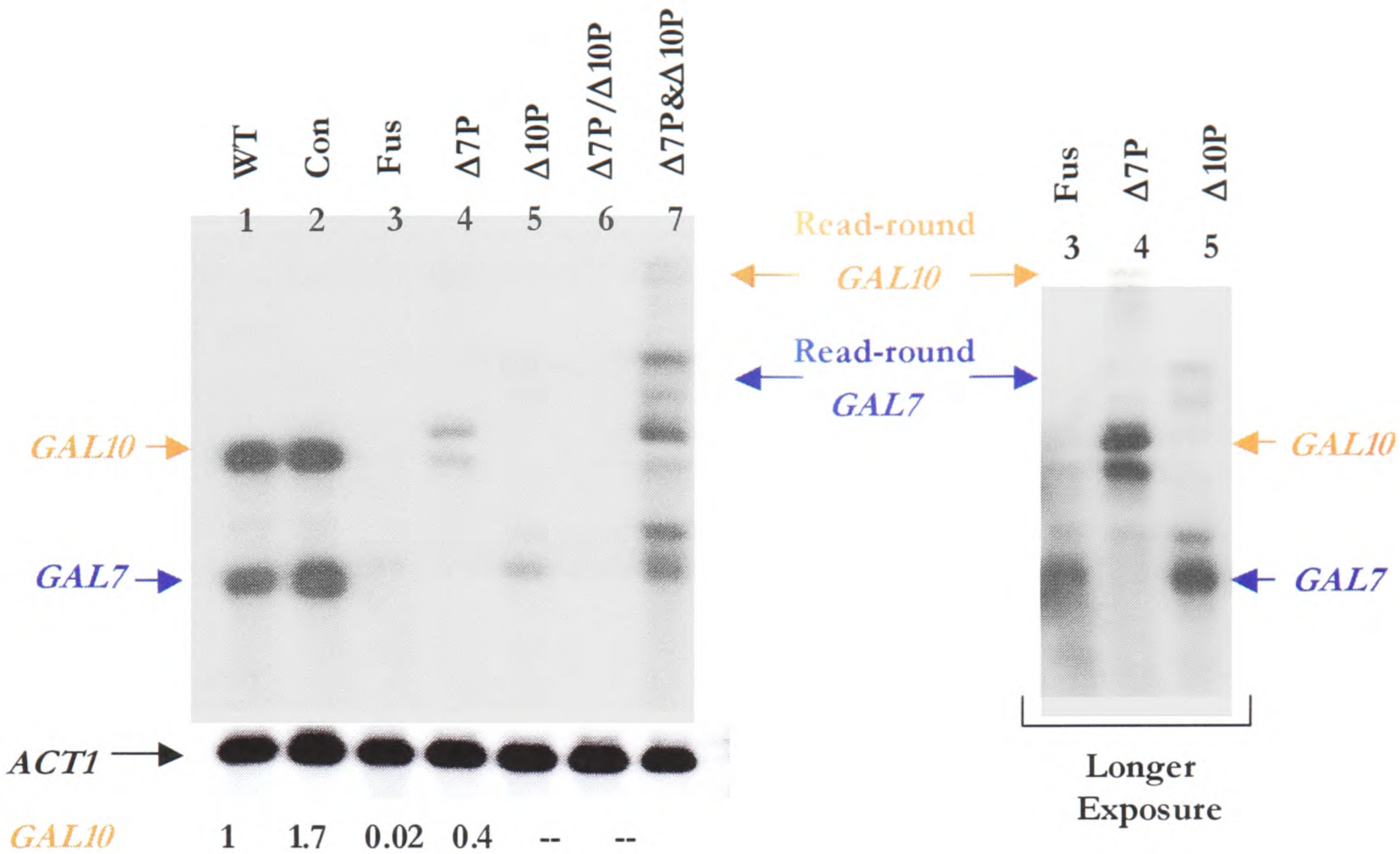


Figure I 2.5 Northern blot with total RNA from pYC10-7 (WT) (lane 1), pYC10-7Con (Con) (lane 2), pYC10-7Fus (Fus) (lane 3), Fus-Δ7P (lane 4), Fus-Δ10P (lane 5), Fus-Δ7P/Δ10P (lane 6), and Fus-Δ7P&Δ10P (lane 7). Probe was a 1.5 kb fragment generated by double restriction digest with *BglII* and *SalI* of the WT *GAL10-7* locus which contained 427 bp of 5' *GAL7* and 403 bp of 3' of *GAL10*. Filter was stripped and re-probed with *ACT1* probe as a loading control. The *ACT1* probe was a 567 bp fragment (277 –844 bp 3' of the *ACT1* start codon). Numbers shown beneath gel indicate quantitated levels of steady state *GAL10* and *GAL7* mRNA, corrected for actin and compared using steady state level of *GAL10* and *GAL7* mRNA from WT as 1.

2.2 Separate transcription from individual promoters

In order to demonstrate that the collision effect in Fus is the result of transcription on the same template, it was necessary to separate *GAL10* and *GAL7* transcription. Additionally, it is possible that the effect seen in Fus is due to the generation of antisense nascent transcripts resulting in RNA interference (RNAi) *in vivo* or could compete with the probes for hybridisation. To control for these events constructs that separated the transcription of each gene in the Fus plasmid were created and TRO was performed in various *in vivo* combinations.

2.2.2 Transcription of *GAL10* with *GAL7* promoter deletion (Fus-Δ7P)

A Fus construct that transcribed only *GAL10* was generated by deletion the *GAL7* TATA box. The plasmid pYC10-7FusΔTATA*GAL7* (Fus-Δ7P) (Figure I 2.6A), had a 9 bp deletion of the TATA

box 5'-GATATAAAA- 3'(*GAL7*: -179 to -188) in the Fus construct. The primers used for the long range PCR (ΔTATAa and ΔTATAb) are identified on the diagram in Figure I 2.6A and the sequence is given in Figure I 2.6B (in green). This small deletion had been shown previously (Greger and Proudfoot 1998) to abolish the induction of *GAL7* without affecting *GAL10* initiation.

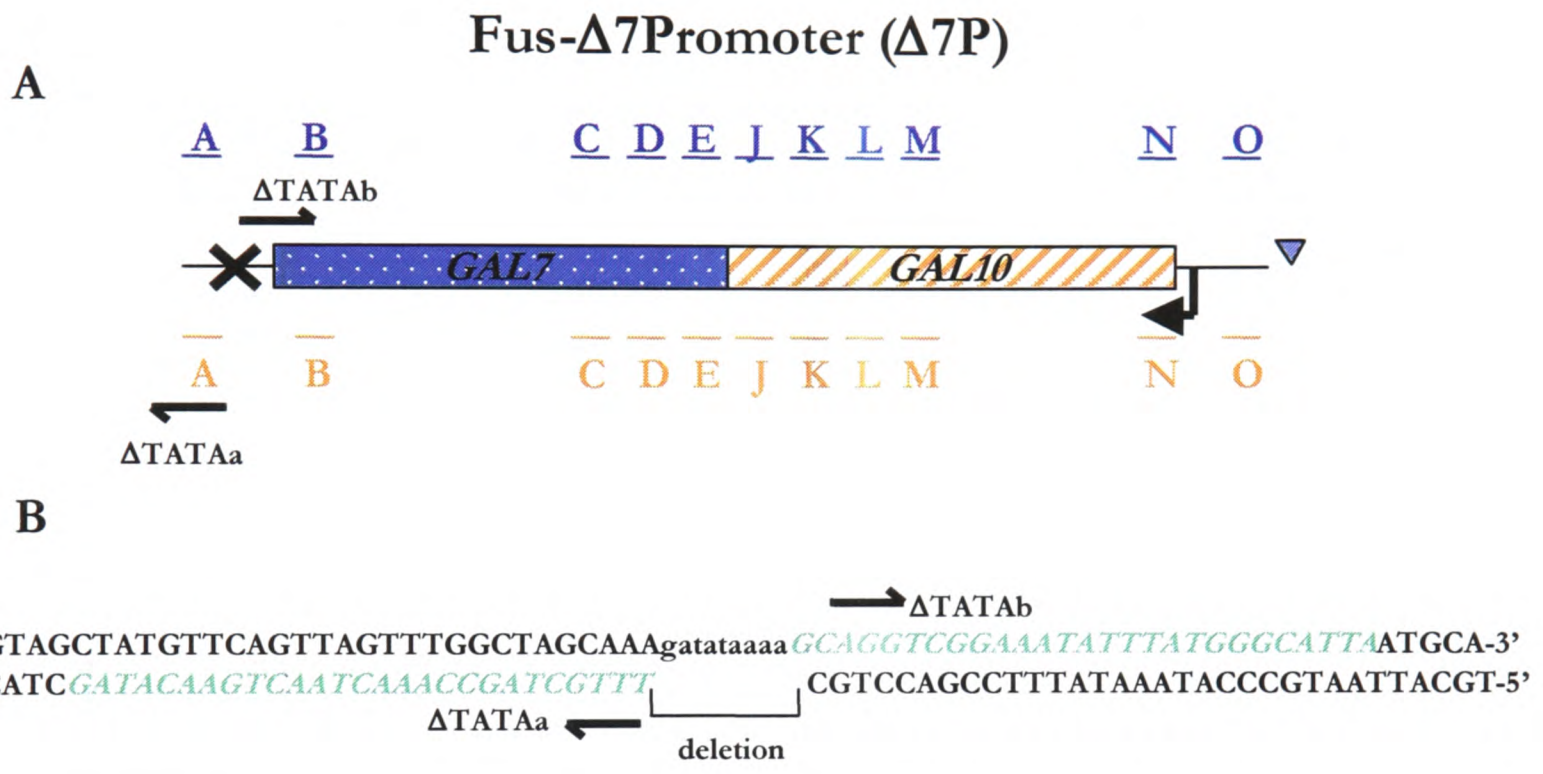


Figure I 2.6 (A) Fus-Δ7Promoter (Δ7P). Primers used in generating the ΔTATA *GAL7* construct ('ΔTATAa' and 'ΔTATAb') are labelled. The direction of transcription for *GAL10* is indicated with an arrow. Location of single stranded M13 probes are indicated. Blue triangle indicates 190 bp *GAL7* poly(A) site placed upstream of *GAL10* in the vector backbone (B) Sequence specifics for Fus-Δ7Promoter (Δ7P). Sequence in lower case letters indicates what was deleted in Δ7P. Green italicised letters indicate sequence, direction and location of primers used in long range PCR

The nascent transcription profile of Fus-Δ7P shows very low transcription of *GAL7* (Figure I 2.7). Additionally, the level of transcription detected by *GAL10* sense probes is much higher than detected previously over the length of the template up to probe A. Beyond probe A nascent transcription appears to terminate. This is in agreement with the steady state data that suggests cryptic poly(A) sites are used to generate a doublet of detectable *GAL10* mRNA bands (Figure I 2.5, lane 4). The longer exposure in Figure I 2.5 shows that there are additional larger bands that have the same doublet pattern, perhaps the result of transcription that reads around the entire plasmid. The decrease of the steady state *GAL10* transcripts from Fus-Δ7P compared to Con or WT presumably reflects the inefficiency of this cryptic mRNA processing. Quantitation of Fus-Δ7P membrane seen in Figure I 2.7 was performed as described previously and is shown in Figure I 2.10.

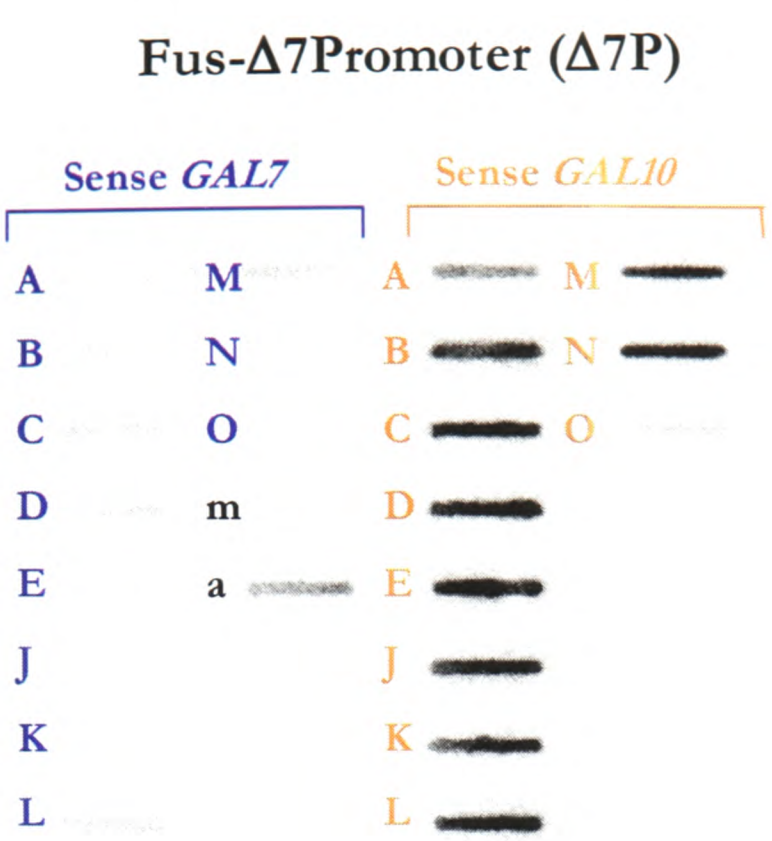


Figure I 2.7 TRO of Fus-Δ7Promoter (Δ7P) transformed cells. Diagram shows the arrangement of the *GAL10* and *GAL7* genes with direction of *GAL10* transcription indicated with an arrow and inactivated *GAL7* promoter is indicated by a cross. Blue triangle indicates 190 bp *GAL7* poly(A) site placed upstream of *GAL10* in the vector backbone. Location of single stranded M13 probes were indicated in Figure I 2.6. Probe 'm' is an M13 probe without an insert serving as a negative hybridisation control. Probe 'a' is an M13 probe with a fragment of the *S.cerevisiae* actin gene serving as a positive transcription control. Quantitation of this filter is shown in Figure I 2.10.

2.2.3 Transcription of *GAL7* with *GAL10* promoter deletion (Fus-Δ10P)

Similar analysis was done with a deletion of the UAS of the *GAL10* gene to generate a Fus construct that transcribed *GAL7* but not *GAL10* (Fus-Δ10P) (Figure I 2.8). The pYC10-7FusΔUAS*GAL10* (Fus-Δ10P) plasmid had the *EcoRI* fragment of Fus replaced with the *EcoRI* fragment from pΔ-UAS_{G10} plasmid described previously (Greger and Proudfoot 1998). This created a deletion in Fus that eliminates –136 to –414 of *GAL10*, removing all UAS sites and probe O (Figure I 2.8B). In this construct there was an up-regulation of transcription across the *GAL7* gene and the near elimination of transcripts from the *GAL10* promoter (Figure I 2.9 and Figure I 2.10).

A



B

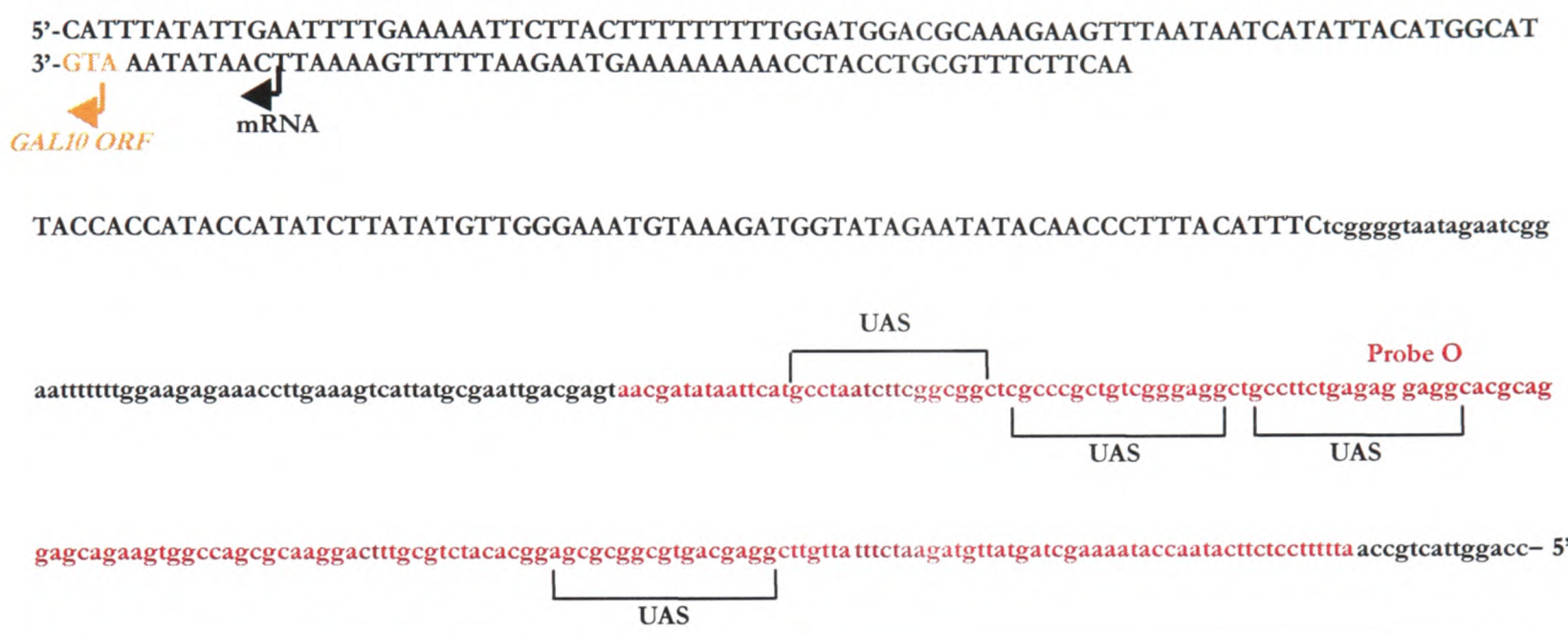


Figure I 2.8 (A) Fus-Δ10Promoter (Δ10P). The direction of transcription for *GAL7* is indicated with an arrow. Location of single stranded M13 probes are indicated. Blue triangle indicates 190 bp *GAL7* poly(A) site placed upstream of *GAL10* in the vector backbone. (B) Sequence specifics for Fus-Δ10Promoter (Δ10P). *GAL10* UAS sequences are indicated. Sequence in lower case letters indicates what was deleted in Δ10P. Yellow sequence indicates *GAL10* ORF. Red sequence indicates the location of probe O.

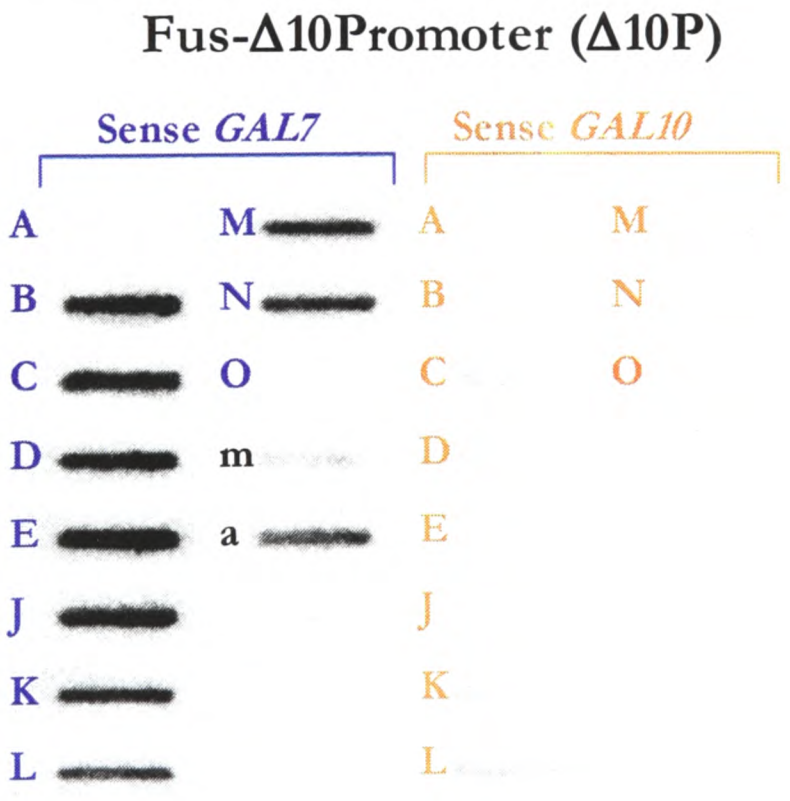


Figure I 2.9 TRO of Fus-Δ10Promoter (Δ10P) transformed cells. Location of single stranded M13 probes are indicated in Figure I 2.8A. Probe 'm' is an M13 probe without an insert serving as a negative hybridisation control. Probe 'a' is an M13 probe with a fragment of the *S.cerevisiae* actin gene serving as a positive transcription control.

Polymerase density over *GAL7* in Fus-Δ10P appears relatively consistent across the length of the probes with the exception of probe O, which was deleted with the *GAL10* promoter region. Steady state analysis of Fus-Δ10P shows that there are low levels of *GAL7* mRNA, most likely because of the inefficient cryptic poly(A) signals that are being used (Figure I 2.5, lane 5). Read-round bands that probably result from utilization of the same cryptic poly(A) sites the second time around the plasmid can also be seen this construct (Figure I 2.5, longer exposure lanes 4 and 5). This suggests that the *GAL7* poly(A) site inserted upstream of *GAL10* did not function as had been expected. Quantitation as described previously was done for Figure I 2.9 and is shown in Figure I 2.10.

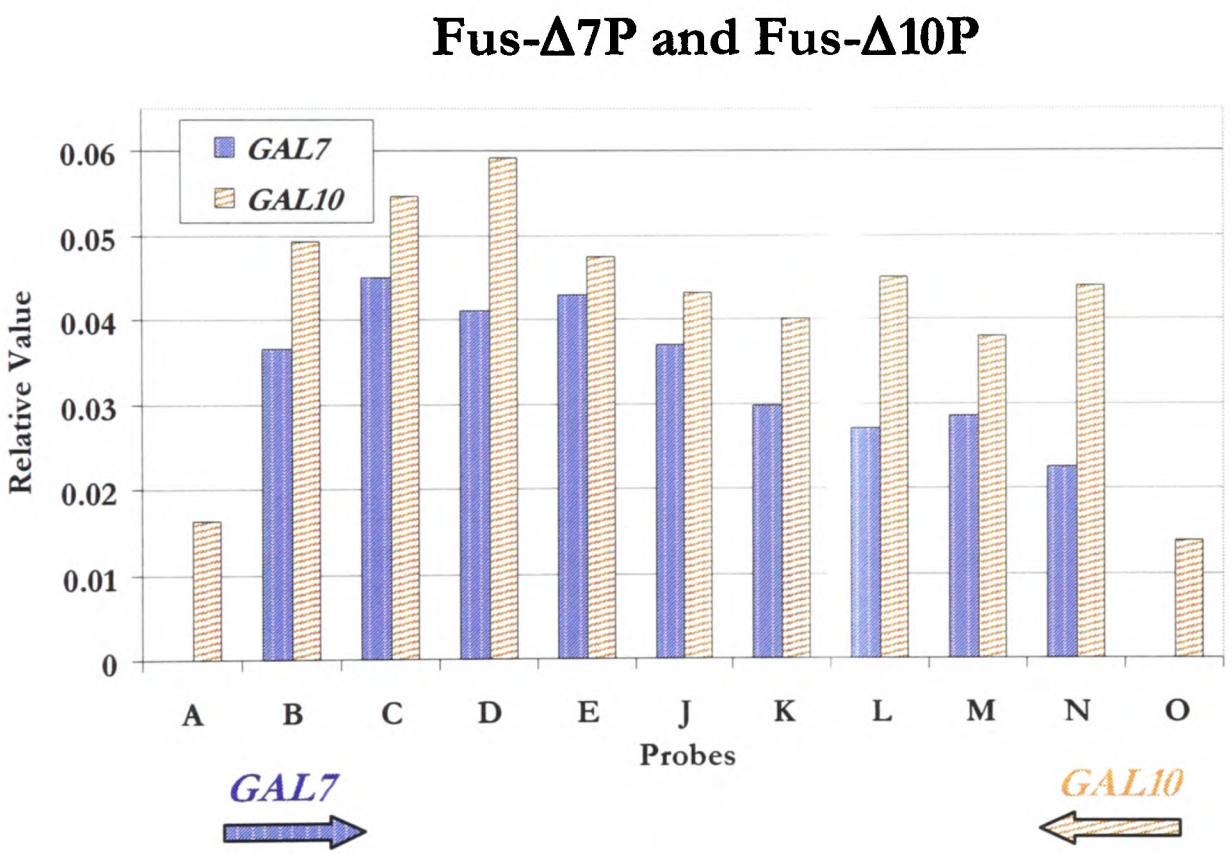


Figure I 2.10 Quantitation of data shown in Figure I 2.7 and Figure I 2.9 - TRO of Δ7P and Δ10P plasmid. Arrows indicate direction of transcription for each gene. Signals were quantified in a Phosphoimager (Molecular Dynamics). The values were corrected for background hybridisation and 'U' content in each probe. Signals were normalized to the actin positive transcription probe 'a' on each filter to allow direct comparison of transcription levels.

2.2.4 Transcription with both *GAL10* and *GAL7* promoters deleted (Fus-Δ7P/Δ10P)

As an additional control, a plasmid was created in which both promoters were deleted (Fus-Δ7P/Δ10P). As expected, the nascent (Figure I 2.11) and steady state (Figure I 2.5, lane 6) data demonstrates that the deletion of both promoters in the same construct results in very little transcription from either promoter. The low level of transcription that is seen is most likely the result of basal transcription that is not the result of galactose induction. If compared to the actin

positive control in that nascent (Figure I 2.11) and the level of steady state transcripts seen in Figure I 2.5 (lane 6) it is evident that the genes are effectively not generating transcripts. This also serves as a control showing that the transcripts generated from the introduced plasmids are the only source of RNA hybridising to the probes.

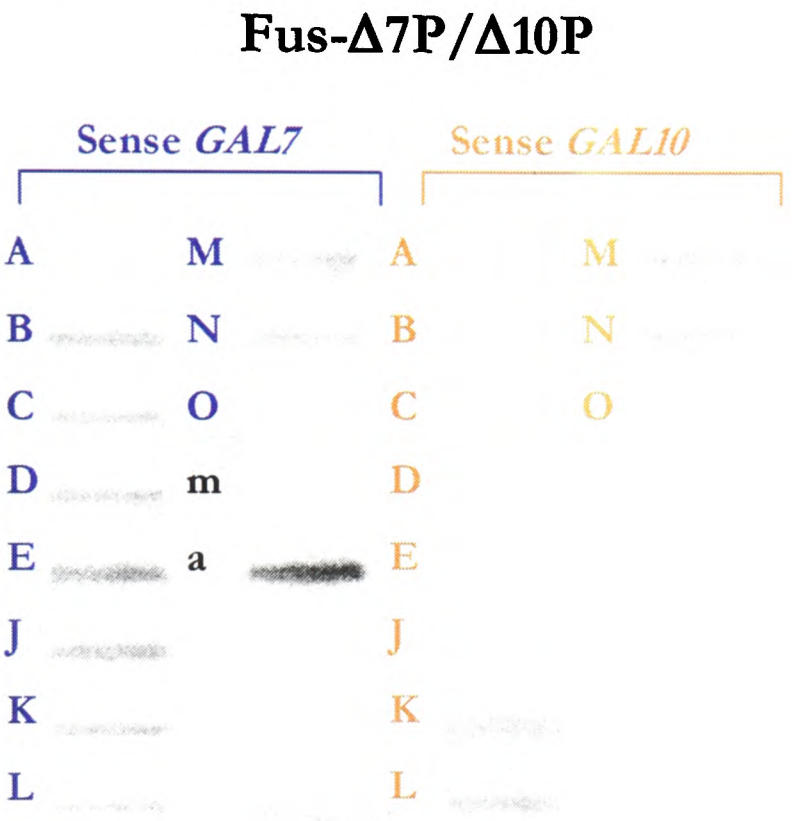


Figure I 2.11 TRO of Fus-Δ7P/Δ10P transformed cells. Plasmid was generated in the same manner as the individual promoter deletions but sequentially on the same plasmid. Location of single stranded M13 probes is indicated in Figure I 2.8A. Probe 'm' is an M13 probe without an insert serving as a negative hybridisation control. Probe 'a' is an M13 probe with a fragment of the *S.cerevisiae* actin gene serving as a positive transcription control.

2.2.5 Separate *GAL10* and *GAL7* transcription in Fus (Fus-Δ7P & Δ10P)

All of the plasmids described so far are centromeric and consequently contain 1-2 copies per cell. In order to verify that apparent collision effects observed were in fact due to transcription on a single template (in *cis*) the transcription on separate templates in the same cell (in *trans*) was compared. This was achieved by transforming the same cell with the both of the plasmids separately transcribing *GAL10* and *GAL7*.

TRO analysis of the Δ*GAL* strain sequentially co-transformed with Fus-Δ7P and Fus-Δ10P was performed. Critically, this should demonstrate whether the TRO profile of separate transcription is different from that which is observed in Fus. At both the nascent (Figure I 2.12) and steady state levels (Figure I 2.5, lane 7) the termination profiles appear to be the sum of both Fus-Δ7P and Fus-

Δ10P plasmids. These data support our transcriptional collision model for interference, as it requires the *GAL10* and *GAL7* genes to be present in a *cis* arrangement. Trivial explanations for the observed inhibition of *GAL10* and *GAL7* transcription through antisense effects are effectively excluded.

Fus-Δ7P&Δ10P

(separate plasmids in the same cell)

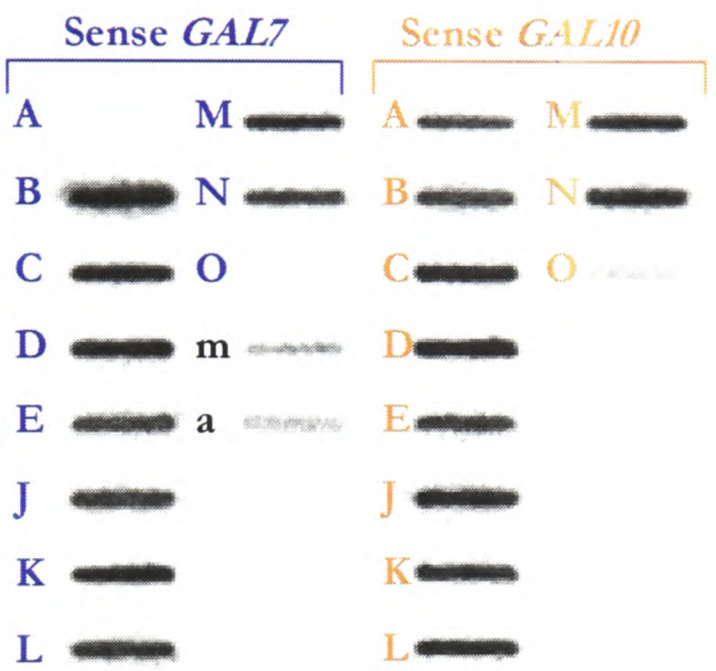


Figure I 2.12 TRO of Fus-Δ7P&Δ10P transformed cells. Location of single stranded M13 probes are indicated in Figure I 2.8A. Probe 'm' is an M13 probe without an insert serving as a negative hybridisation control. Probe 'a' is an M13 probe with a fragment of the *S.cerevisiae* actin gene serving as a positive transcription control.

Discussion

The aim of these studies was to investigate the effect that convergent transcription on a single template has on the transcriptional activity of Pol II in *S. cerevisiae.* In the presence of intact mRNA 3' end processing signals the convergently arranged *GAL7* and *GAL10* genes generate functional transcripts that appear to terminate and be processed in a similar manner to the WT transcripts. However, once the 3' end processing and termination signals are deleted there is a drastic effect on the level of transcriptional processivity in both genes, a phenomenon referred to as transcriptional collision.

During convergent transcription both *GAL7* and *GAL10* initiation appear to increase. In the Fus construct, the level of initiation for *GAL7* continues to be elevated above the level seen in Con or the WT arrangement (compare *GAL7* probe B of Figure I 2.4 with *GAL7* probe B of Figure I 1.4

and Figure I 1.9). *GAL10* initiation is lower in Fus than seen in Con but about the same as what is seen in WT (compare *GAL10* probe N of Figure I 2.4 with *GAL10* probe N of Figure I 1.4 and Figure I 1.9). Interestingly, the level of initiation from both genes in the Fus construct is higher than that which is seen when the genes are transcribed separately (Fus-Δ7P and Fus-Δ10P) which suggests that there is something about the concurrent transcription of convergent genes that can stimulate initiation.

While the precise biological cause of the transcriptional collision effect still remains to be identified, the impact appears to occur in the elongation stage of transcription. One possible explanation is that as transcription proceeds, the polymerases collide and cause one or both to release from the template. Alternatively, a polymerase complex, when confronted with one polymerase on the opposite strand, could stall and be unable to advance further. This could be due to a direct physical impediment to the transcription machinery or an indirect effect caused by super-coiling changes to the DNA template during transcription (Liu and Wang 1987; Giaever *et al.* 1988; Giaever and Wang 1988; Tsao *et al.* 1989; Dunaway and Ostrander 1993). In the latter scenario, a transcription bubble progresses along a template creating positive super-coils. If these positive super-coils were generated from both strands simultaneously, it would create a region of hyper-super-coiling that would be predicted to prevent further advancement in either direction on the template. This theory would also help explain the increased initiation seen when the genes are concurrently transcribed as negative supercoils have been shown to increase transcriptional initiation (Liu and Wang 1987).

It is clear that the transcriptional collision seen in this investigation occurred in *cis* and was alleviated when the opposing polymerase was prevented from transcribing simultaneously. Separate expression of *GAL10* (Fus-Δ7P) and *GAL7* (Fus-Δ10P) in the same cell (Fus-Δ7P&Δ10P) (see Figure I 2.12) showed no interference despite the generation of antisense transcripts. These data demonstrate that simultaneous transcription of *GAL10* and *GAL7* on the same template is occurring. Additionally, this argues against any RNAi effects in this *S.cerevisiae* strain. Interestingly, previous research investigating RNAi in *S.cerevisiae* has detected RNAi in *cis* but not in *trans* (Atkins *et al.* 1994; Cogoni

and Macino 2000; Hammond *et al.* 2001). It seems plausible that these previously identified *cis* RNAi effects may actually be due to transcriptional collision on the same template rather than true RNAi.

Transcriptional collision between convergent genes in budding yeast

Elizabeth M. Prescott and Nick J. Proudfoot*

Sir William Dunn School of Pathology, University of Oxford, South Parks Road, Oxford OX1 3RE, United Kingdom

Communicated by Stanley Fields, University of Washington, Seattle, WA, May 6, 2002 (received for review February 21, 2001)

Transcriptional interference between genes and the regulatory elements of simple eukaryotes such as *Saccharomyces cerevisiae* is an unavoidable consequence of their compressed genetic arrangement. We have shown previously that with the tandem arranged genes *GAL10* and *GAL7*, inefficient transcriptional termination of the upstream gene inhibits initiation of transcription on the downstream gene. We now show that transcriptional interference can occur also with *S. cerevisiae* RNA polymerase II genes arranged convergently. We demonstrate that when the *GAL10* and *GAL7* genes are rearranged in a convergent orientation, transcriptional initiation occurs at full levels. However, as soon as the two transcripts begin to overlap, elongation is restricted, resulting in a severe reduction in steady-state mRNA accumulation. This effect is observed only in cis arrangement, arguing against RNA-interference effects acting on the potential generation of antisense transcripts. These data reinforce the necessity of separating adjacent RNA polymerase II transcription units by efficient termination signals.

The genomic sequence of *Saccharomyces cerevisiae* is highly compact, with over 70% represented by ORFs (1). This high gene density is largely a result of the short intergenic regions that average 618 bp apart for diverging promoters and 517 bp for tandem (head-to-tail) genes using promoter/terminator combinations (2). Convergent genes average only 326 bp apart (2), with several examples of less than 300 bp (3–6). Such a limited intergenic sequence would allow only ≈150 bp per gene for proper termination outside of the neighboring ORF, which is less sequence than the few termination regions defined thus far in *S. cerevisiae* genes (7–9). Analysis of adjacent gene sets in *S. cerevisiae* determined that less than 6% of coexpressed, adjacent genes are in a convergent orientation, with a bias to divergent transcription (10). These statistics suggest that there may be evolutionary pressure to select against convergently arranged, cotranscribed genes (2).

Transcription of RNA polymerase II (PolII) in *S. cerevisiae* has been studied extensively, with both the mechanism of initiation and elongation recently described in atomic detail (11, 12). The mechanism for termination and coupled mRNA 3′-end processing is understood less completely. However, mutations in cleavage/polyadenylation factors prevent properly regulated termination, highlighting the interdependence of these functions (13). The polyadenylation signals in budding yeast genes are less sequence-specific than in mammalian systems. In mammals, an almost invariant AAUAAA hexonucleotide is located 15–20 nt upstream of the cleavage site, which works in conjunction with a less well defined GU- or U-rich element downstream to promote proper 3′-end formation. In *S. cerevisiae*, a more degenerate poly(A) signal is assisted by an upstream efficiency element (consensus UAUAUA) to promote 3′-end formation. Additionally, at 15–30 nt downstream of the cleavage site a positioning element (consensus AAAAAA or AAUAAA) appears to direct cleavage at the preferred site Y(A)n (see ref. 14 for review).

With such a densely packed genome it is not surprising that interference has been detected between transcribing genes (15–17). Transcriptional interference (TI) may occur when an elongating polymerase fails to terminate properly and thus reads into a cotranscribed downstream transcription unit, reducing promoter activity. TI also has been documented to affect the proper use of both ARS and CEN elements and in so doing inhibits either replication or chromosomal segregation in mitosis (18–20). Sequence elements that mediate enhanced meiotic gene conversion also are susceptible to TI (21). At the nascent level, the *Schizosaccharomyces pombe* genes *nmt1* and *nmt2* were shown to transcribe into two poorly defined convergent downstream ORFs (22) with little detectable interference. These *nmt* genes both are induced to transcribe at high levels, whereas the convergent genes *avn2* and *gut2* are ubiquitously expressed at much lower levels. The possibility that these convergent genes are expressed separately remains to be determined. In contrast, steady-state analysis in *S. cerevisiae* has shown that the elimination or constitutive activation of one promoter in a convergent pair of genes often affects the steady-state levels of mRNA generated from the opposing gene (4, 15).

In the tightly regulated metabolic pathway of galactose fermentation, the tandem *GAL10* and *GAL7* genes are affected by TI (16). A fraction of polymerases elongate beyond *GAL10* and have an inhibitory effect on the initiation of the downstream gene, *GAL7*. Furthermore mutations of the *GAL10* poly(A) signal prevent all PolII termination and thus have a more drastic effect on the levels of *GAL7* gene expression. Interestingly, *GAL7* expression is recoverable by overexpression of Gal4p (23). Presumably TI blocks the access of the Gal4p transcription factor to the *GAL7* promoter and in so doing prevents transcriptional activation. The coinduction of these genes by galactose makes an ideal system to study TI in different gene orientations.

In these studies we have manipulated the orientation of the *GAL10* and *GAL7* genes to model the effect of convergent, cotranscribed genes. By eliminating the intergenic region of these convergent genes, we have investigated the effects of improper termination on convergent transcription. Our results graphically illustrate the need for efficient PolII termination. In its absence, inhibition of both genes occurs by a mechanism we refer to as transcriptional collision.

Materials and Methods

S. cerevisiae Strain. All experiments use ΔGAL *S. cerevisiae* strain 6-1/13 (MATα *his*3Δ200 *met*15Δ0 *trp*Δ63 *ura*3Δ0 gal10-7::URA3) (16). The strain was grown on synthetic complete (SC) medium lacking uracil and supplemented with 2% raffinose.

Plasmids. pYC10-7 (wild type, WT) was based on Ycplac22 (16). pYC10-7Con (Con) was created by *Xba*I digestion of pYC10-7 and partial digestion with *Eco*RV. The 7,658-bp product (lacking *GAL7*) was isolated. A *GAL7* fragment was generated by PCR with a unique *Xho*I site 103 bp after the poly(A) site (see *Primers*).

Abbreviations: PolII, RNA polymerase II; TI, transcriptional interference; WT, wild type; TRO, transcriptional run-on; RNAi, RNA interference.

*To whom reprint requests should be addressed. E-mail: nicholas.proudfoot@path.ox.ac.uk.

www.pnas.org/cgi/doi/10.1073/pnas.132270899

:10-7Fus (Fus) was created by ligating long-range PCR
ts generated from Con by using *Pfu* DNA polymerase
igene) and kinased primers Fuse7 and Fuse10 located
/ inside the ORFs for both genes (see *Primers*). The
GAL7 poly(A) site fragment (*GAL7*, 1,118–1,327) in the
d orientation for *GAL7* was inserted into the *Sac*I site of
:tor backbone upstream of the *GAL10* promoter (−442).
:10-7FusΔTATA*GAL7* (Fus-Δ7P) had a 9-bp deletion of
.TA box, 5′-ATATAAAA- 3′ (*GAL7*, −179 to −188), in
s construct generated by long-range PCR with primer
Aa and ΔTATAb (see *Primers*).
:10-7FusΔUAS*GAL10* (Fus-Δ10P) had the *Eco*RI frag-
of Fus replaced with the *Eco*RI fragment from the
.S_{G10} plasmid (16), which created Fus with a deletion that
ates −136 to −414 of *GAL10*, removing all UAS sites.
:10-7FusΔTATA*GAL7*/ΔUAS*GAL10* (Fus-Δ7P/Δ10P)
enerated by replacing the *Eco*RI fragment from the
7P plasmid with the *Eco*RI fragment from Fus-Δ10P.
:10-7FusΔTATA*GAL7*&ΔUAS*GAL10* (Fus-Δ7P&Δ10P)
:otransformation of Fus-Δ7P and Fus-Δ10P. See *Growth*
ransformation of S. cerevisiae Strains below.

:. The primer used were: Con *GAL7* insert, GAL7+100
GAGACTTACAAGCTGCATTGTATTC) and Up*GAL7*
.GTTCATTTCGTTACTTTTG); Fus construct, Fuse7
CTTTCTTTTTACAGTCTTTGTAGATAATG), and
) (GCTGGCAAATCAGGAAAATCTGTAGAC); Δ7P
uct, Δ7TATAa (CTTTGCTAGCCAAACTAACTGAA-
.G), and Δ7TATAb (GCAGGTCGGAAATATTTAT-
:ATTA).

and Transformation of *S. cerevisiae* Strains. Strain 6-1/13
:own on SC medium lacking uracil with 2% raffinose
a). Strains transformed with the plasmids were plated onto
ates lacking uracil and tryptophan (24). Fus-Δ7P&Δ10P
:eated by transforming 6-1/13 carrying the Fus-Δ7P plas-
ith Fus-Δ10P plasmid and selecting on URA^-, TRP^- SC
with 2% galactose (Sigma) and 20 μg/ml ethidium
de. All growth was with URA^-, TRP^- SC medium, and
lactose with this strain.
ore transcriptional run-on (TRO) analysis, strains (except
nd Fus-Δ7P&-Δ10P) were grown on URA^-, TRP^- SC
m and 2% raffinose until early log phase and then induced
% galactose for 30–60 min. WT was grown in URA^- SC
m and 2% raffinose. Fus-Δ7P&-Δ10P were grown on
$^-$, TRP^- SC medium, and 2% galactose.

robe Constructs. Single-stranded DNA probes were made
:able 1) by cloning PCR-generated fragments (using *Pfu*
polymerase, Stratagene) from Con into the double-
ed *Xba*I and *Eco*RI sites of M13mp18 and M13mp19,
generated probes for the sense strands of both *GAL7* and
0. The M13 control probe (m) contained no insert. The
/e control actin probe (a) had a 567-nt fragment containing
44 nt 3′ of the *ACT1* start codon inserted into the *Hinc*II
f M13mp18.

nalysis. TRO analysis was performed as described (25)
t that 100-ml cultures grown to an OD_{600} of ≈0.1 were
ed with 2% galactose. Additionally, the TRO reaction was
ed to proceed for 5 min and then washed briefly with AE
· (50 mM NaAc/10 mM EDTA, pH 5.0). All TRO analyses
reproduced in at least three independent experiments.

ern Blot Analysis. Total RNA was analyzed (16). The North-
:obe was a 1.5-kb fragment generated from the WT plasmid
'II and *Sal*I digestion, which contains 427 bp of 5′ of *GAL7*
03 bp of 3′ of *GAL10*. The *ACT1* probe was a 567-bp

Table 1. TRO M13 probes

Probe	Location	Size, nt
A	*GAL7*: −401/−212	189
B	*GAL7*: 73/252	179
C	*GAL7*: 638/793	155
D	*GAL7*: 800/944	144
E	*GAL7*: 938/1,119	181
F	*GAL7*: 1,110/1,294	184
G	*GAL7*: 1,184/1,354	170
H	*GAL7*: 1,266/1,362	177
	GAL10: 2,406/2,325	
I	*GAL10*: 2,245/2,055	190
J	*GAL10*: 2,070/1,884	186
K	*GAL10*: 1,885/1,727	158
L	*GAL10*: 1,720/1,572	148
M	*GAL10*: 1,537/1,379	158
N	*GAL10*: 330/152	178
O	*GAL10*: −187/−371	184

fragment (277–844 bp 3′ of the *ACT1* start codon) from the *ACT1* gene of *S. cerevisiae*.

Results

Transcription of the *GAL10* and *GAL7* Genes Arranged in Convergent Orientation. The *GAL10* and *GAL7* genes are organized naturally in a tandem arrangement on chromosome II. A previous study determined the transcription profile for nascent transcription in these genes for the WT conformation. These results were obtained by TRO analysis of a centromeric plasmid containing the *GAL10-GAL7* genes (pYC10-7) transformed into the ΔGAL strain (16). This TRO analysis is repeated in Fig. 1*A*. As indicated, higher nascent transcription is seen over the *GAL10* than *GAL7* gene. Also, most transcription fails to terminate beyond the *GAL10* poly(A) signals and reads into and accumulates over the *GAL7* promoter. In the previous study (16) analysis of a plasmid containing the *GAL10-GAL7* genes with the *GAL7* promoter inactivated revealed that most *GAL10* transcription terminates upstream of the *GAL7* transcription start site. Consequently, run-on signals over the *GAL7* gene shown in Fig. 1*A* are likely to derive from transcription initiation on the *GAL7* promoter.

To investigate the effect of concurrent, convergent transcription, we inverted the *GAL7* gene to generate pYC10-7Con (Con, Fig. 1*B*). The natural polyadenylation sites and a portion of the *GAL10-7* intergenic region were maintained. The intergenic region between *GAL10* and *GAL7* (from the 3′ end of the *GAL10* ORF to the 5′ end of the *GAL7* ORF) is 725 bp including the 280-bp promoter region of *GAL7*. In Con, the intergenic region is 533 bp from the 3′ end of the *GAL10* ORF to the 3′ end of the *GAL7* ORF, which includes 240 bp beyond the 3′ end of the *GAL7* ORF and 293 bp downstream of the 3′ end of the *GAL10* ORF.

TRO analysis of the Δ*GAL* yeast strain transformed with Con yields a nascent termination profile similar to the WT situation with a few notable exceptions. Transcription in the *GAL10* and *GAL7* genes appeared to terminate over the same probes regardless of orientation. However, the level of read-through transcription from *GAL10* decreased in Con (compare the *GAL10* probe G of Con with the *GAL7* probe A of WT). *GAL7* transcription terminated shortly after the WT poly(A) site in both orientations. Additionally, the levels of transcription appeared higher for both *GAL10* and *GAL7* genes in Con. Compare the levels of *GAL10* and *GAL7* nascent transcripts to the endogenous actin probe (a) for WT and Con plasmids. In particular for *GAL7*, there was a 2-fold increase in the level of

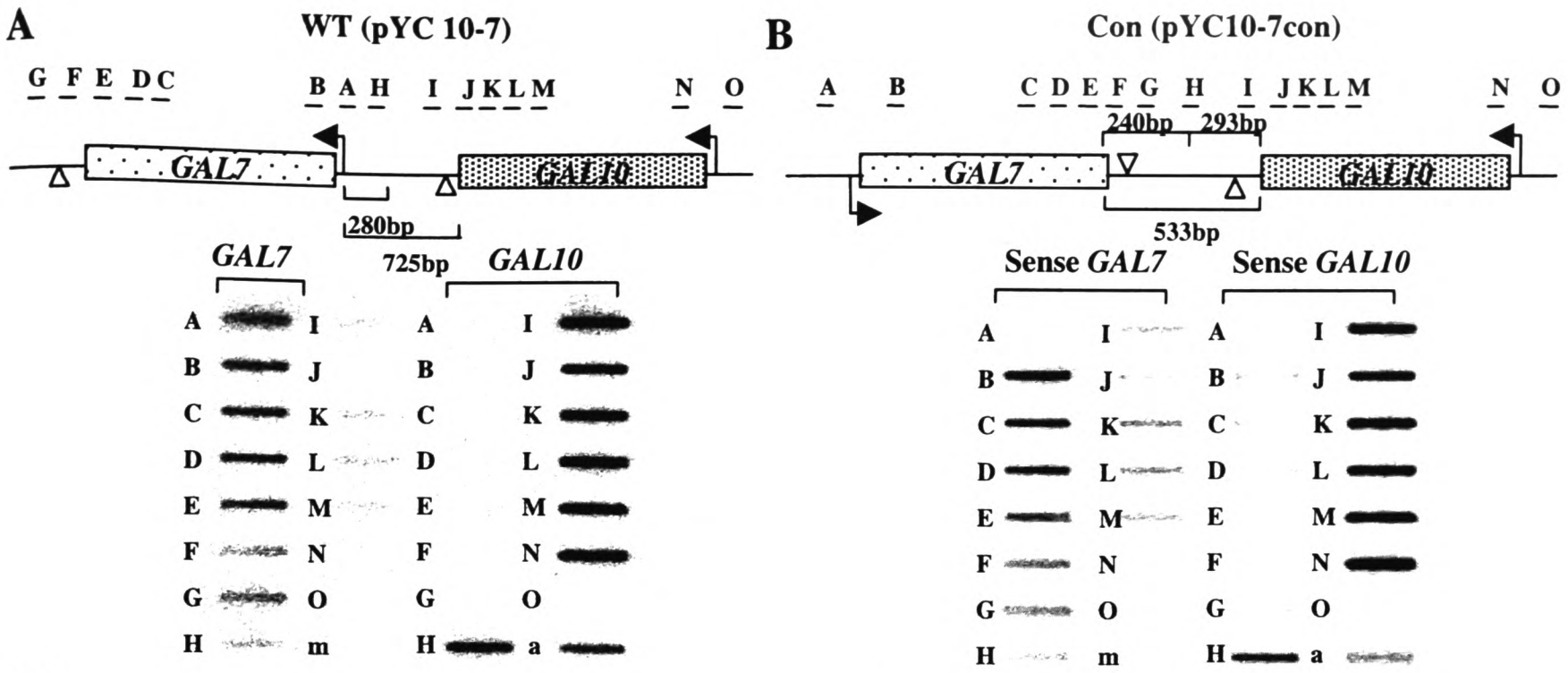


Fig. 1. TRO analysis of transformed *GAL10-7* plasmids. (*A*) TRO analysis of pYC10-7 (WT) transformed cells. The diagram shows the tandem arrangement of the *GAL10* and *GAL7* genes, with the direction of transcription indicated with an arrow. The position of the single-stranded M13 probes used are shown and are as described in Table 1. Probe m is an M13 probe without an insert serving as a negative hybridization control. Probe a is an M13 probe with a fragment of the *S. cerevisiae* actin gene serving as a positive transcription control. The distances between the ORFs of *GAL10* and *GAL7* in the WT orientation are labeled. Open arrowheads indicate polyadenylation sites. The *GAL7* promoter region is indicated also. (*B*) TRO analysis of pYC10-7Con (Con)-transformed cells. The diagram shows convergent orientation of *GAL10* and *GAL7*, with the intergenic region's distance labeled. This construct includes 240 bp downstream of *GAL7* and 293 bp downstream of *GAL10* to make a hybrid intergenic region. Sense and antisense single-stranded M13 DNA probes were selected as appropriate. (*C*) Quantitation of the data shown in *A* and *B*. Signals were quantified in a PhosphorImager (Molecular Dynamics). The values were corrected for background hybridization and U content in each probe. Signals were normalized to the actin positive transcription probe a on each filter to allow direct comparison of transcription levels.

nascent transcription from *GAL7*, which may be the result of alleviation of promoter occlusion occurring from read-through transcription beyond the *GAL10* poly(A) site into the *GAL7* promoter. The relative increase in *GAL7* gene expression is reflected by the steady-state mRNA analysis shown in Fig. 2 (lanes 1 and 2), where *GAL7* mRNA levels increase 2.3-fold with the Con construct. However, we also note that *GAL10* mRNA levels increase 1.7-fold in the Con construct as compared with WT. In this case TI is unlikely to be important, because no TRO signal is detectable over the *GAL10* promoter (probe O) in WT. It therefore is possible that other factors such as supercoiling effects may be involved.

Removal of Termination Signals Promotes Transcriptional Collision in Convergent Transcription. To test the effect of TI in the convergent orientation, the entire intergenic region was deleted from Con to make the plasmid pYC10-7Fus (Fus), effectively fusing together the *GAL7* and *GAL10* genes. The deleted *GAL7* poly(A) site was reinserted into the vector backbone upstream of *GAL10* in the forward orientation for *GAL7*, because this poly(A) signal is known to function bidirectionally. We anticipated that it would act to polyadenylate both the *GAL10* and *GAL7* transcripts in Fus (Fig. 3*A*). This construct left the intact ORFs of each gene but eliminated any termination or processing signals beyond the 3′ end of either ORF. In so doing, sequences complimentary to the probes F, G, H, and I were deleted.

TRO analysis of Δ*GAL* yeast transformed with Fus showed dramatic TI when these genes were transcribed simultaneously (Figs. 3*A* and 4*A*). For *GAL10*, transcription signals in Fus showed full levels at initiation but then dropped abruptly with probe M and more gradually to background levels from probe L to B. The initiation level of *GAL7* transcription was similar to Con, but as with *GAL10* signals it abruptly dropped at probe C. However, they then continued at a significant level, leading to a more gradual decline through the *GAL7* ORF, tailing off beyond the region at which it normally would contain a poly(A) site in

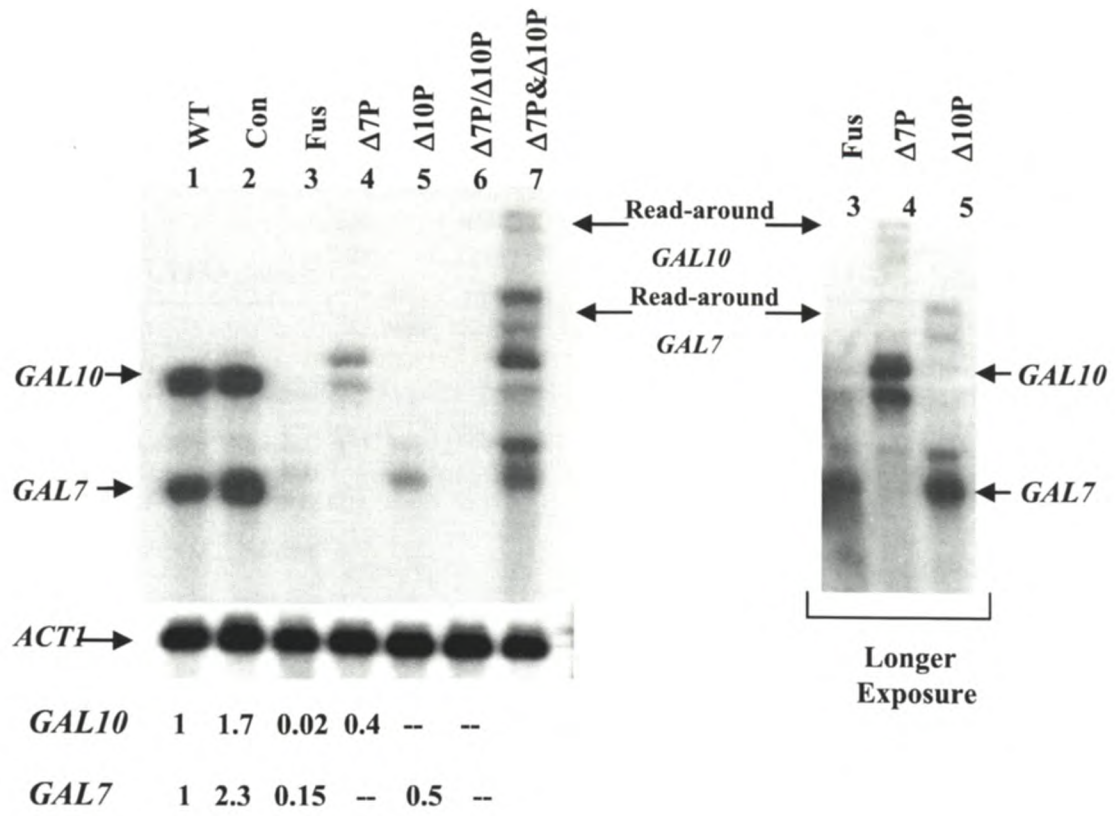


Fig. 2. Steady-state analysis of *GAL* constructs. Northern blot with total RNA from pYC10-7 (WT, lane 1), pYC10-7Con (Con, lane 2), pYC10-7Fus (Fus, lane 3), Fus-Δ7P (lane 4), Fus-Δ10P (lane 5), Fus-Δ7P/Δ10P (lane 6), and Fus-Δ7P&Δ10P (lane 7). Longer exposure shows the read-around bands for Fus-Δ7P (lane 4) and Fus-Δ10P (lane 5). The filter was striped and reprobed with *ACT1* as a loading control. Quantitated levels of steady-state *GAL10* and *GAL7* mRNA are corrected for actin and compared by using the steady-state level of *GAL10* and *GAL7* mRNA from WT as 1.

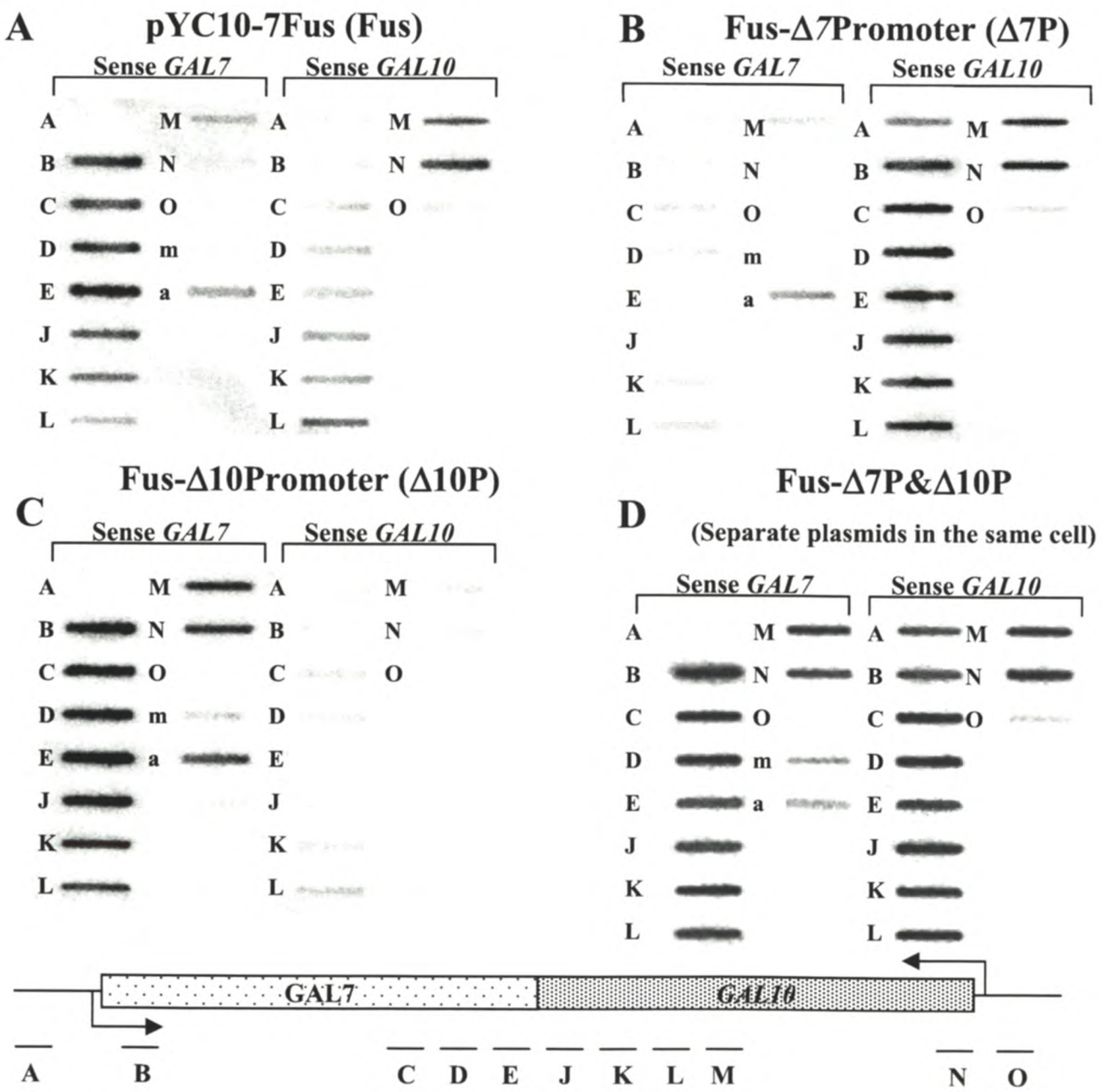


Fig. 3. TRO analysis of transformed plasmids as indicated (*A–D*). The positions of single-stranded M13 probes used are shown in the diagram at the bottom of the figure and are as described in Table 1.

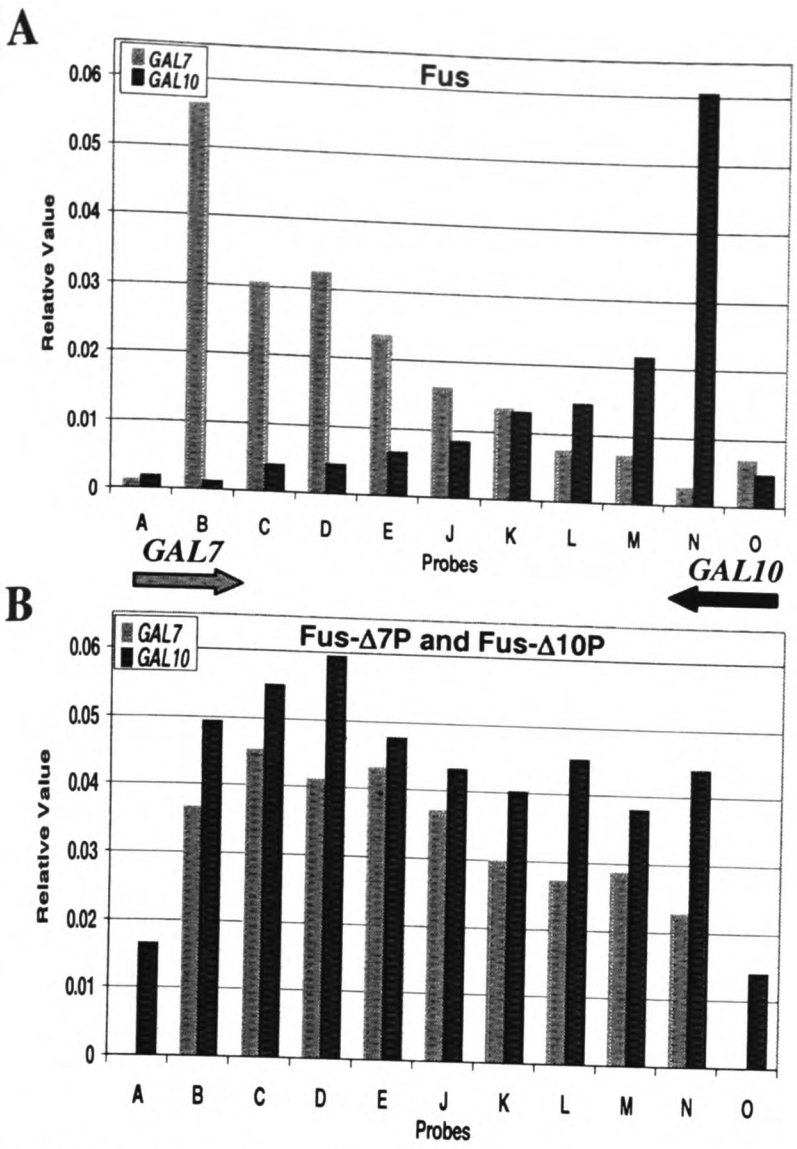


Fig. 4. Quantitation of Fus and Fus-Δ7P and Fus-Δ10P transcription profiles. (*A*) TRO analysis of Fus transformed plasmid. The signals were corrected as described for Fig. 1. Arrows indicate the directions of transcribing polymerases. (*B*) TRO analysis of Fus-Δ7P and Fus-Δ10P. In the Fus-Δ10P plasmid (giving *GAL7* transcription), the O probes have been deleted with the deletion of the *GAL10* promoter. In the *GAL10* transcription profile, decrease in *GAL10* transcription over probe A is likely the result of cryptic polyadenylation that is seen at the steady-state level (Fig. 2, lane 4). Also, the above background levels of transcription over probe O (before the *GAL10* promoter) are most likely to be the result of read-around transcription seen at steady state (Fig. 2, lane 4). The single-stranded M13 probes used are described in Table 1.

the WT construct. Thus TI of convergent transcripts seems to have no effect on initiation but acts during transcriptional elongation. We refer to this process as transcriptional collision. Steady-state analysis of *GAL10* and *GAL7* mRNA produced by the Fus construct is entirely consistent with the TRO analysis. Thus *GAL10* mRNA levels were reduced 20-fold and *GAL7* levels 4-fold as compared with the single transcription unit constructs Fus-Δ7P and Fus-Δ10P, respectively (Fig. 2, compare lane 3 with lanes 4 and 5). It is clear that transcriptional collision of convergent transcripts results in a drastic reduction in gene expression.

Transcriptional Collision Requires Convergent Transcription in cis. Because convergent transcription will generate separate antisense nascent transcripts, it is possible that the effects we observe are caused by an RNA-interference (RNAi) effect *in vivo* or more trivially by prehybridization of the antisense TRO RNA before filter hybridization during the experimental analysis. We therefore sought to demonstrate that transcriptional collision depends on transcription from the same template by creating control constructs that separated the transcription of the genes in the Fus plasmid. Elimination of the TATA box from the *GAL7* promoter inactivated *GAL7* gene transcription (Fus-Δ7P, Fig. 3*B*). The nascent transcription profile in this construct shows very low polymerase density over *GAL7* and higher levels of transcription on *GAL10* over the length of the template up to probe A. At this point nascent transcription begins to terminate (Fig. 3*B*). Steady-state analysis shows that *GAL10* uses cryptic poly(A) sites to generate a doublet of detectable mRNA bands (Fig. 2, lane 4). The lower level of these steady-state *GAL10* transcripts presumably reflects the inefficiency of this cryptic mRNA processing.

Similar analysis was done with the deletion of the UAS of the *GAL10* gene (Fus-Δ10P) showing an up-regulation of transcription across the *GAL7* gene and the elimination of transcripts from the *GAL10* promoter (Fig. 3*C*). Polymerase density appears relatively consistent across the length of the probes (probe O was deleted with the deletion of the *GAL10* promoter region). Steady-state analysis shows that Fus-Δ10P generates low levels of steady-state *GAL7* mRNA again by using inefficient cryptic poly(A) signals (Fig. 2, lane 5). Read-around bands can be seen also in both this construct and Fus-Δ7P (Fig. 2, longer exposure lanes 4 and 5), which reflect transcripts that have used the same cryptic poly(A) sites the second time around the plasmid. Finally, the Fus-Δ7P/Δ10P construct demonstrated that the deletion of both promoters in the same construct resulted in little transcription from either promoter at nascent (data not shown) or steady-state (Fig. 2, lane 6) levels.

With the above single *GAL7* and *GAL10* transcribing plasmids defined, we finally carried out the critical experiment of cotransforming both of these plasmids into the Δ*GAL* deletion strain followed by TRO analysis. At both the nascent (Fig. 3*D*) and steady-state (Fig. 2, lane 7) levels we observe profiles similar to the sum of both Fus-Δ7P and Fus-Δ10P plasmids. These data demonstrate that the transcriptional collision effect we have observed require the *GAL10* and *GAL7* genes to be present in a cis arrangement. Trivial explanations for the observed inhibition of *GAL10* and *GAL7* transcription through antisense effects are effectively excluded.

Discussion

The aim of these studies was to investigate the effect that convergent gene transcription has on PolII transcriptional activity in *S. cerevisiae*. In the presence of intact mRNA 3′-end-processing signals the convergent *GAL7* and *GAL10* genes behaved similarly to the tandem-arranged genes as found in the WT situation. However, once the intergenic region, including the termination signals, was deleted there is a drastic effect on the level of transcriptional processivity in both genes, a phenomenon that we term transcriptional collision.

Both *GAL7* and *GAL10* transcriptional initiation appear unperturbed. Instead, the impact of interference is seen in the elongation stage of the transcription. The precise biological cause for this collision effect remains to be identified. One possibility is that as transcription proceeds, the polymerases collide and cause one or both to release from the template. Alternatively, a polymerase complex, when confronted with one polymerase on the opposite strand, stalls and is unable to advance further. This collision effect could be a direct physical impediment to the transcription machinery or an indirect effect caused by supercoiling changes to the DNA template during transcription (26–29). In the later scenario, a transcription bubble progresses along a template, creating positive supercoils. If these positive supercoils were generated from both strands simultaneously, it would create a region of hyper-supercoiling that would be predicted to prevent further advancement in either direction on the template. Such a mechanism explains the steady-state and nascent drop in detectable transcription.

The transcriptional collision seen in this investigation occurred in cis and was alleviated when the opposing polymerase was prevented from transcribing simultaneously. Separate ex-

pression of *GAL10* (Fus-Δ7P) and *GAL7* (Fus-Δ10P) in the same cell (Fus-Δ7P&Δ10P; see Fig. 3*D*) showed no interference even though it created antisense transcripts. These data demonstrate that simultaneous transcription of these genes is occurring and also argues against any RNAi effects in this *S. cerevisiae* strain. Previous research investigating RNAi in *S. cerevisiae* has detected RNAi in cis but not in trans. It seems plausible that these so-called cis RNAi effects actually may be caused by transcriptional collision on the same template rather than true RNAi (ref. 30 and references therein).

This study also provides further information about the capacity of *GAL10* and *GAL7* to interact in their WT arrangement. In the construct Con, elimination of transcribing polymerases over the promoter for *GAL7* (Fig. 1*B*, probe A) presumably permits the unimpeded binding of Gal4p transcription factor (23) and thus increases transcription of *GAL7*. The level of transcription before the *GAL7* promoter (*GAL7* probe A) in WT clearly illustrates the extent of read-through transcription into the *GAL7* promoter region, which is in contrast to the same region in Con, which showed no detectable signal before the promoter of *GAL7*. The quantified 2-fold increase in *GAL7* nascent transcripts from WT to Con could underrepresent the true increase in transcripts that initiated at the *GAL7* promoter, because the WT level of transcription detected in the *GAL7* region is a combination of transcripts that initiated at *GAL7* and the read-through from the *GAL10* promoter. However, this observed 2-fold increase is consistent with the increase at steady-state levels (Fig. 2, lane 2). Thus we observe an increase of 2.3-fold in *GAL7* mRNA for Con as compared with WT. These data further support the existence of TI in the WT orientation of *GAL10* and *GAL7*.

In this paper we have defined a new type of TI termed transcriptional collision that can be alleviated when transcription on the opposing strand is suppressed. The relevance of this phenomenon to the endogenous genes of *S. cerevisiae* remains to be established. However, it is abundantly clear that coexpressed adjacent genes must guard against TI when arranged in tandem or transcriptional collision when convergent. The role of efficient PolII-termination signals is critical to allow the full expression of such genes. It seems plausible that restricted or regulated gene expression could be effected by the regulation of PolII termination.

We are grateful to members of the Proudfoot lab for advice throughout these studies and to Nick Kent for reading the manuscript. E.M.P. was supported by the HLP Educational Trust. This research was supported by Wellcome Trust Program Grant 062329 (to N.J.P.).

1. Feldmann, H., Aigle, M., Aljinovic, G., Andre, B., Baclet, M. C., Barthe, C., Baur, A., Becam, A. M., Biteau, N., Boles, E., *et al.* (1994) *EMBO J.* **13,** 5795–5809.
2. Dujon, B. (1996) *Trends Genet.* **12,** 263–270.
3. Malavasic, M. J. & Elder, R. T. (1990) *Mol. Cell. Biol.* **10,** 2809–2819.
4. Peterson, J. A. & Myers, A. M. (1993) *Nucleic Acids Res.* **21,** 5500–5508.
5. McGrath, J. P. & Varshavsky, A. (1989) *Nature (London)* **340,** 400–404.
6. Barker, D. G., White, J. H. M. & Johnston, L. H. (1987) *Eur. J. Biochem.* **162,** 659–667.
7. Aranda, A., Perez-Ortin, J. E., Moore, C. & del Olmo, M. (1998) *Nucleic Acids Res.* **26,** 4588–4596.
8. Aranda, A., Perez-Ortin, J. E., Moore, C. & del Olmo, M. (1998) *RNA* **4,** 303–318.
9. Osborne, B. I. & Guarente, L. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 4097–4101.
10. Kruglyak, S. & Tang, H. X. (2000) *Trends Genet.* **16,** 109–111.
11. Fu, J. H., Gnatt, A. L., Bushnell, D. A., Jensen, G. J., Thompson, N. E., Burgess, R. R., David, P. R. & Kornberg, R. D. (1999) *Cell* **98,** 799–810.
12. Poglitsch, C. L., Meredith, G. D., Gnatt, A. L., Jensen, G. J., Chang, W. H., Fu, J. H. & Kornberg, R. D. (1999) *Cell* **98,** 791–798.
13. Birse, C. E., Minvielle-Sebastia, L., Lee, B. A., Keller, W. & Proudfoot, N. J. (1998) *Science* **280,** 298–301.
14. Zhao, J., Hyman, L. & Moore, C. (1999) *Microbiol. Mol. Biol. Rev.* **63,** 405–445.
15. Puig, S., Perez-Ortin, J. E. & Matallana, E. (1999) *Curr. Microbiol.* **39,** 369–373.
16. Greger, I. H. & Proudfoot, N. J. (1998) *EMBO J.* **17,** 4771–4779.
17. Springer, C., Valerius, O., Strittmatter, A. & Braus, G. H. (1997) *J. Biol. Chem.* **272,** 26318–26324.
18. Tanaka, S., Halter, D., Livingstonezatchej, M., Reszel, B. & Thoma, F. (1994) *Nucleic Acids Res.* **22,** 3904–3910.
19. Snyder, M., Sapolsky, R. J. & Davis, R. W. (1988) *Mol. Cell. Biol.* **8,** 2184–2194.
20. Hill, A. & Bloom, K. (1987) *Mol. Cell. Biol.* **7,** 2397–2405.
21. Rocco, V., Demassy, B. & Nicolas, A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 12068–12072.
22. Hansen, K., Birse, C. E. & Proudfoot, N. J. (1998) *EMBO J.* **17,** 3066–3077.
23. Greger, I. H., Aranda, A. & Proudfoot, N. (2000) *Proc. Natl. Acad. Sci. USA* **97,** 8415–8420.
24. Gietz, R. D. & Schiestl, R. H. (1995) *Methods Mol. Cell. Biol.* **5,** 255–269.
25. Birse, C. E., Lee, B. A., Hansen, K. & Proudfoot, N. J. (1997) *EMBO J.* **16,** 3633–3643.
26. Dunaway, M. & Ostrander, E. A. (1993) *Nature (London)* **361,** 746–748.
27. Tsao, Y. P., Wu, H. Y. & Liu, L. F. (1989) *Cell* **56,** 111–118.
28. Giaever, G. N. & Wang, J. C. (1988) *Cell* **55,** 849–856.
29. Liu, L. F. & Wang, J. C. (1987) *Proc. Natl. Acad. Sci. USA* **84,** 7024–7027.
30. Atkins, D., Arndt, G. M. & Izant, J. G. (1994) *Biol. Chem. Hoppe-Seyler* **375,** 721–729.

Part II: rRNA Gene Termination

Preface

In *Saccharomyces cerevisiae* the site of transcriptional termination in the Pol I transcribed rRNA genes is thought to occur downstream of the mature 25S gene 3' end. However, this conclusion is based largely on steady state RNA analysis. In Section 1, the position of nascent termination for rRNA genes is determined. In Section 2, two *S.cerevisiae* strains with variation in Pol I subunit composition are investigated to see how termination is effected. Section 3 investigates the role that Rnt1p, a protein involved in early rRNA maturation, plays in the location of nascent termination of rRNA genes.

Introduction

Pol I termination is a multi-step process that requires varied sequence elements across different species but probably functions through a similar mechanism in all eukaryotes (Reeder 1999). Termination elements for Pol I display strong orientation and factor dependence generating transcripts with 3' ends 11-21 nt upstream of the of the terminator sequence. Pre-rRNA transcripts are rapidly processed to give mature rRNA gene products through a complex processing pathway.

In murine rRNA termination, the 3' end of the transcript has been mapped to 21 nt upstream of a tandemly repeated 18 bp sequence motif referred to as the 'Sal box' or 565 bp downstream of the 28S RNA coding region (Grummt *et al.* 1985). The Sal box contains a consensus motif (AGGTCGACCAGA/TT/ATCCG) that is repeated eight times in the several hundred base pairs of the non-transcribed spacer. Additionally, each Sal box is flanked by long pyrimidine stretches that also play a role in efficient 3' end formation. The Sal box is bound in a sequence specific manner by the termination factor mTTF-1p that shows pronounced homology with the yeast terminator protein, Reb1p, and the human TTF-1p (for review see Mason *et al.* 1998). In *S.cerevisiae*, termination is mediated by a protein binding site (CCGGGTAAAAA) which is bound by the terminator protein Reb1p and a T-rich release element located upstream of this binding site (Lang 2002). Additionally, evidence suggests that trans-acting factors are involved in transcript release from murine and yeast

Pol I systems (Mason *et al.* 1997).

Analysis of the yeast rRNA genes has suggested the presence of several termination sites. Initial studies concluded that Pol I terminates near a position +210 downstream of the 3' end of mature 25S ribosomal RNA (Kempersveenstra *et al.* 1986; Vandersande *et al.* 1989). This conclusion was questioned by subsequent studies which demonstrated that binding of Reb1p, in conjunction with an adjacent T-rich stretch, could induce efficient termination *in vitro* at +93, closer to the 3' end of 25S RNA (Lang and Reeder 1993; Lang *et al.* 1994; Lang and Reeder 1995). Subsequently, the site of Pol I termination *in vivo* was re-examined by a combination of S1 nuclease protection and RT-PCR analysis (Reeder *et al.* 1999). This study lead to the conclusion that roughly 90% of all transcripts terminated at the +93 Reb1p terminator with the remaining 10% terminating at the +250 site which was referred to as the "failsafe" terminator. A possible caveat to the Reeder *et al.* (1999) study was that it analysed the abundance of steady state RNAs, and results could have been skewed if different segments of the transcripts under study had markedly different half-lives. Additionally, these termination sites have been questioned, as some researchers have not been able to detect products terminating at the same sites (Elela *et al.* 1996).

Saccharomyces cerevisiae Pol I consists of 14 subunits with five shared by all three polymerases, two shared between Pol I and Pol III and the remainder unique subunits including two large subunits (A190, A135) and five additional subunits (A49, A43, A34.5, A14 and A12.2) (Lanzendorfer *et al.* 1997). The *RPA12* gene encodes a 13.5 kd subunit (A12) of Pol I whose role in transcription has yet to be established. A12 (Rpa12p) is homologous to Pol II subunit B12.6 (Rpb9p) and Pol III subunit C11 (Rpc11p), which are all homologous to TFIIS (Van Mullem *et al.* 2002). Each of these proteins is tripartite and formed by two zinc-binding domains. Additionally, there are large similarities in the C-terminal regions and also significant similarities in the N-terminal regions that suggest structural and functional homology (Chedin *et al.* 1998).

The Pol III subunit C11 is involved in the intrinsic RNA cleavage activity of the polymerase and termination efficiency (Chedin *et al.* 1998). Pol III missing the *S.cerevisiae* C11 showed a decreased

dwelling time of the polymerase and were not able to accurately identify pause and termination sites (Chedin *et al.* 1998). Similarly, the deletion of the Pol II homolog B12.6 (Rpa9p) has been shown to increase elongation through sequences that signal complexes to pause or arrest (Awrey *et al.* 1997). In addition to increasing read through, Pol II lacking this subunit generated greater quantities of RNA, with both of these qualities suppressed upon the addition of purified Rpa9p (Awrey *et al.* 1997). While Pol II lacking Rpa9p still retains affinity for TFIIS, it is much less responsive to TFIIS cleavage than the wild type Pol II suggesting that the subunit may be responsible for transmitting signals to the elongating ternary complex (Awrey *et al.* 1997). Somewhat paradoxically, TFIIS, which also shares homology with Rpa12p, is known to engage the transcribing Pol II complex and assist in bypassing blocks to elongation by stimulating the Pol II intrinsic cleavage activity (Wind and Reines 2000).

Immunolocalization of subunit A12 on a three-dimensional model of Pol I indicates that it is located outside the groove that is thought to bind DNA on the finger-like projection (Klinger *et al.* 1996; Chedin *et al.* 1998). This projection is thought to be flexible and able to wrap around the bound template to prevent disassociation during progressive synthesis possibly giving A12 a role in modulating the structure of the elongation complex (Awrey *et al.* 1997; Chedin *et al.* 1998). In two-hybrid analysis, Rpa12p was shown to partner with the second largest subunit of Pol I, Rpa135p (Flores *et al.* 1999).

Deletion analysis has shown that only the N-terminal half of the Rpa12p subunit is critical for function. Every partial deletion of the N-terminal domain acts like a null mutant while deletions of the conserved C-terminal domain have no growth effect (Van Mullem *et al.* 2002). Growth defects of Rpa12p null strains include poor growth at low temperature, lack of growth at 34°C or above in addition to a strong sensitivity to 6-azauracil and mycophenolate (Nogi *et al.* 1993; Van Mullem *et al.* 2002). Null deletions of Rpa14p, another non-essential Pol I subunit, are lethal when combined with an N-terminal or null deletion or Rpa12p (Gadal *et al.* 1997).

As observed in Nogi *et al.* (1993), the cellular concentration of the two largest subunit of Pol I,

Rpa190p and Rpa135p, decrease in Δ*RPA12* strain. It was also noted that overexpression of Rpa190p partially suppresses the high temperature sensitivity of the Δ*RPA12* strain. From this it was suggested that the Rpa12p is important in the assembling and/or a structural role in Pol I with the non-incorporated Rpa190p subject to proteolysis (Nogi *et al.* 1993). Interactions with other Pol I subunits increased the speculation that Rpa12p plays a role in the structural maintenance of Pol I.

After transcription of the 35S pre-rRNA in *S.cerevisiae*, processing events create the mature 18S, 25S and 5.8S ribosomal RNAs. An important early step in this processing is cleavage by Rnt1p, reported to occur co-transcriptionally at a tetraloop with a conserved AGNN sequence located within 13 bp - 16 bp of the cleavage sites in the 3' ETS (Chanfreau *et al.* 2000; Nagel and Ares 2000; Lebars *et al.* 2001; Wu *et al.* 2001). Rnt1p is a double strand specific endoribonuclease that shares sequence homology with bacterial RNase III (Elela and Ares 1998) and is involved in cleavage of the primary transcript at the 3' and 5' end of the pre-rRNA as well as the maturation of spliceosomal snRNAs and snoRNAs (Chanfreau *et al.* 1997; Chanfreau *et al.* 1998; Elela and Ares 1998; Qu *et al.* 1999; Seipelt *et al.* 1999). The involvement of Rnt1p in the maturation of other RNA substrates that are also involved in rRNA maturation, such as the U14 a snoRNA, allows Rnt1p to influence rRNA maturation at multiple levels (Chanfreau *et al.* 1998).

The cleavage site for Rnt1p has been shown to occur *in vivo* at positions +14 and +49 relative to the 3' end of the mature 25S rRNA as indicated in Figure II 0.1 (Kufel *et al.* 1999). The Rnt1p deletion strain (*rnt1*) results in a slow growth phenotype and accumulation of incompletely processed transcripts (Elela *et al.* 1996; Chanfreau *et al.* 1998; Kufel *et al.* 1999). The *rnt1* stain also has slow processing of other cleavage sites which are not lethal suggesting that there may be other endonucleases that are able to partially substitute for Rnt1p *in vivo* (Kufel *et al.* 1999).

In vivo steady state analysis shows that in the absence of functional Rnt1p, cleavage events in the 5' ETS and 3' ETS are greatly impaired resulting in the an accumulation of the full length 35S transcript as well as several new cleavage products (Elela *et al.* 1996; Kufel *et al.* 1999; Lamontagne *et al.* 2000). In the temperature sensitive *rnt1* strain, Elela *et al.* (1996) isolated steady state cleavage products at

+195 and +210 nt past the 3' end of the mature 25S rRNA. Additionally, Kufel *et al.* (1999) identified *in vivo* steady state 3' ends located between +180 and +264 in *rnt1*. Both of these 3' ends were further downstream than detected by *in vivo* steady state analysis of wild type strains (Kufel *et al.* 1999; Reeder *et al.* 1999). Elela *et al.* (1996) and Kufel *et al.* (1999) interpreted the 3' ends seen in *rnt1* mutant strain as representing the actual site of termination in rRNA genes without the co-transcriptional processing by Rnt1p.

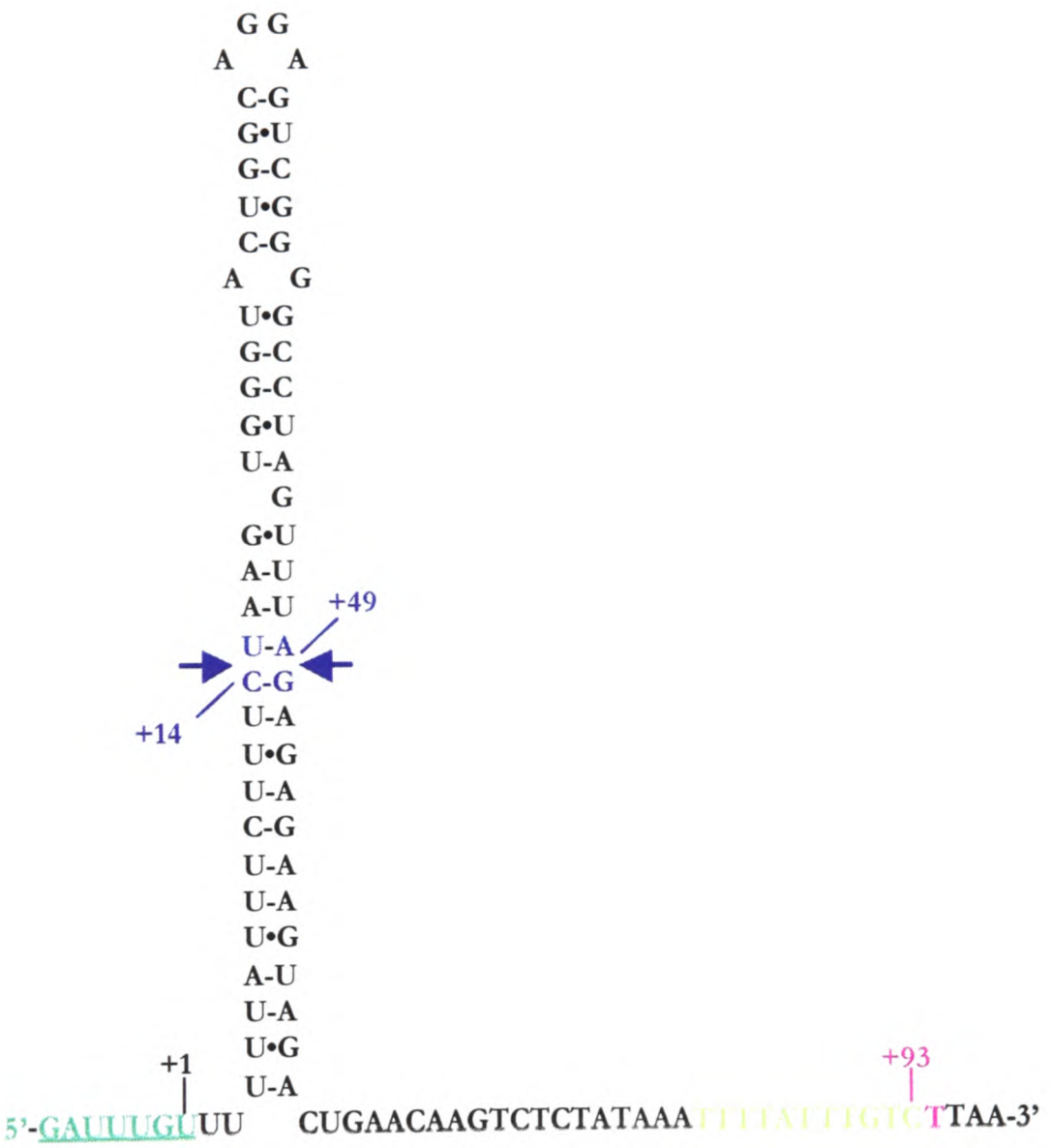


Figure II 0.1 Model for the RNA structure surrounding the 3' ETS. The Rnt1p cleavage sites are blue and indicated by blue arrows. The green sequence indicates the mature 3' end of 25S gene. The +93 termination site is identified in pink. The lime coloured sequence indicates the 10-12 bp release element. Diagram largely reproduced from Kufel *et al.*1999.

The sequence of the region downstream of the mature 3' end of 25S gene is shown in Figure II 0.2 and Figure II 0.1 with various features indicated. The primary termination site at +93 and the failsafe +250 terminator are indicated as described in Reeder *et al.* (1999). The Reb1p binding site and upstream release element are also indicated. Additionally, the cleavage sites for Rnt1p are indicated as described in Kufel *et al.* (1999).



Figure II 0.2 The 25S gene and location of the mature 3' end are indicated in green. Distances shown in grey are from the mature 3' end of 25S gene. The primary termination site at +93 and the 'failsafe' terminator at +250 are indicated in pink. The 11 bp Reb1p binding and pause element is located in the purple box. The lime coloured sequence indicates the 10-12 bp release element. The two Rnt1p cleavage sites at +14 bp and +49 bp are identified in blue.

Part II - Section 1

Transcriptional termination in rRNA genes

In this study, TRO analysis is employed to determine the nascent termination profile downstream of the mature 3' end of the rRNA genes. Three *S.cerevisiae* strains were investigated to minimize any strain specific variations that may occur. The evidence presented here is consistent with a primary termination site located at +93 with the remainder of transcripts terminating further downstream. These data obtained by nascent transcription analysis therefore confirm the previous steady state RNA analysis.

1.1 Probes for rRNA genes

The *S.cerevisiae* rDNA repeat is diagrammed in Figure II 1.1. Position of single stranded probes designed to detect transcription over the rRNA termination region are indicated. Probes 'r2'-'r7' are located in the region just downstream of the mature 3' end of 25S. In addition there are two genic ('18S' and '25S') probes and one promoter proximal ('pro') probe. Initially, only the termination region probes ('r2'-'r7') were employed with the genic ('18S' and '25S') and promoter proximal ('pro') probes being added subsequently to give additional information about overall rRNA gene transcription. In Figure II 1.2 the probes covering the termination region are shown in more detail. All probes are described further in the Materials section.

In addition to the probes containing rDNA sequence, three additional probes were used to monitor non-specific background hybridisation and transcription level. As before probe 'm' is an M13 probe with no insert that serves as a negative control for background hybridisation. Due to strain specific cross hybridisation in MWK028 derived strains, probe 'p' containing a fragment of the *Schizosaccharomyces pombe* ura4 gene was used as a cross species negative control. Probe 'a' contains a fragment of the *ACT1* gene from *Saccharomyces cerevisiae* that serves as a positive transcriptional control as used before.

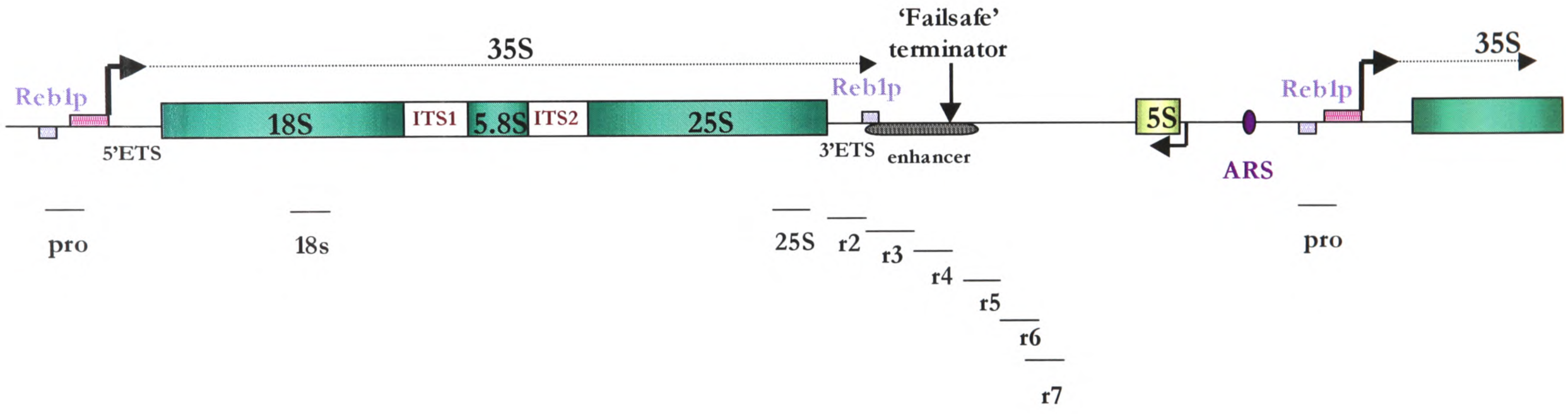


Figure II 1.1 Diagram of rDNA repeat. Direction of the 35S polycistronic transcript is indicated by arrow and dotted line. The 18S, 5.8S and 25S genic regions are indicated in green. The 5' and 3' external transcribed spacers (ETS) are indicated. The internal transcribed spacers (ITS1 and ITS2) are in brown. Pol III transcribed 5S gene is identified by the yellow box with direction of transcription indicated by the arrow. The light purple boxes identify the location Reb1p binding sites. Promoter region of 35S transcript is identified by the pink boxes. The purple oval identifies the location of the autonomously replicating sequence (ARS) . The enhancer region is identified by grey oblong. The replication fork barrier (RFB), intergenic sequence (IGS) and the non-transcribed spacers (NTS1 and NTS2) are not indicated but can be seen in Figure F of the Introduction.

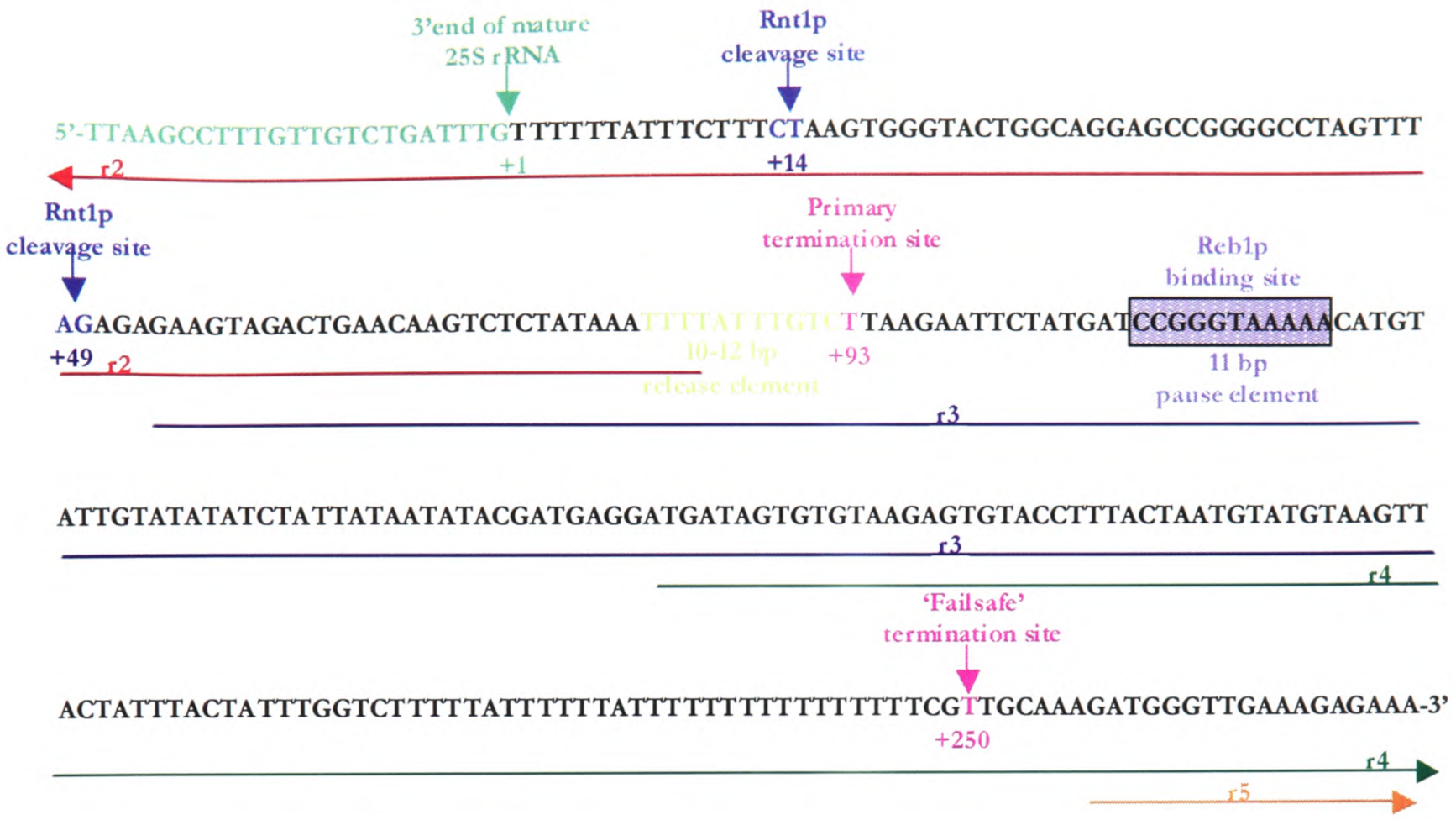


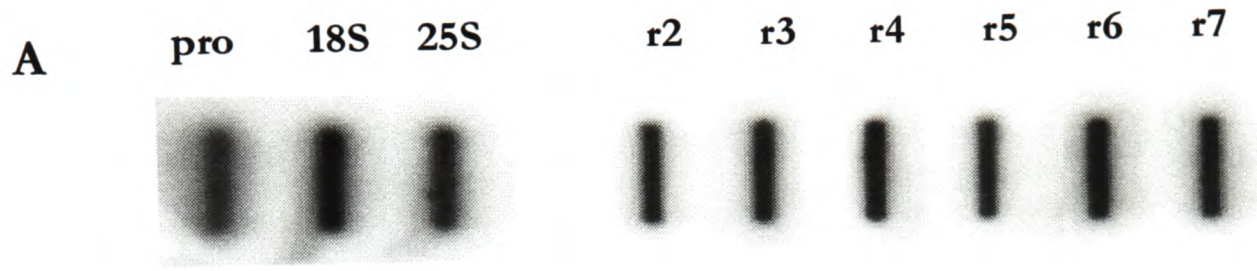
Figure II 1.2 The 25S gene and location of the mature 3' end are indicated in green. Distances shown in grey are from the mature 3'end of 25S gene. The primary termination site at +93 and the 'failsafe' terminator at +250 are indicated in pink. The 11 bp Reb1p binding and pause element is located in the purple box. The lime coloured sequence indicates the 10-12 bp release element. The two Rnt1p cleavage sites at +14 bp and +49 bp are identified in blue. The single stranded M13 probes located downstream of the mature 3' end of the 25S genic region ('r2'-'r7') are shown in more detail. The location of probe 'r2' is indicated in red. The location of probe 'r3' is identified in blue. The location of probe 'r4' is in green. The location of probe 'r5' is in orange. Arrow at the end of probes 'r2', 'r4' and 'r5' indicate that the probe extends beyond listed sequence. The purple box shows the location of the binding of the terminator protein Reb1p.

1.2 Synthetic RNA demonstrates specificity of M13 probes

To demonstrate that each probe was capable of effectively hybridising to the specific transcripts, radiolabeled synthetic RNA was generated and annealed to the probes. This was accomplished by *in vitro* transcription of rRNA fragments cloned into pGem4 vector, followed by hydrolysis and hybridisation as it is performed in the TRO protocol. Probes 'r2'-'r7' were transcribed from one continuous insert in the pGem4 vector while 'pro', '18S' and '25S', were each transcribed from separate vectors. After synthetic transcription with these three non-contiguous probes approximately the same quantity of each radioactive RNA extraction was added to the membrane.

Figure II 1.3A shows an autograph of a membrane to which the radiolabeled synthetic RNA was hybridised. The quantitation shown in Figure II 1.3B is the ratio of the radioactivity bound at each probe per 'U' relative to the level of hybridisation of probe 'r2' for the contiguous 'r2'-'r7' transcript

and '18S' for the non-contiguous genic and promoter proximal probes. From these data (Figure II 1.3 A and B) it is evident that that all the probes hybridise radiolabeled RNA with nearly equivalent efficiency (varying less than two fold). As all probes were not from a contiguous template, the ratio of hybridisation of synthetic RNA was not used to correct subsequent TRO membrane analysis.



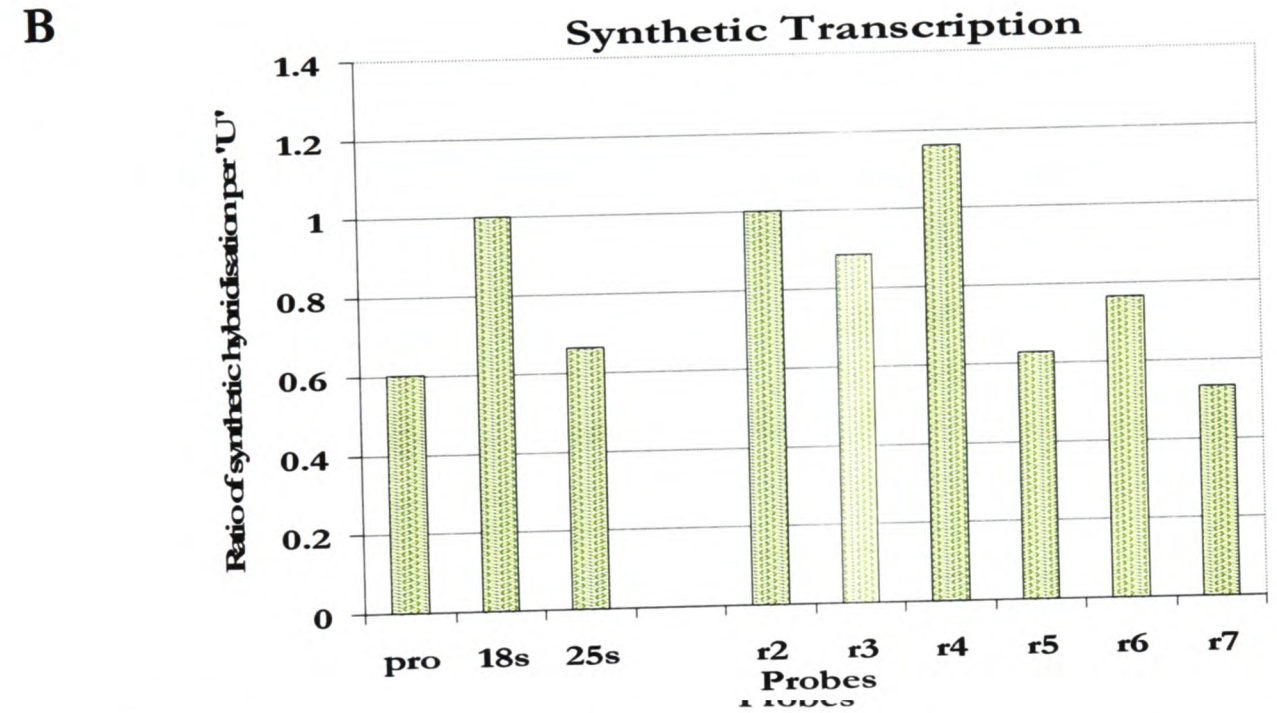


Figure II 1.3 Synthetic transcription over the rRNA genes. (A) A typical hybridisation to membrane of synthetic transcription reaction. The probes 'r2'-'r7' were transcribed from a contiguous synthetic transcript. The genic probes ('18S' and '25S') and the promoter proximal probe ('pro') were on separate templates and equal counts of each transcription reaction were added to the membrane for hybridisation. (B) Graph of the ratio of hybridised radiolabeled transcripts for each probe to the level of hybridisation at probe 'r2' for the continuous transcript of probes 'r2'-'r7'. Probes 'pro', '18S' and '25S' were each compared to the level of radiolabeled transcripts that hybridised to probe '18S'.

1.3 Identification of sites of Pol I nascent rRNA gene termination

TRO analysis as described previously was employed to detect *in vivo* nascent transcript termination downstream of the 3' end of the mature 25S gene. This investigation looks at three different strains, which were all wild type for the rRNA gene transcription. This repetition enabled minimisation of possible strain specific termination effects. Figure II 1.4 indicates where the probes are in relation to sequence elements of the rRNA termination region.

1.3.1 *S.cerevisiae* strain *ΔGAL* transcription termination profile

The rRNA gene termination profile for *S.cerevisiae* strain *ΔGAL* was determined using TRO. Figure II 1.4 shows a typical slot blot hybridisation for a TRO experiment with *ΔGAL* grown in glucose. As before, the quantification shown in Figure II 1.5A was corrected for background, adjusted for the number of U's per probe and the level of actin. The number indicated by the bar graph is the average of seven independent experiments with the standard deviation shown by error bars. The average level of transcription at each probe was compared to the average transcription at probe 'r2' to generate the ratio of termination in Figure II 1.5B. It should be noted when the experiments on this strain were preformed before the probes 'pro', '18S' and '25S' had been made and are therefore not represented in this data.

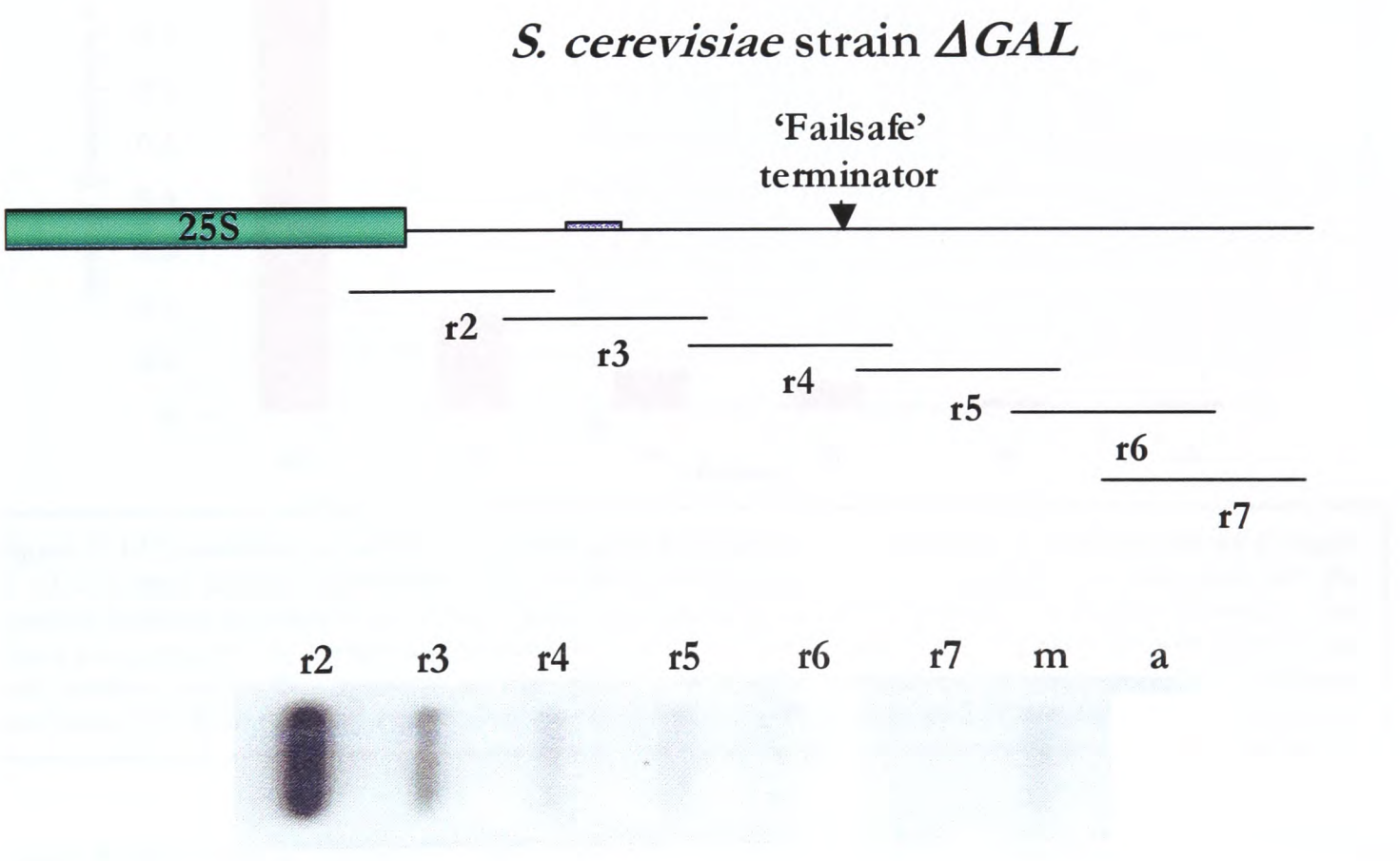
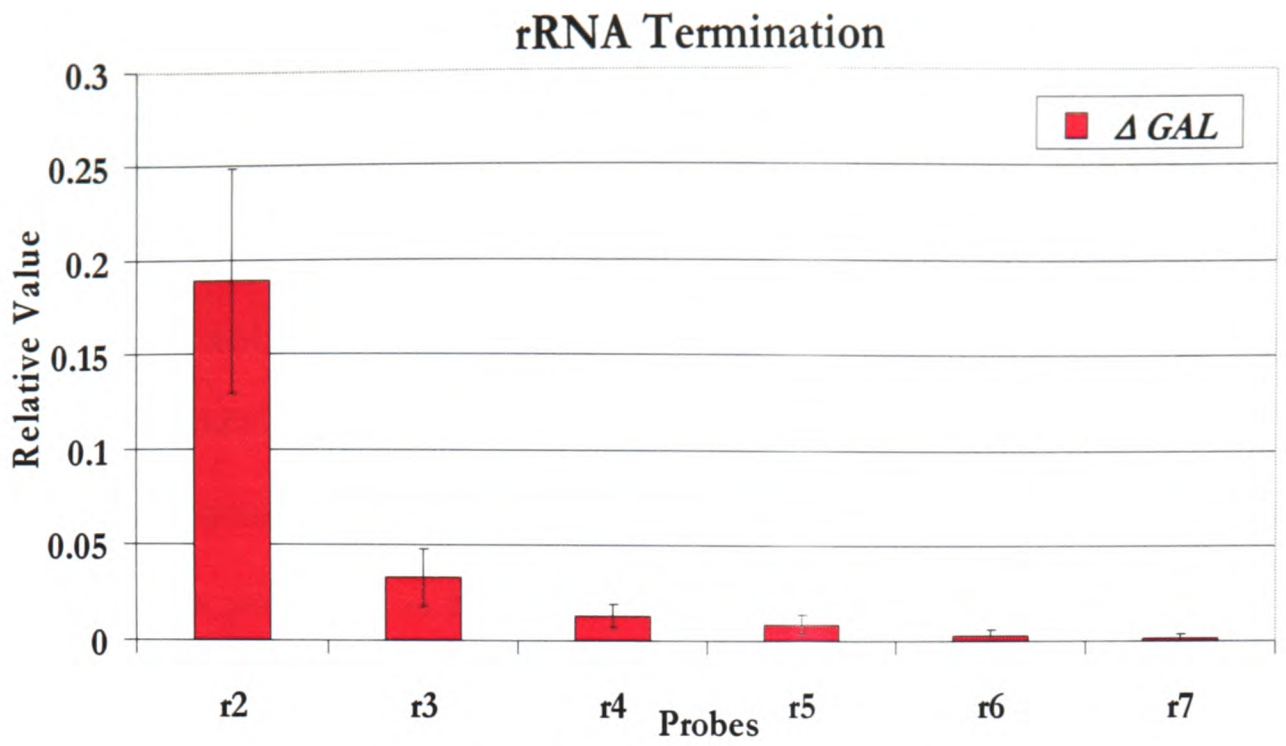


Figure II 1.4 Transcription over rRNA genes in strain *ΔGAL*. A diagram of the location of single stranded probes and a typical hybridisation of radiolabeled rRNA gene transcripts of *S. cerevisiae* strain *ΔGAL* to membrane. The genic probes ('18S' and '25S') and the promoter proximal probe ('pro') had not been generated at the time these experiments were carried out and are thus not represented in this data. Probe 'm' is an M13 probe without an insert serving as a negative hybridisation control. Probe 'a' is an M13 probe with a fragment of the *S.cerevisiae* actin gene serving as a positive transcription control.

A



B

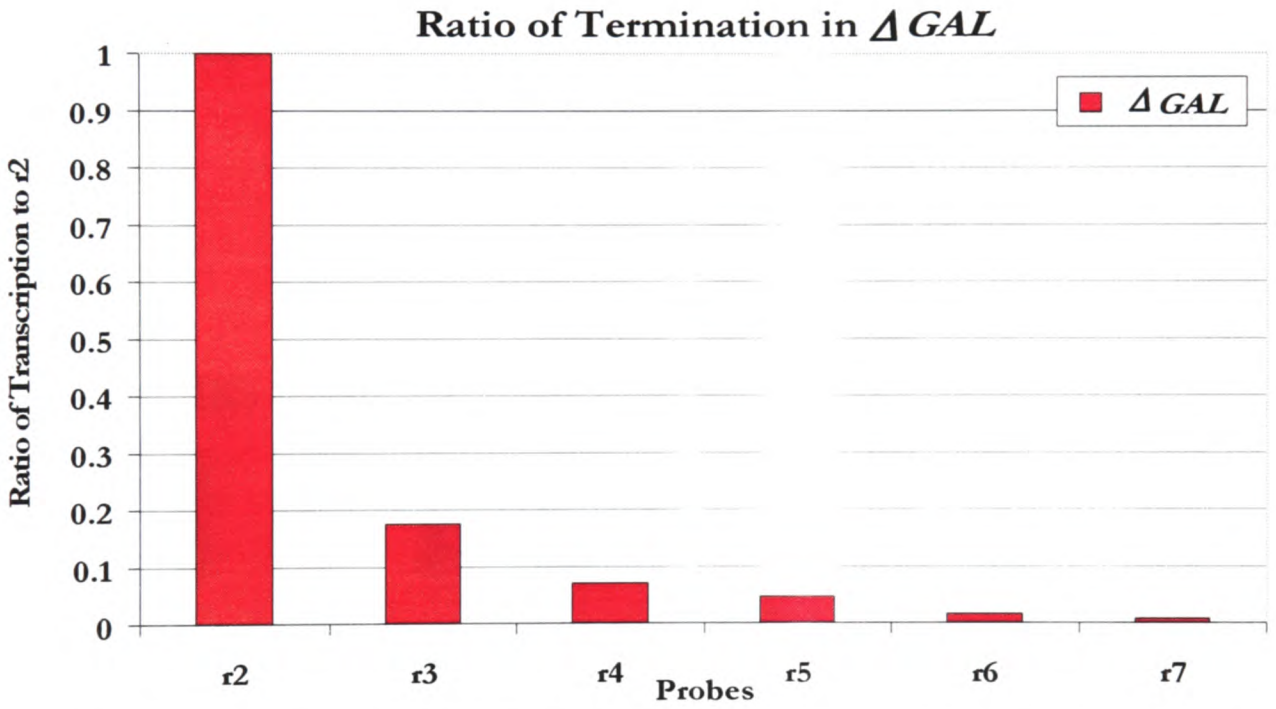


Figure II 1.5 Quantitation of strain Δ*GAL* rRNA gene termination. (A) Quantitation of membrane shown in Figure II 1.4 and other repeated experiments. The average of all seven experiments performed was calculated with the standard deviation indicated by error bars. Signals were quantified in a Phosphoimager (Molecular Dynamics). The values were corrected for background hybridisation and 'U' content in each probe. Signals were normalized to the actin positive transcription probe 'a' on each filter to allow direct comparison of transcription levels between membranes. (B) Ratio of the average transcription at each probe to the average level of transcription over probe 'r2'.

As seen by the ratio in Figure II 1.5B, over 90% of the transcripts terminate before the 'r4' probe, located downstream of the Reb1p binding site. Additional transcripts terminate between probes 'r4' and 'r5', which is consistent with the use of the +250 'failsafe' terminator. The level of transcription drops to near background by 'r6'. These results are consistent with the majority of transcripts terminating at the Reb1 terminator (between 'r2' and 'r4') and the remainder terminating at or shortly after the 'failsafe' terminator (between 'r4' and 'r6'). Therefore, the rRNA genes nascent termination profile confirms previous *in vivo* steady state RNA data.

1.3.2 *S.cerevisiae* strains MWK028 and *RNT1* transcription termination profiles

TRO analysis of *S.cerevisiae* wild type strains MWK028 and *RNT1* were performed as described previously. The MWK028 strain expresses the *SNR6* gene from an ectopic plasmid but exhibits wild type growth (Kaiser and Brow 1995). A typical slot blot hybridisation for a TRO experiment of the six performed with MWK028 is shown in Figure II 1.6A. Additionally, in Figure II 1.6B there is a typical slot blot hybridisation for one of the four TRO performed with *RNT1*. The quantification for both strains was performed as described before and is shown in Figure II 1.7. Figure II 1.8 shows a composite graph of the ratios of transcription termination for each of the wild type *S.cerevisiae* strains described above. As was done previously, the level at each probe was compared to the detectable transcription for that strain at probe 'r2'.

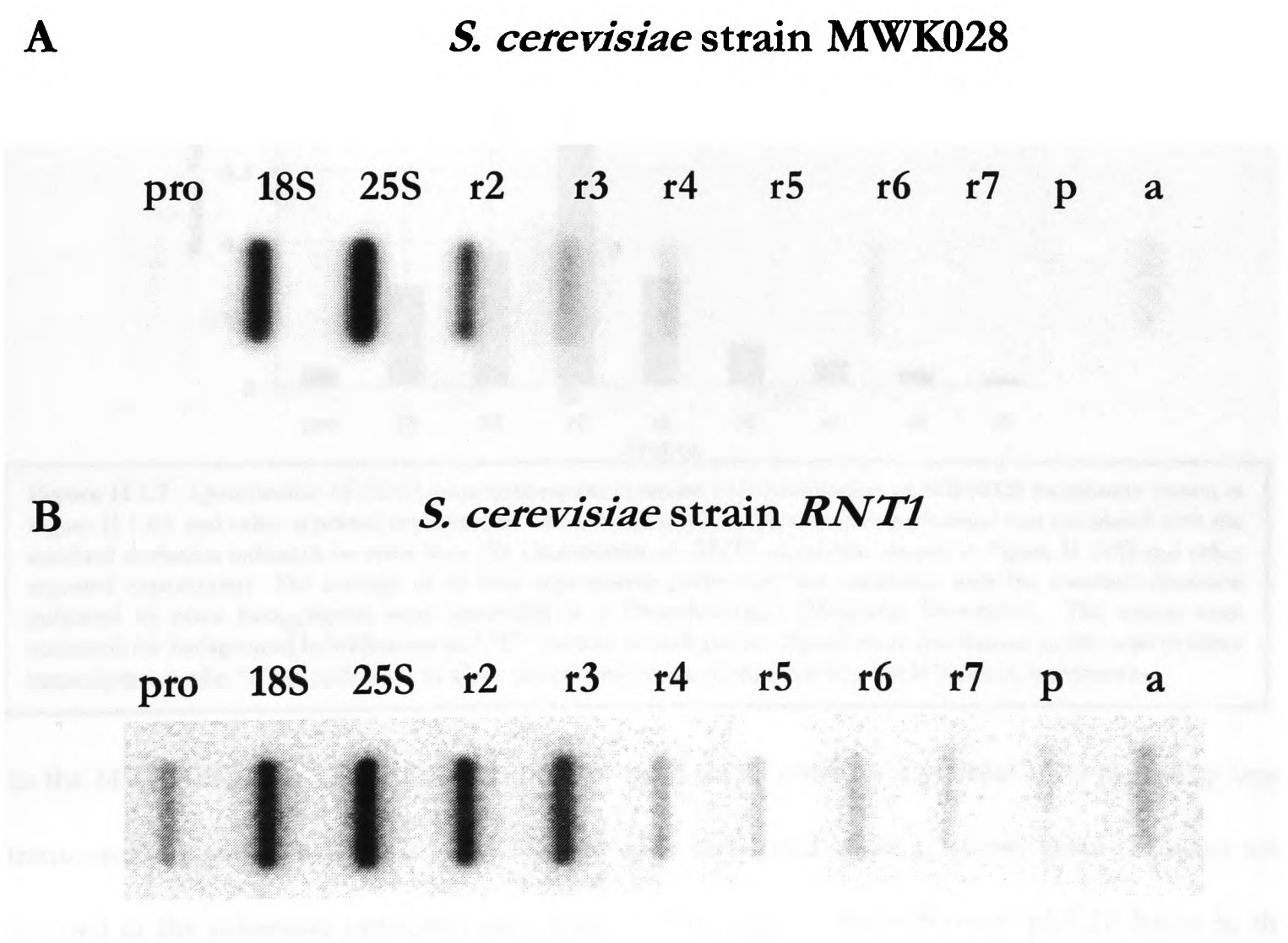
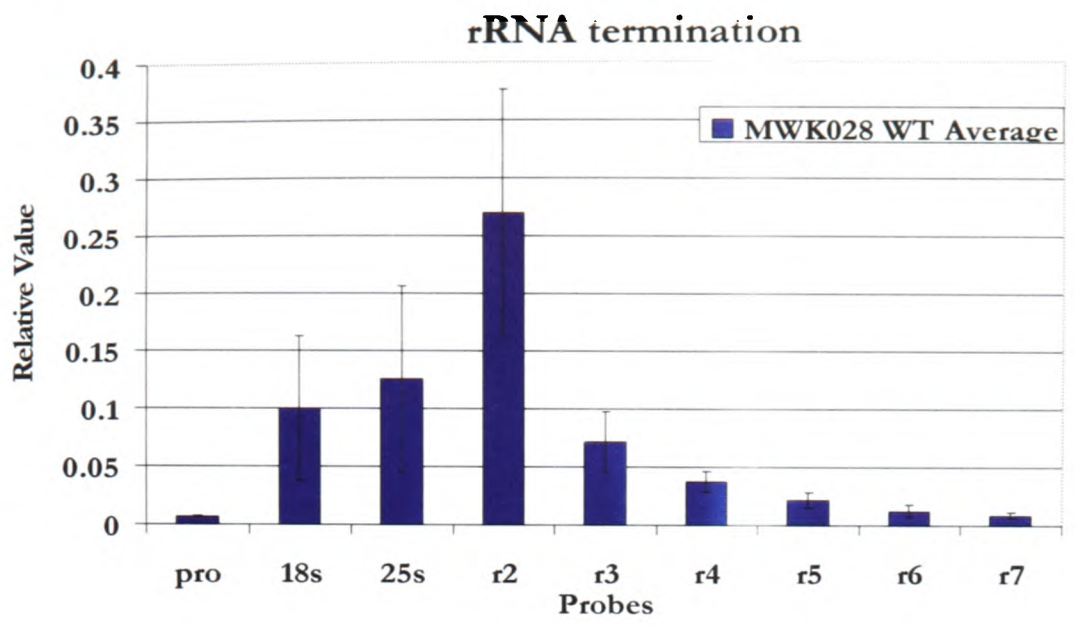


Figure II 1.6 Transcription over rRNA genes. (A) A picture typical hybridisation of radiolabeled rRNA gene transcripts of *S.cerevisiae* strain MWK028 to membrane. (B) A picture typical hybridisation of radiolabeled rRNA gene transcripts of *S.cerevisiae* strain *RNT1*. Due to strain specific cross hybridisation in MWK028 derived strains, probe 'p' containing a fragment of the *Schizosaccharomyces pombe* ura4 gene was used as a cross species negative control. Probe 'a' is an M13 probe with a fragment of the *S.cerevisiae* actin gene serving as a positive transcription control.

A



B

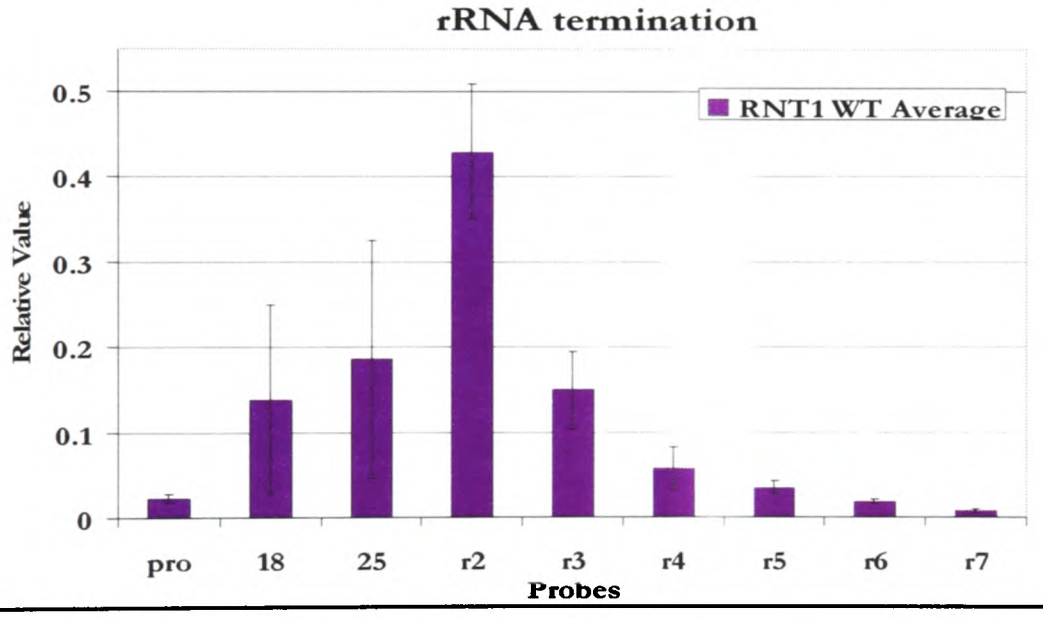


Figure II 1.7 Quantitation of rRNA gene termination in strains (A) Quantitation of MWK028 membrane shown in Figure II 1.6A and other repeated experiments. The average of all six experiments performed was calculated with the standard deviation indicated by error bars. (B) Quantitation of *RNT1* membrane shown in Figure II 1.6B and other repeated experiments. The average of all four experiments performed was calculated with the standard deviation indicated by error bars. Signals were quantified in a Phosphoimager (Molecular Dynamics). The values were corrected for background hybridisation and "U" content in each probe. Signals were normalized to the actin positive transcription probe "a" on each filter to allow direct comparison of transcription levels between membranes.

In the MWK028 strain, the negative control 'm' used for all previous *S.cerevisiae* TRO picked up large transcription signals (data not shown). These were eliminated when a *S.pombe* genic sequence was inserted in the otherwise unbroken poly linker. This suggests that the intact pUC18 linker in the single stranded M13 control was hybridising to the radiolabeled RNA. The elimination of the cross hybridisation with the addition of an insert suggests that the cross hybridisation problem does not significantly affect the other M13 probes containing inserts.

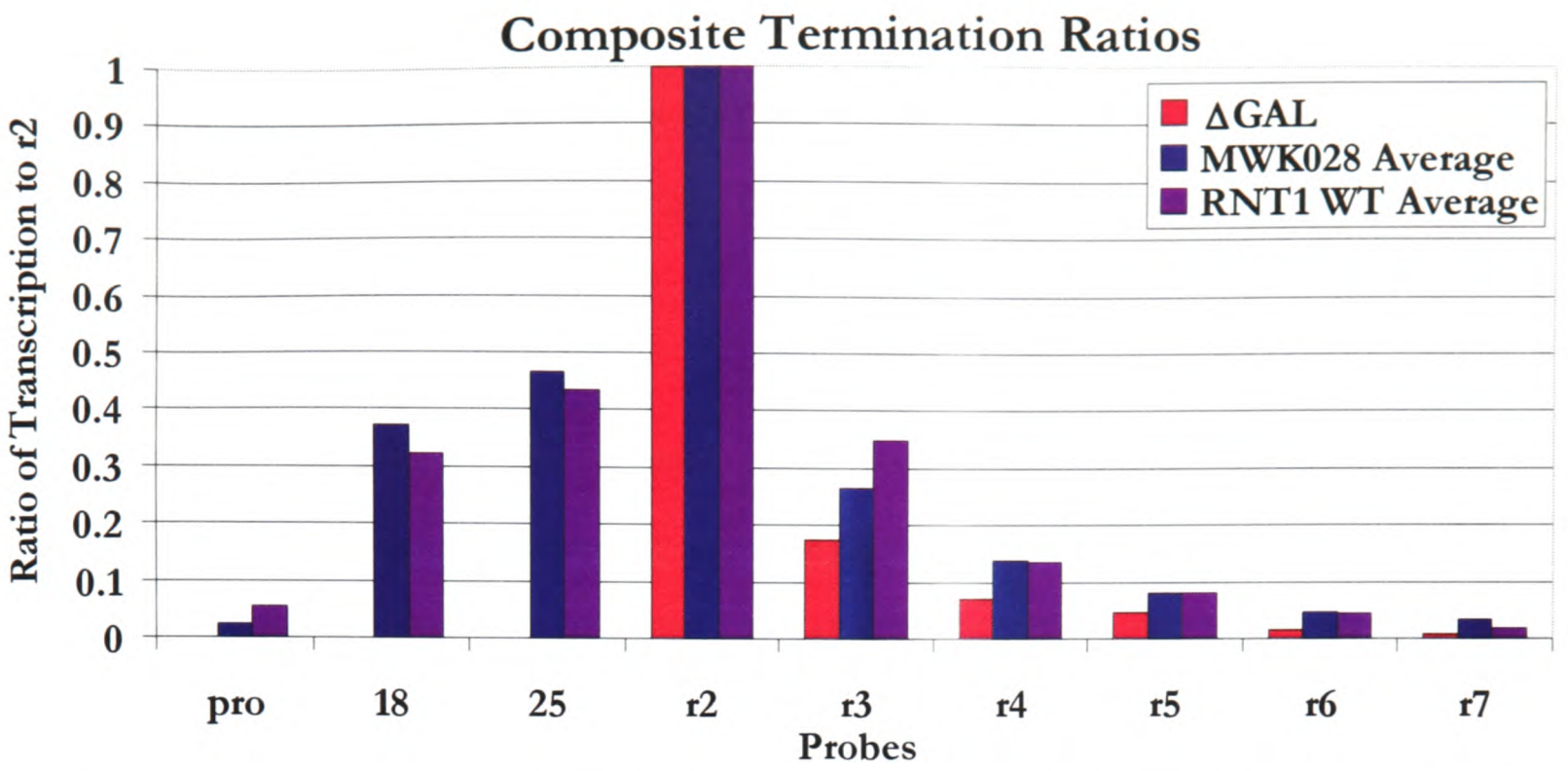


Figure II 1.8 Composite graph of ratios of rRNA gene termination. Ratio of the average transcription termination at each probe to the average level of transcription over probe 'r2' of that strain for strains wild type strains *ΔGAL* (Figure II 1.5B), MWK028 and *RNT1*.

It should be noted the genic probes ('18S' and '25S') and the promoter proximal probe ('pro') are included in these data. All four experiments with *RNT1* were performed with these probes but only three of the six times that TRO was performed on MWK028 were they included. As noted in the Materials section, the '18S', '25S' and 'pro' probes are larger and contain many more radiolabeled U's (54, 61 and 62 respectively) than the termination region probe 'r2' (only 29 U's) resulting in lower relative values when quantitated.

In strain MWK028 about 80% of the transcripts terminate before probe 'r4', which is slightly less than what is seen in *ΔGAL*. The level of transcription extending beyond probe 'r4' is less than 10% of total level of transcription seen at probe 'r2'. In *RNT1* the detectable level of transcription over 'r3' is over 30%, which is much higher than seen in either of the earlier strains. However, the ratio of transcripts seen over 'r4' is about the same as that seen in MWK028 and only slightly elevated from *ΔGAL*. The level of transcription seen over probe 'r2' indicates that there is an accumulation of polymerase density over this region. This signal intensity is indicative of polymerase pausing, which has been reported to occur around the Reb1p binding site. Additionally, both strains *RNT1* and MWK028 appear to have a slightly elevated level of transcription over probes 'r7' and 'pro' relative to *ΔGAL*. This could reflect a low level of strain specific cross hybridisation or, possibly,

read-through transcription that extends through the intergenic spacer and is detected by the promoter proximal probe. However, in *RNT1* the ratio seen over 'pro' is higher than that detected over probe 'r7' suggesting that if there is read through transcription occurring it may build up just upstream of the promoter.

In all strains, only 5-15% of nascent transcription continued beyond probe 'r3' indicating that the majority of transcripts were terminated prior and most likely as a result of the Reb1p terminator previously described (Reeder *et al.* 1999). Additional termination occurs before 'r6' in a region consistent with the +250 'failsafe' terminator. Interestingly, despite the high variation in the nascent transcription detected by the genic probes (as seen by the large standard deviations seen in Figure II 1.5A, II 1.7A and II 1.7B) the ratio of the averages for both strains studied are very similar suggesting that the differences balance out when compiled (discussed in more detail below). Regardless, these data support the steady state results presented previously with a small strain specific variation in the quantity of transcription that passes each termination element.

Discussion

Termination of the rRNA genes at the nascent level occurs in a short and conserved region of the rDNA located downstream of the mature 3' end of the 25S gene. Previous *in vitro* studies showed termination occurring at various different locations. These data support the *in vivo* steady state data presented by Reeder *et al.* (1999) that about 90% of transcription terminates at the +93 primary terminator with additional transcripts termination consistent with the +250 'failsafe' terminator.

The low level of signal detected over probe 'pro' suggests that there may be transcription upstream of the Pol I promoter. One possibility for this occurrence is that not all transcripts are terminating before encountering the next rDNA repeat. The low level of transcription seen over 'r7' also supports this view. However, these experiments cannot rule out transcription generated by any cryptic Pol II promoters known to be located in the IGS. Whether the result of read-through or separately initiated Pol II transcripts, this promoter proximal transcription could play a role in further rounds of initiation or in chromatin remodelling.

As the error bars indicate, there is a larger variation in the level of transcription between different TRO experiments over the genic rRNA probes than is seen over the termination region probes. It is known that the level of rRNA transcription varies with growth and thus was likely a contributing factor to the genic variation resulting from differences in cell cultures at time of TRO. Interestingly, the level of transcription seen across the genic region was usually consistent for both probes within the same TRO experiment, in other words if '18S' was at a low level so was '25S' again suggesting a cell culture effect rather than variation between each probe. The decreased standard deviation seen over the termination region further reinforces the efficient transcriptional termination that has been shown previously.

Interestingly, the average termination profile seen in all three strains show an accumulation of polymerases density over the 'r2' probe. This probe covers the very end of the mature 25S and the region up to the 26 nt down stream of the Reb1p terminator. This build up of polymerases would be consistent with the Reb1p induced pausing, however, the length of the probe makes if difficult to isolate where the potential pausing event occurs. Additionally, the '25S' probe generally detected slightly higher levels of transcription than the '18S' probe. The '25S' probe is located about 500 bp upstream from the mature 3' end of the 25S gene and about 450 bp away from the 5' end of the 'r2' probe. The increase in transcription seen over the '25S' probe could be the result of a large backlog of elongating polymerases that are stalled behind the terminator protein awaiting release. Regardless, this effect is most likely not the result of selective hybridisation. As seen in the synthetic transcription, the '18S' probe appears to hybridise well if not better than the '25S' probe. It is difficult to make quantitative statements from these *in vitro* transcribed RNA as they were made on separate templates but it is clear that both probes work efficiently.

Part II - Section 2

Deletion of Pol I subunit Rpa12p eliminates termination

As shown in the previous section, about 90% of Pol I transcripts terminate at the primary Reb1p terminator at about +93 relative to the mature 3' end of 25S. The remainder of polymerases terminate consistent with a 'failsafe' terminator at about +250. With the site of nascent termination established, the effects of a Pol I subunit deletion were investigated for its effects on the termination in rRNA genes.

This section investigates what role Rpa12p plays in the proper termination of Pol I transcripts in rRNA genes. For reasons described above, it was hypothesized that the Rpa12p subunit of Pol I may be essential for Pol I transcription termination. As overexpression of Rpa190p had been reported to partially compensate for the phenotype of Δ*RPA12*, it was speculated that it might be able to restore termination in the Δ*RPA12* strain.

2.1 Effect of *RPA12* disruption on Pol I termination

The *S.cerevisiae* strain RLY10 has the *RPA12* gene completely deleted by a *HIS3* insertion in MWK028 causing poor growth at 36^0 but normal growth at 21^0 (Nogi *et al.* 1993). RLY10 was grown at 21^0 and TRO analysis was performed to examine the state of Pol I rRNA gene termination. The experiment was repeated six times using probes 'r2'-'r7' and four times with the additional probes 'pro', '18S' and '25S'. Figure II 2.1 illustrates a typical slot blot hybridisation for one experiment and Figure II 2.2A shows the average and standard deviation. As before, quantification was corrected for background, adjusted for the number of U's per probe and the level of actin. The average level of transcription at each probe was compared to the average transcription at probe 'r2' to generate a ratio of termination in Figure II 2.2B

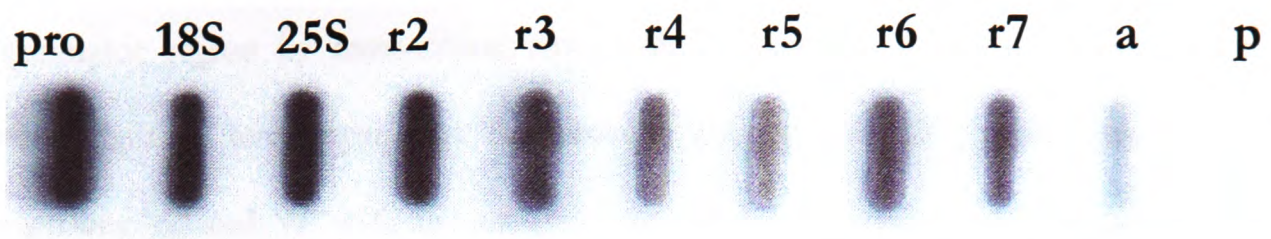


Figure II 2.1 Transcription over rRNA genes in strain RLY10. A picture of a typical hybridisation of radiolabeled rRNA gene transcripts of *S.cerevisiae* strain RLY10 (Δ*RPA12*) to membrane. Probe 'p' containing a fragment of the *Schizosaccharomyces* pombe ura4 gene was used as a cross species negative control. Probe 'a' is an M13 probe with a fragment of the *S.cerevisiae* actin gene serving as a positive transcription control.

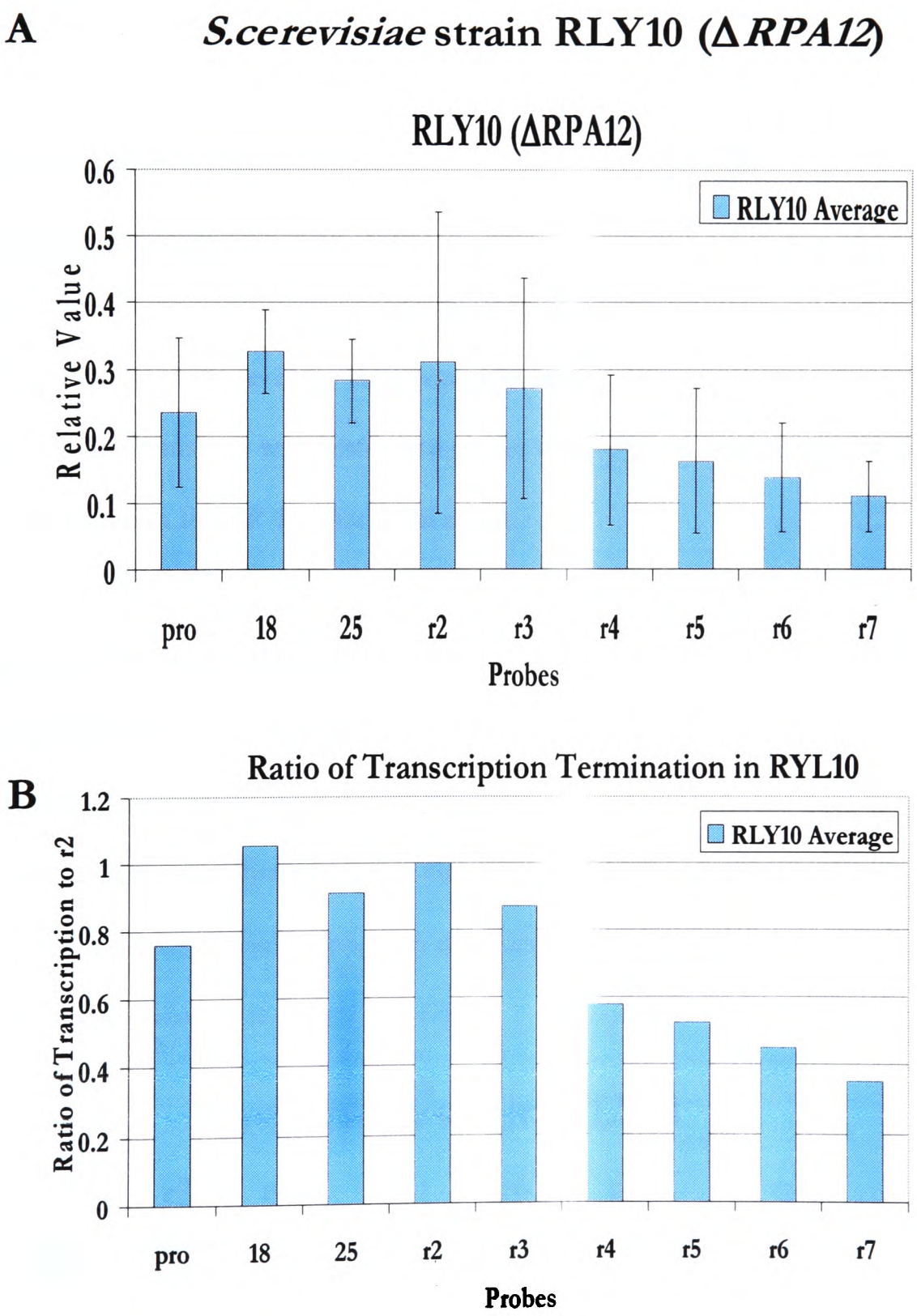


Figure II 2.2 Quantitation of strain RLY10 rRNA gene termination (A) Quantitation of membrane shown in Figure II 2.1 and other repeated experiments with RLY10 (Δ*RPA12*). The average of all six experiments performed was calculated with the standard deviation indicated by error bars. Signals were quantified in a Phosphoimager (Molecular Dynamics). The values were corrected for background hybridisation and 'U' content in each probe. Signals were normalized to the actin positive transcription probe 'a' on each filter to allow direct comparison of transcription levels between membranes. (B) Ratio of the average transcription at each probe to the average level of transcription over probe 'r2'.

From these data it can be seen that the disrupted *RPA12* strain (RLY10) allows read-through of the Pol I terminator region by transcribing past both the primary Reb1p terminator as well as the downstream 'failsafe' terminator. In the parental strain (MWK028), most transcripts terminated between probes 'r2' and 'r4' with signals dropping nearly to background by probe 'r6'. In marked contrast, disrupting *RPA12* allows significant read-through. Also, the accumulation of polymerase density seen over 'r2' in the wild type is not present in RLY10 (see Discussion). Additionally, it appears that transcription can be detected before the promoter of the next transcription unit as seen by the signals over the promoter proximal probe ('pro').

2.2 Over expression of Rpa190p does not restore termination

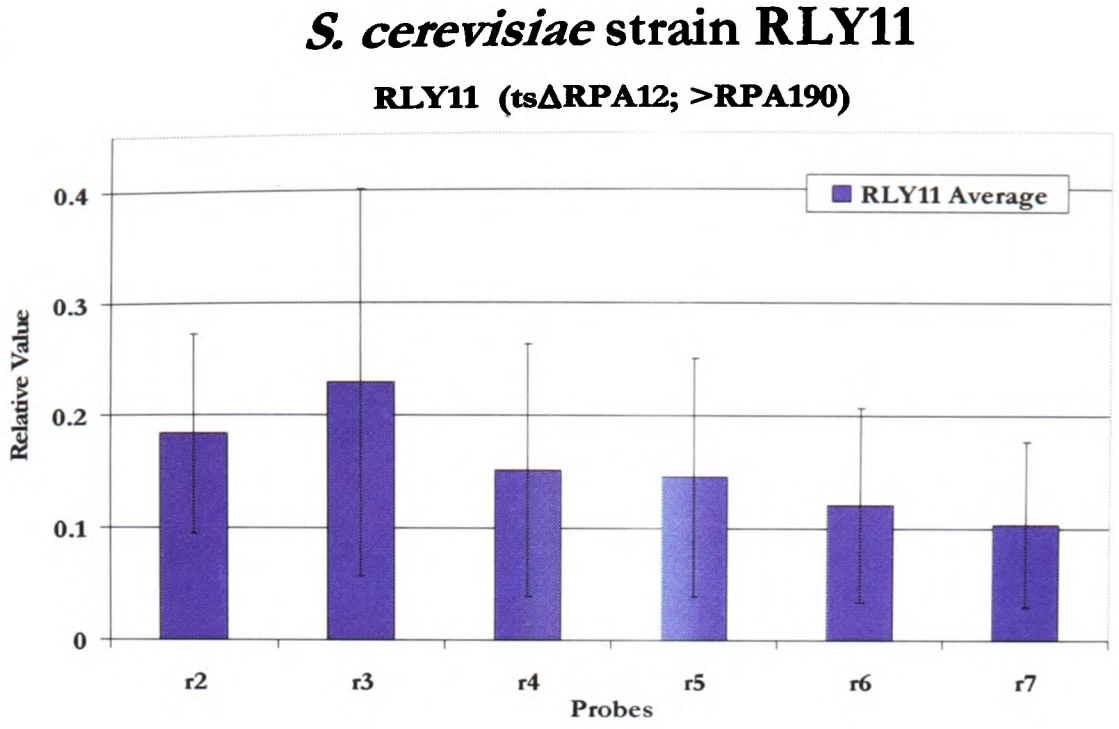
The RLY10 strain was transformed with a high copy yeast plasmid expressing the *RPA190* gene under the control of its own promoter to create strain RLY11. It has been reported that expression of additional Rpa190p partially reverses the phenotype caused by the deletion of the *RPA12* gene (Nogi *et al.* 1993). TRO analysis was performed to see if overexpression of this subunit is able to restore Pol I termination in the *RPA12* disrupted strain.

RLY11 was grown at 21^0 and TRO analysis was performed as before. The experiment was repeated three times for probes 'r2'-'r7'. Figure II 2.3 illustrates a typical slot blot hybridisation for one experiment and Figure II 2.4A shows the average and standard deviation. As before, quantification was corrected for background, adjusted for the number of U's per probe and the level of actin. The average level of transcription at each probe was compared to the average transcription at probe 'r2' to generate a ratio of termination in Figure II 2.4B.



Figure II 2.3 Transcription over rRNA genes in strain RLY11. A picture of a typical hybridisation of radiolabeled rRNA gene transcripts of *S.cerevisiae* strain RLY10 (Δ*RPA12*; *RPA190*) to membrane. Probe 'p' containing a fragment of the *Schizosaccharomyces* pombe ura4 gene was used as a cross species negative control. Probe 'a' is an M13 probe with a fragment of the *S.cerevisiae* actin gene serving as a positive transcription control.

A



B

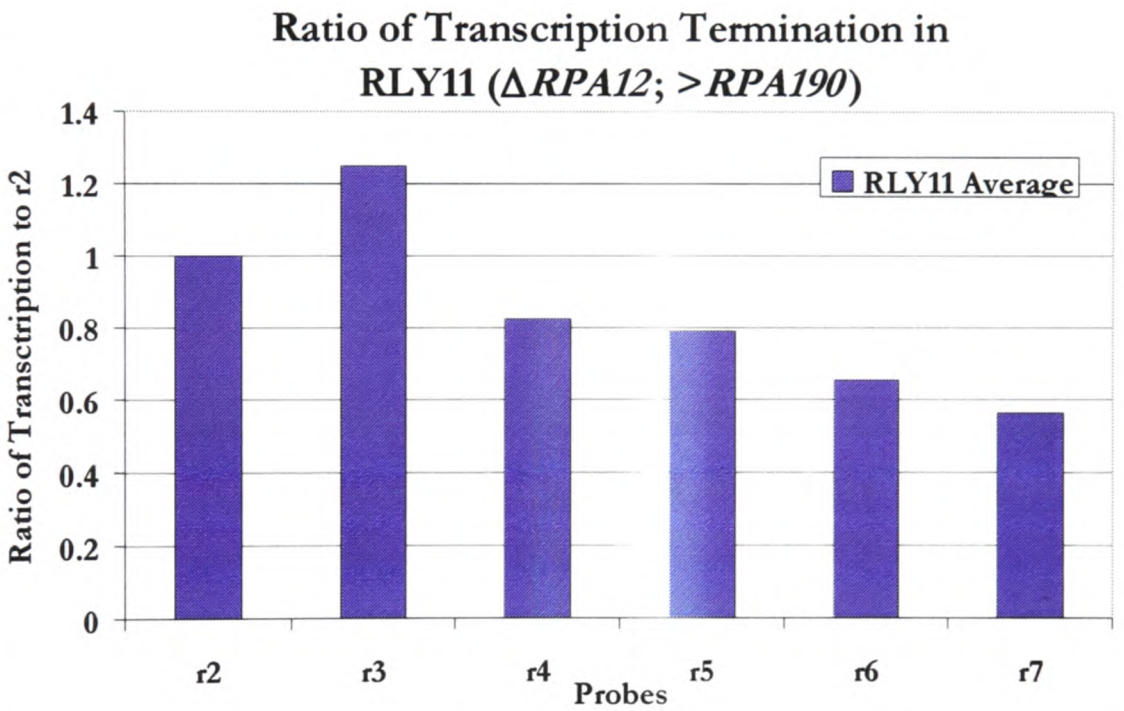


Figure II 2.4 Quantitation of strain RLY11 rRNA transcription termination. (A) Quantitation of membrane shown in Figure II 2.3 and other repeated experiments with RLY11 (Δ*RPA12*; *RPA290*). The average of all 3 experiments performed was calculated with the standard deviation indicated by error bars. Signals were quantified in a Phosphoimager (Molecular Dynamics). The values were corrected for background hybridisation and 'U' content in each probe. Signals were normalized to the actin positive transcription probe 'a' on each filter to allow direct comparison of transcription levels between membranes. (B) Ratio of the average transcription at each probe to the average level of transcription over probe 'r2'.

The termination profile seen from overexpression of Rpa190p in the Δ*RPA12* strain was similar to the Δ*RPA12* strain alone. Restoration of Pol I transcriptional termination does not appear to be the reason for improved growth.

2.3 Composite of MWK028, RLY10 and RLY11

When TRO analysis was performed, the wild type MWK028 strain was analysed at the same time as the mutant strains in order to allow a better comparison. Figure II 2.5A shows a graph comparing

the average and standard deviation for transcription in all three strains. Figure II 2.5B shows the ratio of transcription termination to probe 'r2' for each strain as presented earlier.

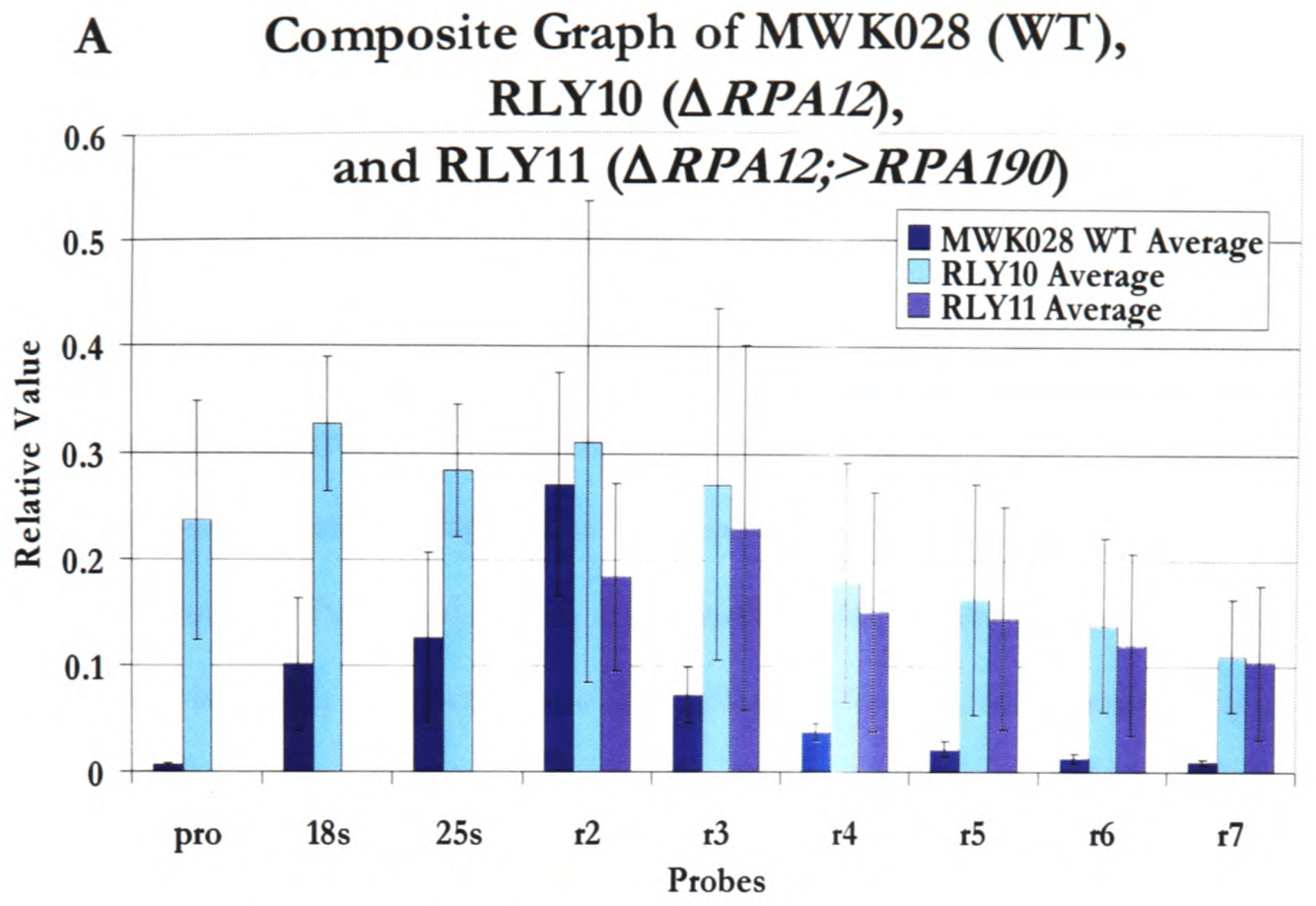


Figure II 2.5 Composite of termination in MWK028, RLY10, RLY11. (A) Composite of graphs as presented previously for MKK028 (Figure II 1.8A), RLY10 (Figure 2.2A), and RLY11 (Figure II 2.4A) (B) Ratio of the average transcription termination at each probe compared to the average level of transcription at probe 'r2' as presented earlier for strain MKK028 (Figure II 1.8B), RLY10 (Figure II 2.2B), and RLY11 (Figure II 2.4B).

The bar graphs in Figure II 2.5 represents the average and consistently seen profiles. The nascent transcription signals from strain RYL10 and RYL11 were much more intense then any of the previously studied wild type strains as can be seen by the relative value for each probe. This suggests

that there was an increase in the level of transcription of the rRNA genes in the *RPA12* disrupted strains. The elevated level of transcription is consistent with data observed from deletions in the Pol II homolog B12.6 as observed by Awrey *et al.* (1997). While it appears that RLY11 may have a slightly elevated level of transcription over RLY10 it is difficult to make direct comparisons without the genic and promoter proximal probes for reference.

Interestingly, from the ratios seen in Figure II 2.5B it appears as if the pausing observed over probe 'r2' in the isogenic wild type strain MWK028 is not occurring in RLY10. The decreased recognition of pause signals in absence of Rpa12p is consistent with the data shown for deletions of the homologous subunits in Pol II and Pol III. Additionally, in RLY10 average transcription over probe '18S' appears to be higher than over probe '25S' giving further support to some pausing occurring over probe '25S' in the wild type. Interestingly, the overexpression of Rpa190p in RLY11 appears to partially recover pausing but further downstream than in the wild type, over probe 'r3' rather than over probe 'r2'.

It should also be noted that there was a greater degree of variability (depicted by the error bars) in the level of transcription in RLY10 and RYL11 than seen over most probes in the wild type strain (MWK028). This can be partially attributed to the fact that the strains were consistently less healthy than the wild type strain resulting in decreased TRO efficiency.

Discussion

Overexpression of Rpa190p, reported to correct for growth defects seen in the *RPA12* deletion (Van Mullem *et al.* 2002), does not recover growth by restoring proper termination of the rRNA genes. Additionally, it appears that efficient termination of Pol I transcribed rRNA genes is not essential for viability with functional rRNA gene products generated from the non-terminated pre-rRNA polycistronic 35S transcript

These results support previous research that yeast can grow, albeit more slowly, when rRNA genes are transcribed solely by Pol II, which neither starts nor stops at the same locations as when

transcribed by Pol I (Nogi *et al.* 1991). This conclusion is also supported by a recent report in where both the Reb1p and failsafe terminator sites were removed from chromosomal rRNA gene copies with no detectable effect on growth or the overall rate of rRNA gene synthesis (Wai *et al.* 2001). The clear implication is that *S.cerevisiae* is able to process functional rRNA genes from any transcript that contains the necessary processing sequences.

It has been demonstrated previously that read-through by Pol I can dislodge transcription factors from downstream Pol I promoters and prevent initiation (Bateman and Paule 1988; Henderson *et al.* 1989). This suggests that a primary purpose of Pol I termination is to prevent such promoter occlusion. Because Pol I terminators are relatively simple, composed of a pause inducing protein and a T-rich release region, it may be that multiple cryptic sites in the yeast intergenic spacer can act to prevent occlusion when the primary terminator is deleted. In support of this idea, Banditt *et al.* (1999) were unable to detect any read-through into the downstream repeat when the enhancer region containing both Reb1p and the failsafe terminator was deleted. However, in Pol I without Rpa12p, nascent transcription is detected over the promoter proximal probe 'pro' suggesting that termination is not occurring at either the primary or any cryptic sites in the intergenic region. This read-through transcription would be consistent with a Rpa12p deficient Pol I that does not efficiently detect pause elements.

All of this suggests that Rpa12p plays multiple roles in rRNA gene transcription. It is likely that Rpa12p is involved in maintaining the overall structure of Pol I. In the deletion strain, this instability may be partially compensated for by increasing the concentration of Rpa190p. However, Rpa12p also plays a second important role in rRNA gene termination with read-through transcription increased significantly in its absence. This cannot be over come with the over expression of Rpa190p but may or may not cause significant growth defects. By comparing these results to those seen in the deletion of the homologous subunits in Pol II and Pol III several mechanisms present themselves. First, the defect in termination could result from an increased rate of transcription that does not allow a long enough dwelling time for termination signal recognition. Alternatively, *ΔRPA12* Pol I may be unable to detect termination signals at any rate of transcription. A third

possibility is that Rpa12p, as has been shown with the homologous Pol II and Pol III subunits, might possess an intrinsic cleavage activity (Awrey *et al.* 1997; Chedin *et al.* 1998). This presents the possibility that Pol I lacking Rpa12p is able to detect pausing but is not able to cleave and release the transcripts and therefore prevents termination.

In determining the validity of these mechanisms the ability and rate of nascent pausing is important. In the wild type construct, there is detectable pausing over probe 'r2' which is presumably the result of Pol I recognition of the Reb1p pause and termination element. In RLY10 (*ΔRPA12*) there is an absence of pausing seen over probe 'r2' suggesting that there is no recognition of the pause element. Additionally, in RLY11 (*ΔRPA12*; >RPA190), the ability to pause near the Reb1p binding site appears to be restored albeit at a position further downstream from that seen in the wild type strain. It could be that in absence of Rpa12p, Pol I is still capable of pausing but is prevented from doing so either through an inability to efficiently detect or act on pause signals.

Part II - Section 3

Deletion of *RNT1* on rRNA gene termination

A *S.cerevisiae* strain with a deletion in Rnt1p, predicted to be involved in co-transcriptional processing, was studied to further investigate nascent termination in rRNA gene transcription. Previous steady state data from strains lacking functional Rnt1p generated transcripts with 3' ends between +195 and +250 relative to the mature 3' end of 25S (Elela *et al.* 1996; Kufel *et al.* 1999; Reeder *et al.* 1999). This was interpreted as indicating the true site of termination (Elela *et al.* 1996; Kufel *et al.* 1999). However, in these studies, the nascent termination profile of a strain lacking Rnt1p (*rnt1*) has been compared to an isogenic wild type strain. From these TRO it can be seen that in absence of Rnt1p, termination occurs further downstream than seen in the wild type strain and extends beyond the reported +250 'failsafe' terminator. This suggests a direct role of Rnt1p in Pol I termination.

3.1 Rnt1p deletion strain termination profile

To investigate the effect of deleting the *RNT1* gene on the termination of rRNA genes in *S.cerevisiae*, TRO analysis was performed on strain *rnt1* in which the *RNT1* gene is deleted by a *TRP1* insertion (Kufel *et al.* 1999). Note that the TRO analysis of the wild type strain (*RNT1*) was described previously in Part II Section 1.

The TRO experiment was repeated four times with Figure II 3.1 illustrating a typical slot blot hybridisation for one experiment and Figure II 3.2A showing the average and standard deviation of all experiments. As before, quantification was corrected for background, adjusted for the number of U's per probe and the level of actin. The average level of transcription at each probe was compared to the average transcription at probe 'r2' to generate a ratio of termination in Figure II 3.2B.

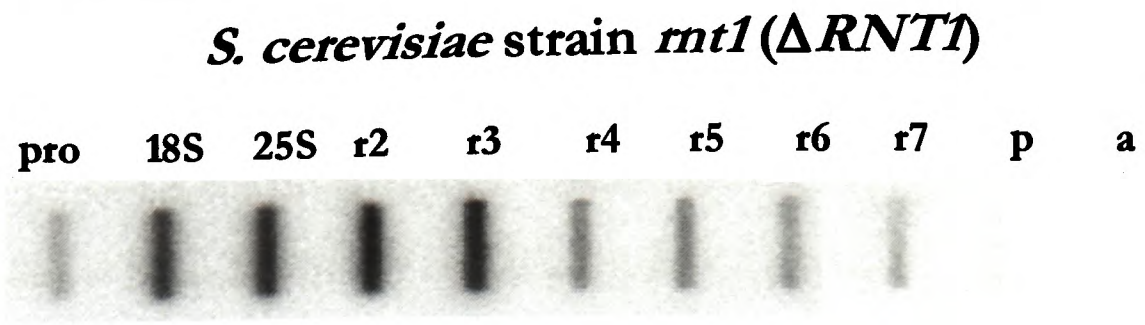


Figure II 3.1 Transcription over rRNA genes in strain *rnt1*. A picture of typical hybridisation of radiolabeled rRNA gene transcripts of *S.cerevisiae* strain *rnt1* to membrane. Probe 'p' containing a fragment of the *Schizosaccharomyces pombe* ura4 gene was used as a cross species negative control. Probe 'a' is an M13 probe with a fragment of the *S.cerevisiae* actin gene serving as a positive transcription control.

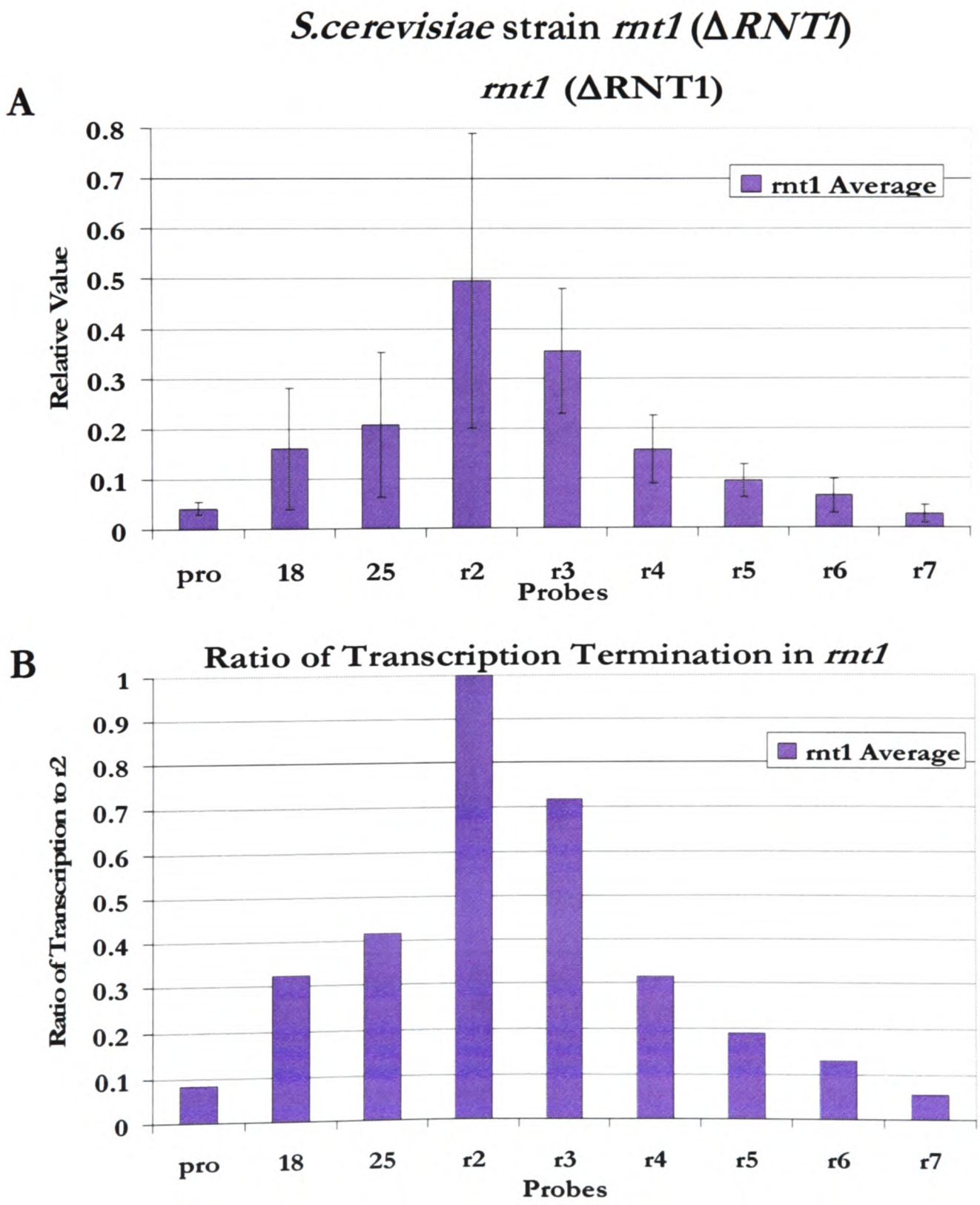


Figure II 3.2 Quantitation of strain *rnt1* rRNA gene termination (A) Quantitation of membrane shown in Figure II 3.2 and other repeated experiments. The average of all 4 experiments performed was calculated with the standard deviation indicated by error bars. Signals were quantified in a Phosphoimager (Molecular Dynamics). The values were corrected for background hybridisation and 'U' content in each probe. Signals were normalized to the actin positive transcription probe 'a' on each filter to allow direct comparison of transcription levels between membranes. (B) Ratio of the average transcription at each probe to the average level of transcription over probe 'r2'.

For better comparison with the isogenic wild type strain, *RNT1* (WT) and rnt1 (Δ*RNT1*) are shown in composite graphs below. The average and standard deviation of all four experiments with both strains are shown in Figure II 3.3A with the ratios as described previously shown in Figure II 3.3B.

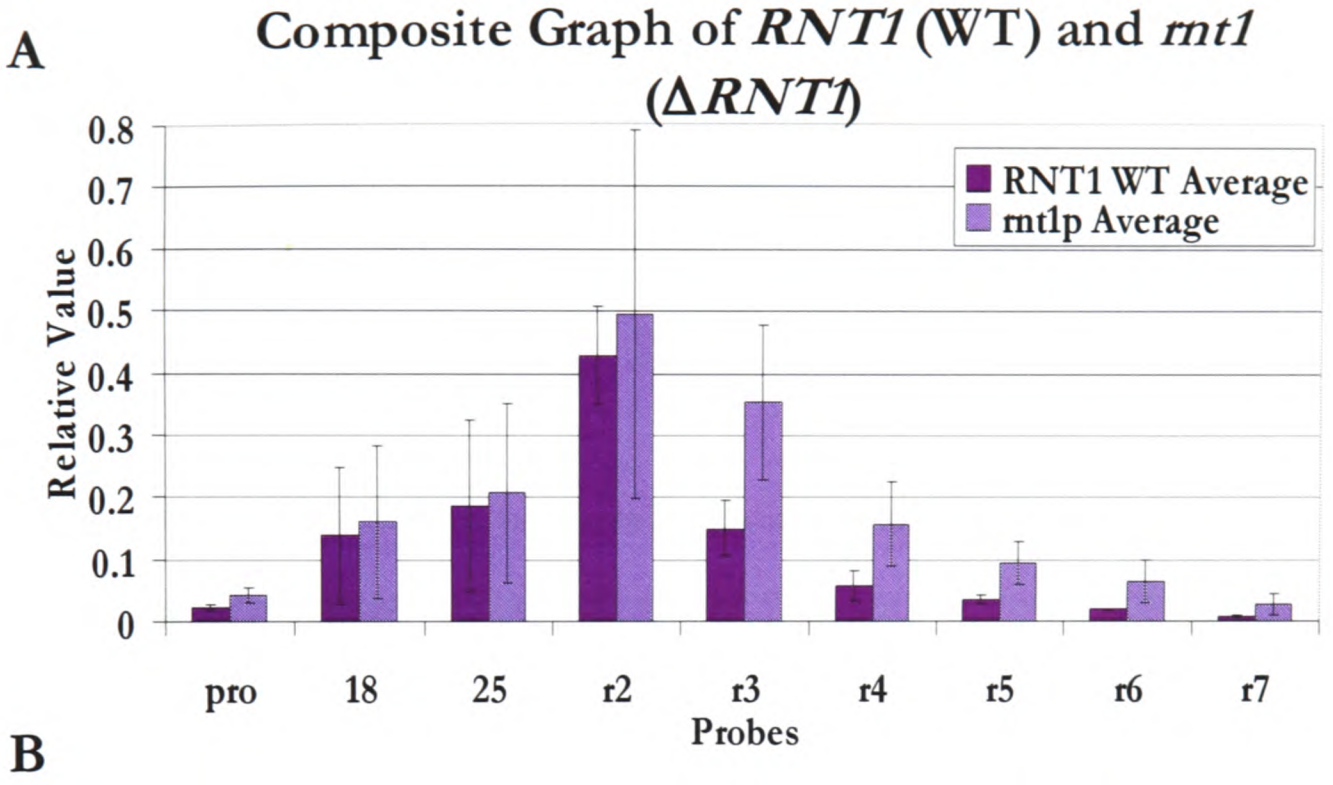


Figure II 3.3 Composite graph of termination in *RNT1* and *rnt1* (A) The average and standard deviation of transcription at each probe for *RNT1* (WT) and *rnt1* (Δ*RNT1*). Individual graphs were shown and described in more detail for *RNT1* (Figure II 1.8B) and *rnt1* (Figure II 3.3A). (B) Ratio of the average transcription termination at each probe compared to the average level of transcription at probe 'r2' of that strain for *RNT1* (WT) and *rnt1* (Δ*RNT1*). Individual graphs were shown and described in more detail for *RNT1* (Figure II 1.9) and *rnt1* (Figure II 3.3B).

From these data it appears that nascent transcript processing by Rnt1p is not required for transcription termination in rRNA genes. However, as seen in Figure II 3.3B there is a 2-fold increase in transcription over probes 'r3'-'r7' in *rnt1* compared to the wild type suggesting that transcribing polymerases are proceeding further downstream than seen in the wild type strain.

About 30% of transcripts pass the Reb1p primary terminator (located in probe 'r3') and continue on to be detected over probe 'r4'. This is more than a two-fold increase from the 12% of transcribing polymerase that continue until probe 'r4' in the wild type strain (Figure II 3.3B). Additionally, almost 20% of polymerases continue beyond the +250 'failsafe' terminator as seen by the signals over probe 'r5' which is more than twice what is seen in *RNT1* (WT). This suggests that nascent termination does occur in the absence of Rnt1p but that the kinetics of the process are significantly altered resulting in an increase in downstream transcription.

Discussion

In some respects, the nascent data seen here is consistent with the steady state data seen before in strains lacking Rnt1p (Elela *et al.* 1996; Kufel *et al.* 1999; Reeder *et al.* 1999). rRNA gene transcription in the *rnt1* strain proceeds beyond the termination sites identified in the wild type strains at the steady state (Elela *et al.* 1996; Kufel *et al.* 1999; Reeder *et al.* 1999) and nascent level (this study). However, the steady state 3' end sites identified by Elela *et al.* (1996) and Kufel *et al.* (1999) as the sites of termination are not consistent with what has been observed here at the nascent level.

The *in vivo* steady state data from a temperature sensitive strain of *RNT1 (rnt1-1)* showed new 3' end transcripts developing after 12 hours at restrictive temperature at the positions +195 and +210 relative to the mature 3' end of the 25S gene (Elela *et al.* 1996). However, in addition to a few other products that were not identified, there was also an observed accumulation of full-length transcript that extends to +284 where the probe used in the protection assay ends. (In relation to the probes used in this study, the +284 probe covers the first 30 nt of the probe 'r5' and covers all but the last 41 nt of probe 'r4'). Reeder *et al.* (1999) also performed RNase protection analysis on RNA isolated from the temperature sensitive strain *rnt1-1*. These data indicated that in the absence of Rnt1p a larger portion of transcripts read-through the +93 primary terminator but do not continue to transcribe beyond the +250 failsafe terminator. A separate attempt by Kufel *et al.* (1999) to identify the 3' end of rRNA gene transcription in the *rnt1* strain used primer extension analysis. In these experiments, transcripts were detected with a probe that hybridises from +162 to +180 but were not detected with a probe overlapping the +250 'failsafe' terminator (+246 to +264) and the very

inning of probe 'r5'. While there is some discrepancy between these different experiments, all se data suggest that the location of 3' ends in the absence of Rnt1p is further downstream than in wild type but generally do not extend beyond +250 relative to the 3' end of mature 25S. In trast, it is clear from the Figure II 3.3A and Figure II 3.3B that there is detectable nascent scription over probe 'r5' in *rnt1* relative to the wild type (*RNT1*). Therefore, nascent scription does appear to extend beyond the sites identified by Elela *et al.* (1996), Reeder *et al.* 99) and Kufel *et al.* (1999).

re are two different scenarios for what the cause of increased polymerase density seen beyond probe 'r2' in the *rnt1* strain. First, the absence of Rnt1p causes a defect in Pol I termination so a higher percentage of the transcripts pass the normal termination signals. However, other tic signals may be recognized further 3' that induce termination. This scenario suggests that a ion of the transcribing polymerases would be induced to pass termination signals due to the nce of an enzyme which functions on the secondary structure of the RNA transcript. A second ario has the polymerases still recognizing the normal termination signals and consequently ing. These paused polymerases may be waiting for a cleavage event to release the transcript and inate. In the absence of Rnt1p the paused polymerase would fail to terminate efficiently due to of Rnt1p cleavage.

ny case, in the absence of Rnt1p termination does occur albeit further downstream than rved in the wild type. It may be that the polymerase encounters and recognizes downstream ination sites like the +250 'failsafe' terminator that induce transcript release. Alternatively, there be other enzymes able to cleave in absence of Rnt1p thereby, possibly, inducing termination. ever, the results presented in Section 2 and other experiments (Allmang and Tollervey 1998) :st that termination is not required for rRNA gene maturation from the pre-rRNA transcript. It ssible that in the absence of Rnt1p, other protein or protein complexes compensate to cleave iascent transcript. If this were the case, the different 3' end sites identified by different iments in the absence of functional Rnt1p could be processing intermediates generated by enzymes.

Conclusion

The aim of this thesis was to present a functional analysis of read-through transcription in *S.cerevisiae* genes transcribed by Pol I and Pol II. TRO analysis was employed to investigate nascent transcription of genes transcribed by both of these polymerases. In Part I, it was demonstrated that concurrent convergent transcription results in transcriptional interference. This TI effect, termed transcriptional collision, shows that proper transcriptional termination and regulation is of vital importance in high-density genomes such as that of *S.cerevisiae.* In Part II, the location of nascent termination in rRNA genes was identified. TRO analysis of mutant yeast strains, which lack factors involved in the Pol I transcription machinery and processing, indicate roles for these factors in transcription termination. While each section has been discussed separately, there are a few points of interest that incorporate both parts of this research.

First, from the results obtained in Part II it appears that Rnt1p plays a role in transcriptional termination of Pol I. The previously reported steady state data (Elela *et al.* 1996; Kufel *et al.* 1999; Reeder *et al.* 1999) demonstrated that 3' end products generated in the absence of Rnt1p were different than those seen in wild type strains. The nascent data presented here confirm these data but further suggest that the defect is at the level of nascent termination rather than merely the absence of processing as had been suggested. This indicates that transcription by Pol I may require efficient processing to achieve proper nascent termination. Interestingly, this is reminiscent of transcriptional termination by Pol II, which is highly dependent on 3' end cleavage/polyadenylation to achieve efficient termination. However, Pol I transcriptional termination still occurs in the absence of Rnt1p suggesting that it may increase efficiency of termination or its role can be played by other factors if necessary. This is in contrast to termination by Pol II, which is closely coupled to processing events and is abolished in the absence of 3' end processing factors.

Another interesting conclusion from these experiments is that TI has different effects on Pol I and Pol II transcribed genes. In Pol II, interference can play an important role in impeding efficient transcription. Whether through preventing initiation (promoter occlusion) or elongation

(transcriptional collision), TI can greatly alter the quantity of functional transcripts generated from Pol II transcribed genes. In contrast, Pol I appears to be less effected by TI. The strains lacking *RPA12* experience severe read-through into the promoter but are still viable. This suggests rRNA gene transcripts generated by Pol I are processed into functional product regardless of the where the transcript begins or ends. The varying impact of TI may be more the result of the fundamentally different processing pathways employed by each polymerase. The processing pathway for rRNA gene maturation involves the utilization of a polycistronic pre-rRNA. Therefore, it is not surprising that any polycistronic transcripts with the appropriate processing signal intact is processed into functional product. However, in the highly regulated processing pathway of mRNA, there is a need for each transcript to be identified as functional before being exported and translated into protein products.

The different impact of TI may be somewhat explained by considering the genes that are transcribed by each polymerase. First, Pol I transcribed genes are located in a large array of rDNA repeats. Any read-through transcription would likely generate an identical product rather than a product that is involved in a separate regulation pathway. In contrast, Pol II transcribed genes are diverse and involve variable regulatory pathways, increasing the probability that an inappropriately transcribed gene could have deleterious effects. Second, large quantities of rRNA gene products are needed so that variation in the number of transcripts may be less significant to the overall viability of the cell. However Pol II transcripts, often highly regulated and found at much lower levels, may be more sensitive to slight variations in transcript number. Unregulated functional transcripts could impact cell processes. Third, any decrease in initiation at a Pol I promoter due to promoter occlusion will have little effect on the overall level of transcript generated if the read through transcript is also able to form functional product. This is not the case for Pol II transcribed genes where the gene that is being interfered with is most likely different from the adjacent gene that is doing the interfering.

In conclusion, similarities exist between transcription complexes for Pol I and Pol II in *S.cerevisiae.* However, as would be expected, there are also many differences that influence the efficient transcription of their respective genes.

Future Prospects

Several aspects of the research described in this thesis can be explored in more detail. In the short term, steady state analysis of the termination region in the mutant Pol I strains presented in Part II will be performed to further characterise the transcriptional defects observed. Additionally, the curious role of *RPA12* in Pol I transcription and termination would be interesting to explore. If overexpression of *RPA190* is able to partially recover the growth defect caused by *RPA12* deletion without restoring termination it suggests that the growth defect results solely from the instability of Pol I. Mutagenesis of the A12 subunit of Pol I may enable separation of the two functions in addition to assessing the impact of the absence of termination on viability. Similar nascent experiments performed on yeast strains lacking the homologous subunits in Pol II and Pol III (B26 and C11 respectively) could generate a better understanding of the role these proteins may also perform in transcriptional termination.

Further investigations into the mechanism of transcriptional collision seen in Part I would also be of interest. It would be valuable to determine if the polymerases are impeding each other directly or through supercoiling effects. Do the elongating Pol II complexes directly collide, so inducing arrest, or are they kept from reaching the middle of the template because of dense supercoiling? Additionally, identification of positions in the yeast genome where transcriptional collision may occur naturally would establish the biological importance of this effect. The few co-transcribed convergent gene pairs so far identified (Kruglyak and Tang 2000) are poorly characterized and could be explored to determine the *in vivo* significance of transcription collision.

References

Abe, A., Y. Hiraoka and T. Fukasawa (1990). "Signal Sequence for Generation of mRNA 3' End in the *Saccharomyces cerevisiae GAL7* Gene." EMBO Journal **9**(11): 3691-3697.

Allmang, C. and D. Tollervey (1998). "The role of the 3' external transcribed spacer in yeast pre-rRNA processing." Journal of Molecular Biology **278**(1): 67-78.

Aprikian, P., B. Moorefield and R. H. Reeder (2000). "TATA binding protein can stimulate core-directed transcription by yeast RNA polymerase I." Molecular and Cellular Biology **20**(14): 5269-5275.

Aprikian, P., B. Moorefield and R. H. Reeder (2001). "New model for the yeast RNA polymerase I transcription cycle." Molecular and Cellular Biology **21**(15): 4847-4855.

Aranda, A., J. E. Perez-Ortin, C. Moore and M. del Olmo (1998 a). "The yeast *FBP1* poly(A) signal functions in both orientations and overlaps with a gene promoter." Nucleic Acids Research **26**(20): 4588-4596.

Aranda, A., J. E. Perez-Ortin, C. Moore and M. L. del Olmo (1998b). "Transcription termination downstream of the *Saccharomyces cerevisiae FPB1* poly(A) site does not depend on efficient 3' end processing." RNA **4**(3): 303-318.

Aranda, A. and N. Proudfoot (2001). "Transcriptional termination factors for RNA polymerase II in yeast." Molecular Cell **7**(5): 1003-1011.

Aranda, A. and N. J. Proudfoot (1999). "Definition of transcriptional pause elements in fission yeast." Molecular and Cellular Biology **19**(2): 1251-1261.

Atkins, D., G. M. Arndt and J. G. Izant (1994). "Antisense Gene Expression in Yeast." Biological Chemistry Hoppe-Seyler **375**(11): 721-729.

Awrey, D. E., R. G. Weilbaecher, S. A. Hemming, S. M. Orlicky, C. M. Kane and A. M. Edwards (1997). "Transcription elongation through DNA arrest sites - A multistep process involving both RNA polymerase II subunit *RPB9* and TFIIS." Journal of Biological Chemistry **272**(23): 14747-14754.

Banditt, M., T. Koller and J. M. Sogo (1999). "Transcriptional activity and chromatin structure of enhancer deleted rRNA genes in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **19**(7): 4953-4960.

Barabino, S. M. L. and W. Keller (1999). "Last but not least: Regulated poly(A) tail formation." Cell **99**(1): 9-11.

Barker, D. G., J. H. M. White and L. H. Johnston (1987). "Molecular Characterization of the DNA Ligase Gene, Cdc17, from the Fission Yeast *Schizosaccharomyces pombe*." European Journal of Biochemistry **162**(3): 659-667.

Bateman, E. and M. R. Paule (1988). "Promoter Occlusion During rRNA Transcription." Cell **54**(7): 985-992.

Beelman, C. A. and R. Parker (1995). "Degradation of mRNA in Eukaryotes." Cell **81**(2): 179-183.

Berger, S. L. (2001). "An embarrassment of niches: the many covalent modifications of histones in transcriptional regulation." Oncogene **20**(24): 3007-3013.

Birse, C. E., B. A. Lee, K. Hansen and N. J. Proudfoot (1997). "Transcriptional termination signals for RNA polymerase II in fission yeast." EMBO Journal **16**(12): 3633-3643.

Birse, C. E., L. Minvielle-Sebastia, B. A. Lee, W. Keller and N. J. Proudfoot (1998). "Coupling termination of transcription to messenger RNA maturation in yeast." Science **280**(5361): 298-301.

Blackwood, E. M. and J. T. Kadonaga (1998). "Going the distance: A current view of enhancer action." Science **281**(5373): 60-63.

Blattner, F. R., G. Plunkett, C. A. Bloch, N. T. Perna, V. Burland, M. Riley, J. ColladoVides, J. D. Glasner, C. K. Rode, G. F. Mayhew, J. Gregor, N. W. Davis, H. A. Kirkpatrick, M. A. Goeden, D. J. Rose, B. Mau and Y. Shao (1997). "The complete genome sequence of Escherichia coli K-12." Science **277**(5331): 1453-1474.

Bordi, L., F. Cioci and G. Camilloni (2001). "In vivo binding and hierarchy of assembly of the yeast RNA polymerase I transcription factors." Molecular Biology of the Cell **12**(3): 753-760.

Borggrefe, T., R. Davis, A. Bareket-Samish and R. D. Kornberg (2001). "Quantitation of the RNA polymerase II transcription machinery in yeast." Journal of Biological Chemistry **276**(50): 47150-47153.

Brown, S. A., A. N. Imbalzano and R. E. Kingston (1996). "Activator-dependent regulation of transcriptional pausing on nucleosomal templates." Genes & Development **10**(12): 1479-1490.

Buratowski, S. (2000). "Snapshots of RNA polymerase II transcription initiation." Current Opinion in Cell Biology **12**(3): 320-325.

Burge, C. B., R. A. Padgett and P. A. Sharp (1998). "Evolutionary fates and origins of U12-type introns." Molecular Cell **2**(6): 773-785.

Calvo, O. and J. L. Manley (2001). "Evolutionarily conserved interaction between CstF-64 and PC4 links transcription, polyadenylation, and termination." Molecular Cell **7**(5): 1013-1023.

Chanfreau, C., M. Buckle and A. Jacquier (2000). "Recognition of a conserved class of RNA tetraloops by *Saccharomyces cerevisiae* RNase III." PNAS **97**(7): 3142-3147.

Chanfreau, G., S. AbouElela, M. Ares and C. Guthrie (1997). "Alternative 3'-end processing of U5 snRNA by RNase III." Genes & Development **11**(20): 2741-2751.

Chanfreau, G., P. Legrain and A. Jacquier (1998). "Yeast RNase III as a key processing enzyme in small nucleolar RNAs metabolism." Journal of Molecular Biology **284**(4): 975-988.

Chedin, S., M. Riva, P. Schultz, A. Sentenac and C. Carles (1998). "The RNA cleavage activity of RNA polymerase III is mediated by an essential TFIIS-like subunit and is important for transcription termination." Genes & Development **12**(24): 3857-3871.

Cho, E. J., T. Takagi, C. R. Moore and S. Buratowski (1997). "mRNA capping enzyme is recruited to the transcription complex by phosphorylation of the RNA polymerase II carboxy-terminal domain." Genes & Development **11**(24): 3319-3326.

Cogoni, C. and G. Macino (2000). "Post-transcriptional gene silencing across kingdoms." Current Opinion in Genetics & Development **10**(6): 638-643.

Colgan, D. F. and J. L. Manley (1997). "Mechanism and regulation of mRNA polyadenylation." Genes & Development **11**(21): 2755-2766.

Conaway, J. W. and R. C. Conaway (2000). "Structural biology - Light at the end of the channel." Science **288**(5466): 632-633.

Conaway, J. W., A. Shilatifard, A. Dvir and R. C. Conaway (2000). "Control of elongation by RNA polymerase II." Trends in Biochemical Sciences **25**(8): 375-380.

Conaway, R. C. and J. W. Conaway (1993). "General Initiation Factors for RNA Polymerase II." Annual Review of Biochemistry **62**: 161-190.

Conconi, A., R. M. Widmer, T. Koller and J. M. Sogo (1989). "2 Different Chromatin Structures Coexist in rRNA Genes Throughout the Cell Cycle." Cell **57**(5): 753-761.

Connelly, S. and J. L. Manley (1988). "A Functional mRNA Polyadenylation Signal Is Required for Transcription Termination by RNA Polymerase II." Genes & Development **2**(4): 440-452.

Conradwebb, H. and R. A. Butow (1995). "A Polymerase Switch in the Synthesis of rRNA in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **15**(5): 2420-2428.

Cook, P. R. (1994). "RNA Polymerase - Structural Determinant of the Chromatin Loop and the Chromosome." Bioessays **16**(6): 425-430.

Cramer, P., D. A. Bushnell and R. D. Kornberg (2001). "Structural basis of transcription: RNA polymerase II at 2.8 angstrom resolution." Science **292**(5523): 1863-1876.

Cramer, P., A. Srebrow, S. Kadener, S. Werbajh, M. de la Mata, G. Melen, G. Nogues and A. R. Kornblihtt (2001). "Coordination between transcription and pre-mRNA processing." FEBS Letters **498**(2-3): 179-182.

Dahmus, M. E. (1996). "Reversible phosphorylation of the C-terminal domain of RNA polymerase II." Journal of Biological Chemistry **271**(32): 19009-19012.

Dammann, R., R. Lucchini, T. Koller and J. M. Sogo (1993). "Chromatin Structures and Transcription of rDNA in Yeast *Saccharomyces cerevisiae*." Nucleic Acids Research **21**(10): 2331-2338.

Dammann, R., R. Lucchini, T. Koller and J. M. Sogo (1995). "Transcription in the Yeast rRNA Gene Locus: Distribution of the Active Gene Copies and Chromatin Structure of Their Flanking Regulatory Sequences." Molecular and Cellular Biology **15**(10): 5294-5303.

Dichtl, B. and W. Keller (2001). "Recognition of polyadenylation sites in yeast pre-mRNAs by cleavage and polyadenylation factor." EMBO Journal **20**(12): 3197-3209.

Dujon, B. (1996). "The yeast genome project: What did we learn?" Trends in Genetics **12**(7): 263-270.

Dunaway, M. and E. A. Ostrander (1993). "Local Domains of Supercoiling Activate a Eukaryotic Promoter *In vivo*." Nature **361**(6414): 746-748.

Dvir, A., J. W. Conaway and R. C. Conaway (2001). "Mechanism of transcription initiation and promoter escape by RNA polymerase II." Current Opinion in Genetics & Development **11**(2): 209-214.

Dye, M. J. and N. J. Proudfoot (2001). "Multiple transcript cleavage precedes polymerase release in termination by RNA polymerase II." Cell **105**(5): 669-681.

Elela, S. A. and M. Ares (1998). "Depletion of yeast RNase III blocks correct U2 3' end formation and results in polyadenylated but functional U2 snRNA." Embo Journal **17**(13): 3738-3746.

Elela, S. A., H. Igel and M. Ares (1996). "RNase III cleaves eukaryotic pre-rRNA at a U3 snoRNP-dependent site." Cell **85**(1): 115-124.

Elion, E. A. and J. R. Warner (1984). "The Major Promoter Element of rRNA Transcription in Yeast Lies 2 Kb Upstream." Cell **39**(3): 663-673.

Elion, E. A. and J. R. Warner (1986). "An RNA Polymerase I Enhancer in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **6**(6): 2089-2097.

Elion, E. A. and J. R. Warner (1986). "An RNA Polymerase I Enhancer in Yeast." Journal of Cellular Biochemistry: 183-183.

Evers, R. and I. Grummt (1995). "Molecular Co-evolution of Mammalian Ribosomal Gene Terminator Sequences and the Transcription Termination Factor Ttf-I." Proceedings of the National Academy of Sciences of the United States of America **92**(13): 5827-5831.

Fath, S., P. Milkereit, G. Peyroche, M. Riva, C. Carles and H. Tschochner (2001). "Differential roles of phosphorylation in the formation of transcriptional active RNA polymerase I." PNAS **98**(25): 14334-14339.

Feldmann, H., M. Aigle, G. Aljinovic, B. Andre, M. C. Baclet, C. Barthe, A. Baur, A. M. Becam, N. Biteau, E. Boles, T. Brandt, M. Brendel, M. Bruckner, F. Bussereau, C. Christiansen, R. Contreras, M. Crouzet, C. Ciepluch, N. Demolis, T. Delaveau, F. Doignon, H. Domdey, S. Dusterhus, E. Dubois, B. Dujon, M. Elbakkoury, K. D. Entian, M. Feuermann, W. Fiers, G. M. Fobo, C. Fritz, H. Gassenhuber, N. Glansdorff, A. Goffeau, L. A. Grivell, M. Dehaan, C. Hein, C. J. Herbert, C. P. Hollenberg, K. Holmstrom, C. Jacq, M. Jacquet, J. C. Jauniaux, J. L. Jonniaux, T. Kallesoe, P. Kiesau, L. Kirchrath, P. Kotter, S. Koroll, S. Liebl, M. Logghe, A. J. E. Lohan, E. J. Louis, Z. Y. Li, M. J. Maat, L. Mallet, G. Mannhaupt, F. Messenguy, T. Miosga, F. Molemans, S. Muller, F. Nasr, B. Obermaier, J. Perea, A. Pierard, E. Piravandi, F. M. Pohl, T. M. Pohl, S. Potier, M. Proft, B. Purnelle, M. R. Rad, M. Rieger, M. Rose, I. Schaaffgerstenschlager, B. Scherens, C. Schwarzlose, J. Skala, P. P. Slonimski, P. H. M. Smits, J. L. Souciet, H. Y. Steensma, R. Stucka, A. Urrestarazu, Q. J. M. Vanderaart, L. Vandyck, A. Vassarotti, I. Vetter, F. Vierendeels, S. Vissers, G. Wagner, P. Dewergifosse, K. H. Wolfe, M. Zagulski, F. K. Zimmermann, H. W. Mewes and K. Kleine (1994). "Complete DNA Sequence of Yeast Chromosome II." Embo Journal **13**(24): 5795-5809.

Flores, A., J. F. Briand, O. Gadal, J. C. Andrau, L. Rubbi, V. Van Mullem, C. Boschiero, M. Goussot, C. Marck, C. Carles, P. Thuriaux, A. Sentenac and M. Werner (1999). "A protein-protein interaction map of yeast RNA polymerase III." PNAS **96**(14): 7815-7820.

Furtergraves, E. M. and B. D. Hall (1990). "DNA Sequence Elements Required for Transcription Initiation of the *Schizosaccharomyces pombe* Adh Gene in *Saccharomyces cerevisiae*." Molecular & General Genetics **223**(3): 407-416.

Gadal, O., S. MariotteLabarre, S. Chedin, E. Quemeneur, C. Carles, A. Sentenac and P. Thuriaux (1997). "A34.5, a nonessential component of yeast RNA polymerase I, cooperates with subunit A14 and DNA topoisomerase I to produce a functional rRNA synthesis machine." Molecular and Cellular Biology **17**(4): 1787-1795.

Giaever, G. N., L. Snyder and J. C. Wang (1988). "DNA Supercoiling *In vivo*." Biophysical Chemistry **29**(1-2): 7-15.

Giaever, G. N. and J. C. Wang (1988). "Supercoiling of Intracellular DNA Can Occur in Eukaryotic Cells." Cell **55**(5): 849-856.

Gietz, R. D. and R. H. Schiestl (1995). "Transforming yeast with DNA." Methods in Molecular and Cellular Biology **5**(5): 255-269.

Gnatt, A. L., P. Cramer, J. H. Fu, D. A. Bushnell and R. D. Kornberg (2001). "Structural basis of transcription: An RNA polymerase II elongation complex at 3.3 angstrom resolution." Science **292**(5523): 1876-1882.

Goffeau, A., B. G. Barrell, H. Bussey, R. W. Davis, B. Dujon, H. Feldmann, F. Galibert, J. D. Hoheisel, C. Jacq, M. Johnston, E. J. Louis, H. W. Mewes, Y. Murakami, P. Philippsen, H. Tettelin and S. G. Oliver (1996). "Life with 6000 genes." Science **274**(5287): 563-467.

Greger, I. H., A. Aranda and N. Proudfoot (2000). "Balancing transcriptional interference and initiation on the *GAL7* promoter of *Saccharomyces cerevisiae*." PNAS **97**(15): 8415-8420.

Greger, I. H. and N. J. Proudfoot (1998). "Poly(A) signals control both transcriptional termination and initiation between the tandem *GAL10* and *GAL7* genes of Saccharomyces cerevisiae." EMBO Journal **17**(16): 4771-4779.

Grummt, I., H. Sorbaz, A. Hofmann and E. Roth (1985). "Spacer Sequences Downstream of the 28S-RNA Coding Region Are Part of the Mouse rDNA Transcription Unit." Nucleic Acids Research **13**(7): 2293-2304.

Guo, Z. J., P. Russo, D. F. Yun, J. S. Butler and F. Sherman (1995). "Redundant 3' End-Forming Signals for the Yeast Cyc1 mRNA." PNAS **92**(10): 4211-4214.

Guo, Z. J. and F. Sherman (1995). "3' End Forming Signals of Yeast Messenger RNA." Molecular and Cellular Biology **15**(11): 5983-5990.

Guo, Z. J. and F. Sherman (1996). "3'-end forming signals of yeast mRNA." Trends in Biochemical Sciences **21**(12): 477-481.

Guo, Z. J. and F. Sherman (1996). "Signals sufficient for 3'end formation of yeast mRNA." Molecular and Cellular Biology **16**(6): 2772-2776.

Hammond, S. M., A. A. Caudy and G. J. Hannon (2001). "Post-transcriptional gene silencing by double-stranded RNA." Nature Reviews Genetics **2**(2): 110-119.

Hampsey, M. (1998). "Molecular genetics of the RNA polymerase II general transcriptional machinery." Microbiology and Molecular Biology Reviews **62**(2): 465-503.

Hampsey, M. and D. Reinberg (1999). "RNA polymerase II as a control panel for multiple co-activator complexes." Current Opinion in Genetics & Development **9**(2): 132-139.

Hansen, K., C. E. Birse and N. J. Proudfoot (1998). "Nascent transcription from the nmt1 and nmt2 genes of *Schizosaccharomyces pombe* overlaps neighbouring genes." EMBO Journal **17**(11): 3066-3077.

Henderson, S. L., K. Ryan and B. Sollnerwebb (1989). "The Promoter-Proximal rDNA Terminator Augments Initiation by Preventing Disruption of the Stable Transcription Complex Caused by Polymerase Read-In." Genes & Development **3**(2): 212-223.

Henkin, T. M. (2000). "Transcription termination control in bacteria." Current Opinion in Microbiology **3**(2): 149-153.

Hill, A. and K. Bloom (1987). "Genetic Manipulation of Centromere Function." Molecular and Cellular Biology **7**(7): 2397-2405.

Irniger, S. and G. H. Braus (1994). "Saturation Mutagenesis of a Polyadenylylation Signal Reveals a Hexanucleotide Element Essential for mRNA 3' End Formation in *Saccharomyces cerevisiae*." PNAS **91**(1): 257-261.

Ishihama, A., M. Kimura and H. Mitsuzawa (1998). "Subunits of yeast RNA polymerases: structure and function." Current Opinion in Microbiology **1**(2): 190-196.

Jansa, P. and I. Grummt (1999). "Mechanism of transcription termination: PTRF interacts with the largest subunit of RNA polymerase I and dissociates paused transcription complexes from yeast and mouse." Molecular and General Genetics **262**(3): 508-514.

Jeong, S. W., W. H. Lang and R. H. Reeder (1995). "The Release Element of the Yeast Polymerase I Transcription Terminator Can Function Independently of Reb1p." Molecular and Cellular Biology **15**(11): 5929-5936.

Jeong, S. W., W. H. Lang and R. H. Reeder (1996). "The yeast transcription terminator for RNA polymerase I is designed to prevent polymerase slippage." Journal of Biological Chemistry **271**(27): 16104-16110.

Johnson, S. P. and J. R. Warner (1989). "Unusual Enhancer Function in Yeast rRNA Transcription." Molecular and Cellular Biology **9**(11): 4986-4993.

Johnston, M. (1987). "A Model Fungal Gene Regulatory Mechanism - the GAL Genes of *Saccharomyces cerevisiae*." Microbiological Reviews **51**(4): 458-476.

Johnston, M. and M. Carlson (1992). Regulation of carbon and phosphate utilization. The Molecular Biology of the yeast *Saccharomyces cerevisiae*. Cold Spring Harbor, New York, Cold Spring Harbor Laboratory Press. **2:** 193-281.

Kaiser, M. W. and D. A. Brow (1995). "Lethal Mutations in a Yeast U6 RNA Gene-B Block Promoter Element Identify Essential Contacts With Transcription Factor-IIIc." Journal of Biological Chemistry **270**(19): 11398-11405.

Keener, J., C. A. Josaitis, J. A. Dodd and M. Nomura (1998). "Reconstitution of yeast RNA polymerase I transcription in vitro from purified components - TATA binding protein is not required for basal transcription." Journal of Biological Chemistry **273**(50): 33795-33802.

Kempersveenstra, A. E., J. Oliemans, H. Offenberg, A. F. Dekker, P. W. Piper, R. J. Planta and J. Klootwijk (1986). "3'-End Formation of Transcripts From the Yeast rRNA Operon." EMBO Journal **5**(10): 2703-2710.

Kenzior, A. and W. Folk (2001). "RNA polymerase I holoenzymes." Trends in Plant Science **6**(3): 87-88.

Keys, D. A., B. S. Lee, J. A. Dodd, T. T. Nguyen, L. Vu, E. Fantino, L. M. Burson, Y. Nogi and M. Nomura (1996). "Multiprotein transcription factor UAF interacts with the upstream element of the yeast RNA polymerase I promoter and forms a stable preinitiation complex." Genes & Development **10**(7): 887-903.

Keys, D. A., L. A. Vu, L. S. Steffan, J. A. Dodd, R. T. Yamamoto, Y. Nogi and M. Nomura (1994). "*RRN6* and *RRN7* Encode Subunits of a Multiprotein Complex Essential for the Initiation of rDNA Transcription by RNA Polymerase I in *Saccharomyces cerevisiae*." Genes & Development **8**(19): 2349-2362.

Khochbin, S., A. Verdel, C. Lemercier and D. Seigneurin-Berny (2001). "Functional significance of histone deacetylase diversity." Current Opinion in Genetics & Development **11**(2): 162-166.

Kim, R. A. and J. C. Wang (1989). "A Subthreshold Level of DNA Topoisomerases Leads to the Excision of Yeast rDNA As Extrachromosomal Rings." Cell **57**(6): 975-985.

Kim, T. K., T. H. Kim and T. Maniatis (1998). "Efficient recruitment of TFIIB and CBP-RNA polymerase II holoenzyme by an interferon-beta enhanceosome in vitro." PNAS **95**(21): 12191-12196.

Klinger, C., J. Huet, D. Y. Song, G. Petersen, M. Riva, E. K. F. Bautz, A. Sentenac, P. Oudet and P. Schultz (1996). "Localization of yeast RNA polymerase I core subunits by immunoelectron microscopy." EMBO Journal **15**(17): 4643-4653.

Koleske, A. J. and R. A. Young (1995). "The RNA Polymerase II Holoenzyme and Its Implications for Gene Regulation." Trends in Biochemical Sciences **20**(3): 113-116.

Kressler, D., P. Linder and J. de la Cruz (1999). "Protein trans-acting factors involved in ribosome biogenesis in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **19**(12): 7897-7912.

Kruglyak, S. and H. X. Tang (2000). "Regulation of adjacent yeast genes." Trends in Genetics **16**(3): 109-111.

Kufel, J., B. Dichtl and D. Tollervey (1999). "Yeast Rnt1p is required for cleavage of the pre-rRNA in the 3 ' ETS but not the 5 ' ETS." RNA **5**(7): 909-917.

Kulkens, T., C. Vandersande, A. F. Dekker, H. Vanheerikhuizen and R. J. Planta (1992). "A System to Study Transcription by Yeast RNA Polymerase I within the Chromosomal Context - Functional Analysis of the Ribosomal DNA Enhancer and the Rbp1/Reb1 Binding Sites." Embo Journal **11**(12): 4665-4674.

Kuo, M. H., E. vom Baur, K. Struhl and C. D. Allis (2000). "Gcn4 activator targets Gcn5 histone acetyltransferase to specific promoters independently of transcription." Molecular Cell **6**(6): 1309-1320.

Lalo, D., J. S. Steffan, J. A. Dodd and M. Nomura (1996). "*RRN11* encodes the third subunit of the complex containing Rrn6p and Rrn7p that is essential for the initiation of rDNA transcription by yeast RNA polymerase I." Journal of Biological Chemistry **271**(35): 21062-21067.

Lamontagne, B., A. Tremblay and S. Abou Elela (2000). "The N-terminal domain that distinguishes yeast from bacterial RNase III contains a dimerization signal required for efficient double-stranded RNA cleavage." Molecular and Cellular Biology **20**(4): 1104-1115.

Landick, R. (1997). "RNA polymerase slides home: Pause and termination site recognition." Cell **88**(6): 741-744.

Lang, W. H., B. E. Morrow, Q. D. Ju, J. R. Warner and R. H. Reeder (1994). "A Model for Transcription Termination by RNA Polymerase I." Cell **79**(3): 527-534.

Lang, W. H. and R. H. Reeder (1993). "The Reb1 Site Is an Essential Component of a Terminator for RNA Polymerase I in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **13**(1): 649-658.

Lang, W. H. and R. H. Reeder (1995). "Transcription Termination of RNA Polymerase I Due to a T-Rich Element Interacting with Reb1p." PNAS **92**(21): 9781-9785.

Langst, G., P. B. Becker and I. Grummt (1998). "TTF-I determines the chromatin architecture of the active rDNA promoter." EMBO Journal **17**(11): 3135-3145.

Lanzendorfer, M., A. Smid, C. Klinger, P. Schultz, A. Sentenac, C. Carles and M. Riva (1997). "A shared subunit belongs to the eukaryotic core RNA polymerase." Genes & Development **11**(8): 1037-1047.

Lebars, I., B. Lamontagne, S. Yoshizawa, S. Abou Elela and D. Fourmy (2001). "Solution structure of conserved AGNN tetraloops: insights into Rnt1p RNA processing." EMBO Journal **20**(24): 7250-7258.

Lee, T. I. and R. A. Young (2000). "Transcription of eukaryotic protein-coding genes." Annual Review of Genetics **34**: 77-137.

Leloir, L. F. (1951). "The enzymatic transformation of uridine diphosphate glucose into a galactose derivative." Archives of Biochem. **33**: 186-190.

Lin, C. W., B. Moorefield, J. Payne, P. Aprikian, K. Mitomo and R. H. Reeder (1996). "A novel 66-kilodalton protein complexes with Rrn6, Rrn7, and TATA-binding protein to promote polymerase I transcription initiation in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **16**(11): 6436-6443.

Liu, L. F. and J. C. Wang (1987). "Supercoiling of the DNA Template During Transcription." PNAS **84**(20): 7024-7027.

Liu, Y., J. A. Ranish, R. Aebersold and S. Hahn (2001). "Yeast nuclear extract contains two major forms of RNA polymerase II mediator complexes." Journal of Biological Chemistry **276**(10): 7169-7175.

Logan, J., E. Falckpedersen, J. E. Darnell and T. Shenk (1987). "A Poly(a) Addition Site and a Downstream Termination Region Are Required for Efficient Cessation of Transcription by RNA Polymerase II in the Mouse Beta-Maj-Globin Gene." PNAS **84**(23): 8306-8310.

Lohr, D., P. Venkov and J. Zlatanova (1995). "Transcriptional Regulation in the Yeast GAL Gene Family - a Complex Genetic Network." Faseb Journal **9**(9): 777-787.

Lucchini, R. and J. M. Sogo (1998). The Dynamic Structure of Ribosome RNA Gene Chromatin. Transcription of Ribosomal RNA Genes by Eukaryotic RNA Polymerase I. M. R. Paule. Berlin, Springer-Verlag: 255-276.

Madhani, H. D. and C. Guthrie (1992). "A Novel Base-Pairing Interaction between U2 and U6 Snrnas Suggests a Mechanism for the Catalytic Activation of the Spliceosome." Cell **71**(5): 803-817.

Malavasic, M. J. and R. T. Elder (1990). "Complementary Transcripts from 2 Genes Necessary for Normal Meiosis in the Yeast *Saccharomyces cerevisiae*." Molecular and Cellular Biology **10**(6): 2809-2819.

Mason, S. W., E. E. Sander, R. Evers and I. Grummt (1998). Termination of Mammalian RNA Polymerase I Transcription. Transcription of Ribosomal RNA Genes by Eukaryotic RNA Polymerase I. M. R. Paule. Berlin, Springer-Verlag: 179-194.

Mason, S. W., E. E. Sander and I. Grummt (1997). "Identification of a transcript release activity acting on ternary transcription complexes containing murine RNA polymerase I." EMBO Journal **16**(1): 163-172.

McGrath, J. P. and A. Varshavsky (1989). "The Yeast Ste6 Gene Encodes a Homolog of the Mammalian Multidrug Resistance P-Glycoprotein." Nature **340**(6232): 400-404.

McStay, B. and R. H. Reeder (1990). "A DNA-Binding Protein Is Required for Termination of Transcription by RNA Polymerase I in *Xenopus laevis*." Molecular and Cellular Biology **10**(6): 2793-2800.

McStay, B., G. J. Sullivan and C. Cairns (1997). "The Xenopus RNA polymerase I transcription factor, UBF, has a role in transcriptional enhancement distinct from that at the promoter." EMBO Journal **16**(2): 396-405.

Milkereit, P. and H. Tschochner (1998). "Specialized form of RNA polymerase I, essential for initiation and growth-dependent regulation of rRNA synthesis, is disrupted during transcription." EMBO Journal **17**(13): 3692-3703.

Moore, M. J. and P. A. Sharp (1993). "Evidence for 2 Active Sites in the Spliceosome Provided by Stereochemistry of pre-mRNA Splicing." Nature **365**(6444): 364-368.

Morales, V., C. Giamarchi, C. Chailleux, F. Moro, V. Marsaud, S. Le Ricousse and H. Richard-Foy (2001). "Chromatin structure and dynamics: Functional implications." Biochimie **83**(11-12): 1029-1039.

Morrow, B. E., S. P. Johnson and J. R. Warner (1993). "The rRNA Enhancer Regulates rRNA Transcription in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **13**(2): 1283-1289.

Muscarella, D. E., V. M. Vogt and S. E. Bloom (1987). "Characterization of rRNA Synthesis in a Gene Dosage Mutant - the Relationship of Topoisomerase I and Chromatin Structure to Transcriptional Activity." Journal of Cell Biology **105**(4): 1501-1513.

Myer, V. E. and R. A. Young (1998). "RNA polymerase II holoenzymes and subcomplexes." Journal of Biological Chemistry **273**(43): 27757-27760.

Nagel, R. and M. Ares (2000). "Substrate recognition by a eukaryotic RNase III: The double-stranded RNA-binding domain of Rnt1p selectively binds RNA containing a 5 '-AGNN-3 ' tetraloop." RNA **6**(8): 1142-1156.

Narlikar, G. J., H. Y. Fan and R. E. Kingston (2002). "Cooperation between complexes that regulate chromatin structure and transcription." Cell **108**(4): 475-487.

Nogi, Y., R. Yano, J. Dodd, C. Carles and M. Nomura (1993). "Gene *RRN4* in *Saccharomyces cerevisiae* Encodes the A12.2 Subunit of RNA Polymerase I and Is Essential Only at High Temperatures." Molecular and Cellular Biology **13**(1): 114-122.

Nogi, Y., R. Yano and M. Nomura (1991). "Synthesis of Large rRNAs By RNAPolymerase II in Mutants of *Saccharomyces cerevisiae* Defective in RNA Polymerase I." PNAS **88**(9): 3962-3966.

Nomura, M. (1998). Transcription Factors Used by *Saccharomyces cerevisiae* RNA Polymerase I and the Mechanism of Initiation. Transcription of Ribosomal RNA Genes by Eukaryotic RNA Polymerase I. M. R. Paule. Berlin, Springer-Verlag: 155-172.

Nudler, E. (1999). "Transcription elongation: Structural basis and mechanisms." Journal of Molecular Biology **288**(1): 1-12.

Nudler, E., I. Gusarov, E. Avetissova, M. Kozlov and A. Goldfarb (1998). "Spatial organization of transcription elongation complex in *Escherichia coli.*" Science **281**(5375): 424-428.

Nudler, E., A. Mustaev, E. Lukhtanov and A. Goldfarb (1997). "The RNA-DNA hybrid maintains the register of transcription by preventing backtracking of RNA polymerase." Cell **89**(1): 33-41.

Oakes, M., I. Siddiqi, L. Vu, J. Aris and M. Nomura (1999). "Transcription factor UAF, expansion and contraction of ribosomal DNA (rDNA) repeats, and RNA polymerase switch in transcription of yeast rDNA." Molecular and Cellular Biology **19**(12): 8559-8569.

Orphanides, G., T. Lagrange and D. Reinberg (1996). "The general transcription factors of RNA polymerase II." Genes & Development **10**(21): 2657-2683.

Orphanides, G. and D. Reinberg (2000). "RNA polymerase II elongation through chromatin." Nature **407**(6803): 471-475.

Osborne, B. I. and L. Guarente (1989). "Mutational Analysis of a Yeast Transcriptional Terminator." PNAS **86**(11): 4097-4101.

Osheim, Y. N., E. B. Mougey, J. Windle, M. Anderson, M. Oreilly, O. L. Miller, A. Beyer and B. SollnerWebb (1996). "Metazoan rDNA enhancer acts by making more genes transcriptionally active." Journal of Cell Biology **133**(5): 943-954.

Otero, G., J. Fellows, Y. Li, T. de Bizemont, A. M. G. Dirac, C. M. Gustafsson, H. Erdjument-Bromage, P. Tempst and J. Q. Svejstrup (1999). "Elongator, a multisubunit component of a novel RNA polymerase II holoenzyme for transcriptional elongation." Molecular Cell **3**(1): 109-118.

Pape, L. K., J. J. Windle and B. Sollnerwebb (1990). "Half Helical Turn Spacing Changes Convert a Frog into a Mouse rDNA Promoter - a Distant Upstream Domain Determines the Helix Face of the Initiation Site." Genes & Development **4**(1): 52-62.

Park, J. M., J. Werner, J. M. Kim, J. T. Lis and Y. J. Kim (2001). "Mediator, not holoenzyme, is directly recruited to the heat shock promoter by HSF upon heat shock." Molecular Cell **8**(1): 9-19.

Paule, M. R. and R. J. White (2000). "Transcription by RNA polymerases I and III." Nucleic Acids Research **28**(6): 1283-1298.

Pavco, P. A. and D. A. Steege (1991). "Characterization of Elongating T7 and Sp6 RNA Polymerases and Their Response to a Roadblock Generated by a Site Specific DNA Binding Protein." Nucleic Acids Research **19**(17): 4639-4646.

Peterson, J. A. and A. M. Myers (1993). "Functional Analysis of mRNA 3' End Formation Signals in the Convergent and Overlapping Transcription Units of the *Saccharomyces cerevisiae* Genes *RHO1* and *MRP2*." Nucleic Acids Research **21**(23): 5500-5508.

Peyroche, G., P. Milkereit, N. Bischler, H. Tschochner, P. Schultz, A. Sentenac, C. Carles and M. Riva (2000). "The recruitment of RNA polymerase I on rDNA is mediated by the interaction of the A43 subunit with Rrn3." EMBO Journal **19**(20): 5473-5482.

Pokholok, D. K., N. M. Hannett and R. A. Young (2002). "Exchange of RNA polymerase II initiation and elongation factors during gene expression *in vivo*." Molecular Cell **9**(4): 799-809.

Preker, P. J., M. Ohnacker, L. MinvielleSebastia and W. Keller (1997). "A multisubunit 3' end processing factor from yeast containing poly(A) polymerase and homologues of the subunits of mammalian cleavage and polyadenylation specificity factor." EMBO Journal **16**(15): 4727-4737.

Proudfoot, N. J. (1989). "How RNA Polymerase II Terminates Transcription in Higher Eukaryotes." Trends in Biochemical Sciences **14**(3): 105-110.

Proudfoot, N. J., A. Furger and M. J. Dye (2002). "Integrating mRNA processing with transcription." Cell **108**(4): 501-512.

Puig, S., J. E. Perez-Ortin and E. Matallana (1999). "Transcriptional and structural study of a region of two convergent overlapping yeast genes." Current Microbiology **39**(6): 369-373.

Qu, L. H., A. Henras, Y. J. Lu, H. Zhou, W. X. Zhou, Y. Q. Zhu, J. Zho, Y. Henry, M. Caizergues-Ferrer and J. P. Bachellerie (1999). "Seven novel methylation guide small nucleolar RNAs are processed from a common polycistronic transcript by Rat1p and RNase III in yeast." Molecular and Cellular Biology **19**(2): 1144-1158.

Reeder, R. H. (1999). Regulation of RNA polymerase I transcription in yeast and vertebrates. Progress in Nucleic Acid Research and Molecular Biology, Vol 62. **62:** 293-327.

Reeder, R. H., P. Guevara and J. G. Roan (1999). "Saccharomyces cerevisiae RNA polymerase I terminates transcription at the Reb1 terminator *in vivo*." Molecular and Cellular Biology **19**(11): 7369-7376.

Reeder, R. H. and W. Lang (1994). "The Mechanism of Transcription Termination by RNA Polymerase I." Molecular Microbiology **12**(1): 11-15.

Reeder, R. H. and W. H. Lang (1997). "Terminating transcription in eukaryotes: lessons learned from RNA polymerase I." Trends in Biochemical Sciences **22**(12): 473-477.

Reeder, R. H. and W. H. Lang (1998). Stopping RNA Polymerase I. Transcription of Ribosomal RNA Genes by Eukaryotic RNA Polymerase I. M. R. Paule. Berlin, Springer-Verlag**:** 173-178.

Reid, J. L., V. R. Iyer, P. O. Brown and K. Struhl (2000). "Coordinate regulation of yeast ribosomal protein genes is associated with targeted recruitment of Esa1 histone acetylase." Molecular Cell **6**(6): 1297-1307.

Reines, D., R. C. Conaway and J. W. Conaway (1999). "Mechanism and regulation of transcriptional elongation by RNA polymerase II." Current Opinion in Cell Biology **11**(3): 342-346.

Rocco, V., B. Demassy and A. Nicolas (1992). "The *Saccharomyces cerevisiae ARG4* Initiator of Meiotic Gene Conversion and Its Associated Double-Strand DNA Breaks Can Be Inhibited by Transcriptional Interference." PNAS **89**(24): 12068-12072.

Roeder, R. G. (1996). "The role of general initiation factors in transcription by RNA polymerase II." Trends in Biochemical Sciences **21**(9): 327-335.

Sambrook, J., E. F. Fritsch and T. Maniatis (1989). Molecular Cloning: A Laboratory Manual. Cold Spring Harbour, NY, Cold Spring Harbor Laboratory Press.

Sander, E. E. and I. Grummt (1997). "Oligomerization of the transcription termination factor TTF-I: Implications for the structural organization of ribosomal transcription units." Nucleic Acids Research **25**(6): 1142-1147.

Schultz, M. C., Y. C. Soo and R. H. Reeder (1993). "*In vitro* Definition of the Yeast RNA Polymerase I Enhancer." Molecular and Cellular Biology **13**(5): 2644-2654.

Schultz, P., H. Celia, M. Riva, A. Sentenac and P. Oudet (1993). "Three Dimensional Model of Yeast RNA Polymerase I Determined by Electron Microscopy of Two Dimensional Crystals." EMBO Journal **12**(7): 2601-2607.

Seipelt, R. L., B. H. Zheng, A. Asuru and B. C. Rymond (1999). "U1 snRNA is cleaved by RNase III and processed through an Sm site-dependent pathway." Nucleic Acids Research **27**(2): 587-595.

Shatkin, A. J. and J. L. Manley (2000). "The ends of the affair: Capping and polyadenylation." Nature Structural Biology **7**(10): 838-842.

Shilatifard, A. (1998). "Factors regulating the transcriptional elongation activity of RNA polymerase II." Faseb Journal **12**(14): 1437-1446.

Shilatifard, A. (1998). "The RNA polymerase II general elongation complex." Biological Chemistry **379**(1): 27-31.

Siddiqi, I. N., J. A. Dodd, L. Vu, K. Eliason, M. L. Oakes, J. Keener, R. Moore, M. K. Young and M. Nomura (2001). "Transcription of chromosomal rRNA genes by both RNA polymerase I and II in yeast uaf30 mutants lacking the 30 kDa subunit of transcription factor UAF." EMBO Journal **20**(16): 4512-4521.

Snyder, M., R. J. Sapolsky and R. W. Davis (1988). "Transcription Interferes with Elements Important for Chromosome Maintenance in *Saccharomyces cerevisiae*." Molecular and Cellular Biology **8**(5): 2184-2194.

Springer, C., O. Valerius, A. Strittmatter and G. H. Braus (1997). "The adjacent yeast genes *ARO4* and *HIS7* carry no intergenic region." Journal of Biological Chemistry **272**(42): 26318-26324.

Steffan, J. S., D. A. Keys, J. A. Dodd and M. Nomura (1996). "The role of TBP in rDNA transcription by RNA Polymerase I in *Saccharomyces cerevisiae*: TBP is required for upstream activation factor-dependent recruitment of core factor." Genes & Development **10**(20): 2551-2563.

Steffan, J. S., D. A. Keys, L. Vu and M. Nomura (1998). "Interaction of TATA-binding protein with upstream activation factor is required for activated transcription of ribosomal DNA by RNA polymerase I in Saccharomyces cerevisiae *in vivo*." Molecular and Cellular Biology **18**(7): 3752-3761.

Sun, J. S. and J. L. Manley (1995). "A Novel U2-U6 SnRNA Structure Is Necessary for Mammalian mRNA Splicing." Genes & Development **9**(7): 843-854.

Svejstrup, J. Q. (2002). "Chromatin elongation factors." Current Opinion in Genetics & Development **12**(2): 156-161.

Tanaka, S., D. Halter, M. Livingstonezatchej, B. Reszel and F. Thoma (1994). "Transcription through the Yeast Origin of Replication Ars1 Ends at the AbfI Binding-Site and Affects Extrachromosomal Maintenance of Minichromosomes." Nucleic Acids Research **22**(19): 3904-3910.

Tollervey, D. and T. Kiss (1997). "Function and synthesis of small nucleolar RNAs." Current Opinion in Cell Biology **9**(3): 337-342.

Tsao, Y. P., H. Y. Wu and L. F. Liu (1989). "Transcription Driven Supercoiling of DNA Direct Biochemical Evidence From *In vitro* Studies." Cell **56**(1): 111-118.

Tschochne, H. and P. Milkereit (1997). "RNA Polymerase I from *S. cerevisiae* depends on an additional factor to release terminated transcripts from the template." FEBS Letters **410**(2-3): 461-466.

Uptain, S. M., C. M. Kane and M. J. Chamberlin (1997). "Basic mechanisms of transcript elongation and its regulation." Annual Review of Biochemistry **66**: 117-172.

Van Mullem, V., E. Landrieux, J. Vandenhaute and P. Thuriaux (2002). "Rpa12p, a conserved RNA Polymerase I subunit with two functional domains." Molecular Microbiology **43**(5): 1105-1113.

Vandersande, C., T. Kulkens, A. B. Kramer, I. J. Dewijs, H. Vanheerikhuizen, J. Klootwijk and R. J. Planta (1989). "Termination of Transcription by Yeast RNA Polymerase I." Nucleic Acids Research **17**(22): 9127-9146.

Venema, J. and D. Tollervey (1999). "Ribosome synthesis in *Saccharomyces cerevisiae*." Annual Review of Genetics **33**: 261-311.

Venter, J. C., M. D. Adams, E. W. Myers, P. W. Li, R. J. Mural, G. G. Sutton, H. O. Smith, M. Yandell, C. A. Evans, R. A. Holt, J. D. Gocayne, P. Amanatides, R. M. Ballew, D. H. Huson, J. R. Wortman, Q. Zhang, C. D. Kodira, X. Q. H. Zheng, L. Chen, M. Skupski, G. Subramanian, P. D. Thomas, J. H. Zhang, G. L. G. Miklos, C. Nelson, S. Broder, A. G. Clark, C. Nadeau, V. A. McKusick, N. Zinder, A. J. Levine, R. J. Roberts, M. Simon, C. Slayman, M. Hunkapiller, R. Bolanos, A. Delcher, I. Dew, D. Fasulo, M. Flanigan, L. Florea, A. Halpern, S. Hannenhalli, S. Kravitz, S. Levy, C. Mobarry, K. Reinert, K. Remington, J. Abu-Threideh, E. Beasley, K. Biddick, V. Bonazzi, R. Brandon, M. Cargill, I. Chandramouliswaran, R. Charlab, K. Chaturvedi, Z. M. Deng, V. Di Francesco, P. Dunn, K. Eilbeck, C. Evangelista, A. E. Gabrielian, W. Gan, W. M. Ge, F. C. Gong, Z. P. Gu, P. Guan, T. J. Heiman, M. E. Higgins, R. R. Ji, Z. X. Ke, K. A. Ketchum, Z. W. Lai, Y. D. Lei, Z. Y. Li, J. Y. Li, Y. Liang, X. Y. Lin, F. Lu, G. V. Merkulov, N. Milshina, H. M. Moore, A. K. Naik, V. A. Narayan, B. Neelam, D. Nusskern, D. B. Rusch, S. Salzberg, W. Shao, B. X. Shue, J. T. Sun, Z. Y. Wang, A. H. Wang, X. Wang, J. Wang, M. H. Wei, R. Wides, C. L. Xiao, C. H. Yan, A. Yao, J. Ye, M. Zhan, W. Q. Zhang, H. Y. Zhang, Q. Zhao, L. S. Zheng, F. Zhong, W. Y. Zhong, S. P. C. Zhu, S. Y. Zhao, D. Gilbert, S. Baumhueter, G. Spier, C. Carter, A. Cravchik, T. Woodage, F. Ali, H. J. An, A. Awe, D. Baldwin, H. Baden, M. Barnstead, I. Barrow, K. Beeson, D. Busam, A. Carver, A. Center, M. L. Cheng, L. Curry, S. Danaher, L. Davenport, R. Desilets, S. Dietz, K. Dodson, L. Doup, S. Ferriera, N. Garg, A. Gluecksmann, B. Hart, J. Haynes, C. Haynes, C. Heiner, S. Hladun, D. Hostin, J. Houck, T. Howland, C. Ibegwam, J. Johnson, F. Kalush, L. Kline, S. Koduru, A. Love, F. Mann, D. May, S. McCawley, T. McIntosh, I. McMullen, M. Moy, L. Moy, B. Murphy, K. Nelson, C. Pfannkoch, E. Pratts, V. Puri, H. Qureshi, M. Reardon, R. Rodriguez, Y. H. Rogers, D. Romblad, B. Ruhfel, R. Scott, C. Sitter, M. Smallwood, E. Stewart, R. Strong, E. Suh, R. Thomas, N. N. Tint, S. Tse, C. Vech, G. Wang, J. Wetter, S. Williams, M. Williams, S. Windsor, E. Winn-Deen, K. Wolfe, J. Zaveri, K. Zaveri, J. F. Abril, R. Guigo, M. J. Campbell, K. V. Sjolander, B. Karlak, A. Kejariwal, H. Y. Mi, B. Lazareva, T. Hatton, A. Narechania, K. Diemer, A. Muruganujan, N. Guo, S. Sato, V. Bafna, S. Istrail, R. Lippert, R. Schwartz, B. Walenz, S. Yooseph, D. Allen, A. Basu, J. Baxendale, L. Blick, M. Caminha, J. Carnes-Stine, P. Caulk, Y. H. Chiang, M. Coyne, C. Dahlke, A. D. Mays, M. Dombroski, M. Donnelly, D. Ely, S. Esparham, C. Fosler, H. Gire, S. Glanowski, K. Glasser, A. Glodek, M. Gorokhov, K. Graham, B. Gropman, M. Harris, J. Heil, S.

Henderson, J. Hoover, D. Jennings, C. Jordan, J. Jordan, J. Kasha, L. Kagan, C. Kraft, A. Levitsky, M. Lewis, X. J. Liu, J. Lopez, D. Ma, W. Majoros, J. McDaniel, S. Murphy, M. Newman, T. Nguyen, N. Nguyen, M. Nodell, S. Pan, J. Peck, M. Peterson, W. Rowe, R. Sanders, J. Scott, M. Simpson, T. Smith, A. Sprague, T. Stockwell, R. Turner, E. Venter, M. Wang, M. Y. Wen, D. Wu, M. Wu, A. Xia, A. Zandieh and X. H. Zhu (2001). "The sequence of the human genome." Science **291**(5507): 1304-1351.

Vogelauer, M., F. Cioci and G. Camilloni (1998). "DNA protein interactions at the *Saccharomyces cerevisiae* 35 S rRNA promoter and in its surrounding region." Journal of Molecular Biology **275**(2): 197-209.

Vogelauer, M., J. S. Wu, N. Suka and M. Grunstein (2000). "Global histone acetylation and deacetylation in yeast." Nature **408**(6811): 495-498.

Vu, L., I. Siddiqi, B. S. Lee, C. A. Josaitis and R. Nomura (1999). "RNA polymerase switch in transcription of yeast rDNA: Role of transcription factor UAF (upstream activation factor) in silencing rDNA transcription by RNA polymerase II." PNAS **96**(8): 4390-4395.

Wahle, E. and W. Keller (1996). "The biochemistry of polyadenylation." Trends in Biochemical Sciences **21**(7): 247-250.

Wahle, E. and U. Ruegsegger (1999). "3 '-End processing of pre-mRNA in eukaryotes." FEMS Microbiology Reviews **23**(3): 277-295.

Wai, H., K. Johzuka, L. Vu, K. Eliason, T. Kobayashi, T. Horiuchi and M. Nomura (2001). "Yeast RNA polymerase I enhancer is dispensable for transcription of the chromosomal rRNA gene and cell growth, and its apparent transcription enhancement from ectopic promoters requires Fob1 protein." Molecular and Cellular Biology **21**(16): 5541-5553.

Warner, J. R. (1999). "The economics of ribosome biosynthesis in yeast." Trends in Biochemical Sciences **24**(11): 437-440.

White, R. J. (2001). Gene Transcription: Mechanism and Control. Oxford, Blackwell Science Ltd.

Wind, M. and D. Reines (2000). "Transcription elongation factor SII." Bioessays **22**(4): 327-336.

Woychik, N. A. and M. Hampsey (2002). "The RNA polymerase II machinery: Structure illuminates function." Cell **108**(4): 453-463.

Wu, H. H., P. K. Yang, S. E. Butcher, S. Kang, G. Chanfreau and J. Feigon (2001). "A novel family of RNA tetraloop structure forms the recognition site for *Saccharomyces cerevisiae* RNase III." EMBO Journal **20**(24): 7240-7249.

Wu, H. Y., S. Shyy, J. C. Wang and L. F. Liu (1988). "Transcription Generates Positively and Negatively Supercoiled Domains in the Template." Cell **53**(3): 433-440.

Xie, W. Q. and L. I. Rothblum (1992). "Domains of the Rat rDNA Promoter Must Be Aligned Stereospecifically." Molecular and Cellular Biology **12**(3): 1266-1275.

Yager, T. D. and P. H. Vonhippel (1991). "A Thermodynamic Analysis of RNA Transcript Elongation and Termination in *Escherichia coli*." Biochemistry **30**(4): 1097-1118.

Yamamoto, R. T., Y. Nogi, J. A. Dodd and M. Nomura (1996). "*RRN3* gene of *Saccharomyces cerevisiae* encodes an essential RNA polymerase I transcription factor which interacts with the polymerase independently of DNA template." EMBO Journal **15**(15): 3964-3973.

Young, B. A., T. M. Gruber and C. A. Gross (2002). "Views of transcription initiation." Cell **109**(4): 417-420.

Yudkovsky, N., J. A. Ranish and S. Hahn (2000). "A transcription reinitiation intermediate that is stabilized by activator." Nature **408**(6809): 225-229.

Zawel, L., K. P. Kumar and D. Reinberg (1995). "Recycling of the General Transcription Factors During RNA Polymerase II Transcription." Genes & Development **9**(12): 1479-1490.

Zhao, J., L. Hyman and C. Moore (1999). "Formation of mRNA 3 ' ends in eukaryotes: Mechanism, regulation, and interrelationships with other steps in mRNA synthesis." Microbiology and Molecular Biology Reviews **63**(2): 405-445.