Title
A genetic analysis of mRNA splicing in Saccharomyces cerevisiae

Author
Parker, Roy,

Publication Date
1985

Peer reviewed|Thesis/dissertation

A Genetic Analysis of mRNA Splicing in *Saccharomyces cerevisiae*

by

Roy Parker

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

GENETICS

in the

GRADUATE DIVISION

of the

UNIVERSITY OF CALIFORNIA

San Francisco

Date **University Librarian**

Degree Conferred: . . JUN. 16 1985 . . . . . .

ACKNOWLEDGMENTS

In order to complete this thesis I have been aided in a multitude of ways by numerous people. In particular, I would like to thank my mother and father for their love and support over the years, and for providing me with an early enthusiasm for science. I also thank my advisor, Christine Guthrie, for her helpful advise, and almost undying patience throughout my graduate career.

My collaborators in this thesis work have included: John Rossi and Alessandra Cellini at the City of Hope Research Institute, who initiated the deletion analysis of the actin intron; and Joe Tamm, who concluded the analysis of the temperarture sensitive lethal suppressors.

I would also like to thank: Lucita Esperas for being a surrogate mother and sparring partner; Sherry Parker, my sister, for numerous dinners; Harold Svidlow, for teaching me more about volleyball then he ever knew; Leah Brindley, for forcing me to learn how to windsurf; Miki Nakanishi, for keeping me in shape (at least for a few years); Larry Cousens, for sharing a elegant apartment; Bruce Patterson, because without a hawk-boy there can be no hawk-king; Kathy Wilson, for introducing me to the pleasures of avian pets; and last and also least, Glenn Middleton and Rob Jensen, for nothing but trouble.

A Genetic Analysis of mRNA Splicing in *Saccharomyces cerevisiae*

Roy Parker

ABSTRACT

We have undertaken a genetic analysis of mRNA splicing in yeast. The cornerstone of this approach was the construction of a translational fusion between the yeast actin gene and the *HIS4* gene such that expression of the *HIS4* gene product requires proper splicing of the actin intron. We have utilized this gene both to screen *in vivo* for splicing mutations, and to characterize deletions constructed *in vitro* for their effect on splicing.

In an *in vivo* screen, we have isolated a point mutation which alters splicing of the fusion transcript (Chapter 1). The mutation is a G to A transition in the fifth position of the highly conserved sequence /GTAPyGT, found at the 5' junctions of all yeast introns (Chapter 2). The biochemical consequences of this mutation are threefold: 1) a decreased amount of correctly spliced mRNA is produced; 2) full-length precursor accumulates; and 3) a second intron-containing species accumulates, the 5' terminus of which is 6 bases 5' to the normal 5' junction. Surprisingly, this novel terminus arises via cleavage at a sequence that bears no resemblance to a consensus 5' junction. Moreover, cleavage at this abnormal site does *not* lead to the production of ligated mature mRNA, although this aberrant intermediate does appear to be in a lariat structure. The behavior of this mutant

argues that initial recognition of the 5' junction and subsequent cleavage are separable events and, furthermore, that requirements for 3' endonucleolytic cleavage may be more complex than previously imagined.

In a complementary approach we have used gene fusions to construct a series of internal deletions which remove the TACTAAC box or in addition TACTAAC-like sequences within the intron (Chapter 3). This work was carried out in collaboration with the laboratory of John Rossi (City of Hope). Analysis of the fusion transcripts produced in these deletions has allowed to us to conclude that the TACTAAC-like sequence, TACTAAG, can substitute, though inefficiently, for the correct TACTAAC box in the splicing process. This demonstrates that cryptic TACTAAC boxes can be used and suggests they at least have the potential to play a physiological role.

Finally, we have attempted to identify components of the splicing machinery by the selection of extragenic suppressors to these cis-acting splicing mutations (Chapter 4). However, by focussing on suppressors which confer a temperature sensitive lethal phenotype, we have restricted ourselves to a class of recessive suppressors whose gene products may be involved in the degradation of unstable transcripts. These suppressors define five complementation groups. Interestingly, the conditional lethal phenotype of these suppressors suggests that proper and efficient degradation of RNA is required for cell growth.

TABLE OF CONTENTS

LIST OF TABLES Page

LIST OF FIGURES Page

INTRODUCTION

Prelude

In writing this introduction I have attempted, not to review the entire field of RNA processing, but instead to highlight the key advances in mRNA splicing and their relationship to the results in this thesis. Unfortunately, there is no good recent review of RNA processing to which to refer more interested readers. However, I recommend Cech (1983) and Keller (1984) as potential starting points for a more in depth review of the RNA processing literature.

In the beginning....

RNA processing is a fundamental step in the expression of essentially all eukaryotic gene products. RNA processing refers to a collection of post-transcriptional modifications to the primary transcript. These alterations include such now classic reactions as modification of internal nucleotides, cleavage of 5' and 3' nucleotides, and addition of cap and polyA structures. In 1977 the RNA processing field, in addition to the entire scientific community, was astounded by the discovery that the genes of eukaryotes are not necessarily colinear with the mature functional products they encode. Instead, many genes contain interruptions in the coding sequence referred to as intervening sequences, or introns.

Introns have been described in each of the three major classes of

eukaryotic genes, those genes coding for tRNAs, rRNAs, and mRNAs. The variety of introns is underscored in that these three types of genes also define three structurally different classes of introns (Cech, 1983). The discovery of introns immediately raised two fundamental questions: 1) How are the intron sequences recognized? That is, what are the molecular signals which identify the intron to the splicing machinery? 2) What is the mechanism which removes the intron sequences and ligates the exons together? As reviwed by Cech (1983), the answers to these questions are fundamentally different for each of the three classes of introns. Though each of these mechanisms is inherently interesting, the focus of this dissertation is a genetic analysis of the splicing of introns from mRNA precursors in the yeast Saccharomyces cerevisiae.

Early experiments in the field of mRNA splicing focused on two important observations. First, sequencing numerous mammalian introns revealed conserved sequences at the 5' and 3' intron/exon junctions (reviewed by Mount, 1982). This led to the suggestion that the intron was demarcated by these consensus sequences (Breathnach et al., 1978). From this, a specific proposal was put forth in which the 5' end of a member of the small nuclear RNAs (snRNAs), U1, recognized the intron junctions by direct Watson-Crick base pairing (Lerner et al., 1980; Rogers and Wall, 1980).

The second early observation was that deletion of the consensus sequences at a splice junction invariably led to the activation of one

or more "cryptic" junctions (reviewed by Mount and Steitz, 1983). That is, sequences which resemble the consensus sequence are utilized instead of the mutated junction in the splicing process. This suggests a requirement for a consensus sequence in the use of a splice junction. However, it is easy to find sequences which agree with the consensus rules but are not normally used as splice sites. Thus, while splice junctions are required to conform to the consensus rules, fulfilling this criterion in itself is not sufficient to determine the point of splicing.

Early understanding of the splicing process was slow due to the lack of satisfactory experimental systems. The two most potent weapons in the molecular biologists' arsenal, a powerful genetic approach and an in vitro system, had not been brought to bear on the splicing problem. Yeast has the most powerful genetic system of any eukaryote. Thus, a logical extension of the discovery of introns in some pre-mRNAs in yeast (Ng and Abelson, 1980; Gallwitz and Sures, 1980), and the point of this thesis, was to undertake a genetic approach to this previously intractable problem.

The spliced genes in yeast consist almost exclusively of essential genes (Teem et al., 1984), which are therefore difficult to experimentally manipulate. However, we were able to circumvent this problem by the construction of a chimeric gene containing sequences from the yeast actin gene, including a 309 base pair intervening sequence (Ng and Abelson, 1980; Gallwitz and Sures, 1980), and the

structural gene for HIS4 (Donahue et al., 1982). As described in Chapter One, the extensive genetics of the HIS4 locus enabled us to begin to approach splicing genetically by attempting to identify mutations within the intron which affect splicing.

Since this time several key results have clarified our understanding of certain aspects of the splicing problem. Progress in the field of mRNA splicing has been greatly aided by the accumulation of numerous intron sequences from yeast. The 5' consensus sequence, though subject to variation about a consensus sequence in metazoan introns, is an essentially 100% conserved hexanucleotide in yeast (Langford and Gallwitz, 1983). This lack of variance in yeast splicing signals allowed the identification of an additional conserved sequence within the intron ,the TACTAAC box, a heptanucleotide located in all yeast introns 4-53 nucleotides upstream of the AG at the 3' end of the intron.

The importance of these conserved sequences in the splicing process was demonstrated by the observation that deletion of either the 5' consensus sequence (Gallwitz, 1982) or the TACTAAC box (Pikielny et al., 1983; Langford and Gallwitz, 1983) abolished splicing and led to the accumulation of full-length precursor. As described in Chapter Two, we have been successful in identifying a point mutation in the 5' consensus sequence which reduces splicing to 30-50% of wild-type levels. This confirms, and extends, the hypothesis that these sequences are important for the splicing process.

From the results discussed above, we can partially define a yeast mRNA intron as containing a 5' consensus sequence (GTAPyGT), a TACTAAC box, and an AG at the 3' intron/exon junction preceded by a polypyrimidine stretch. Though two of these features, a 5' consensus sequence and a polypyrimidine stretch followed by an AG, had been recognized in metazoan introns, there was no evidence to suggest an analog of the TACTAAC box.

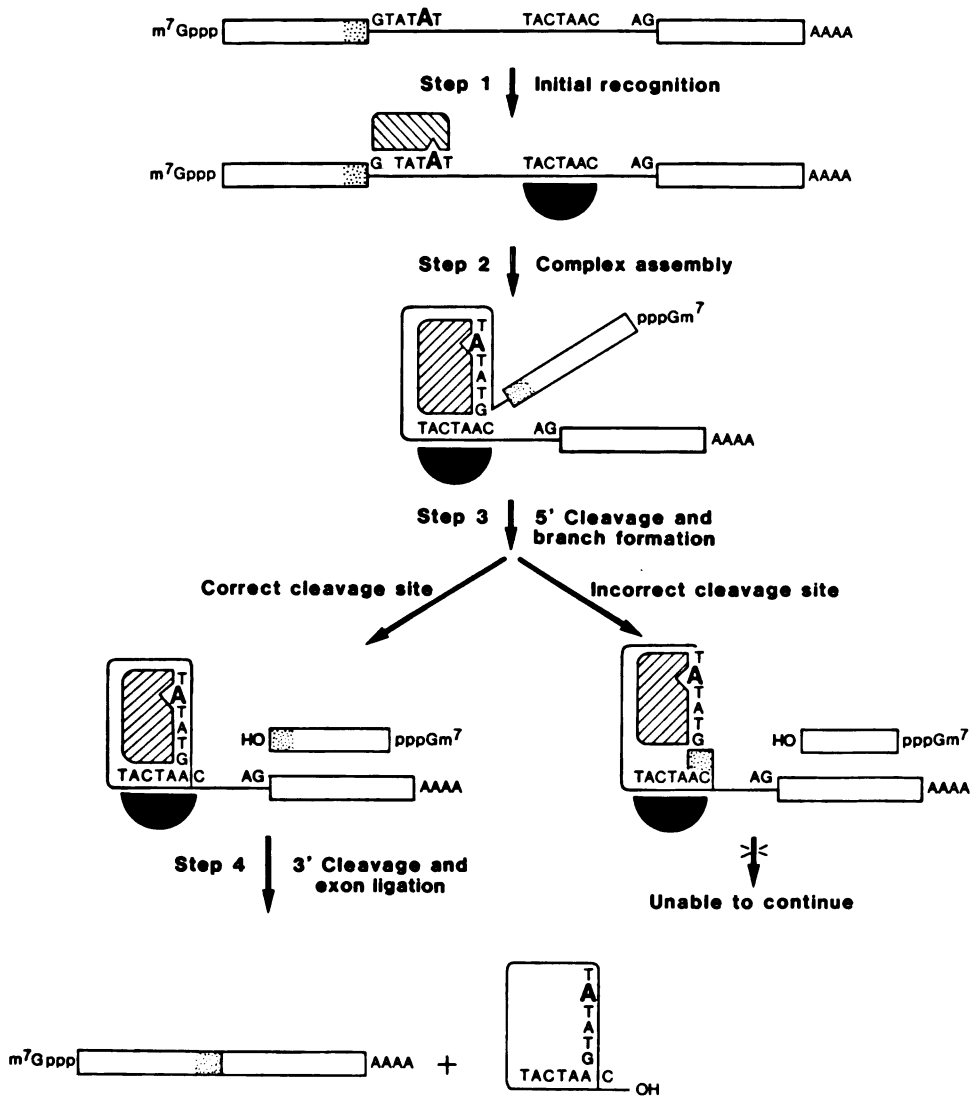
This apparent difference was resolved with the development of extracts which carry out the splicing reaction in vitro. A detailed analysis of the transcripts in both mammalian and yeast extracts (Ruskin et al., 1984; Padgett et al., 1984) and more recently the characterization of intermediates in the splicing process in vivo (Domdey et al., 1984; Rodriguez et al., 1984; Zeitlin and Efstratiadis, 1984) has shown that splicing of pre-mRNA in both systems occurs via a common intermediate. This intermediate has been termed a lariat, because the 5' end of the intron is joined by a 2'-5' phosphodiester bond to a site within the intron. The location of this 2'-5' linkage, or branch (cf. Wallace and Edmonds, 1983), is within the TACTAAC box in yeast (Domdey et al., 1984). In mammalian cells, branch points seem to fall in comparable locations (with respect to the 3' junction), but an obvious consensus sequence is less readily identifiable (Padgett et al., 1984; Ruskin et al., 1984; Keller and Noon, 1984). The distintive nature of these intermediates argues that splicing must proceed by fundamentally similar mechanisms in yeast and mammals, despite the stricter adherence to a specified primary sequence at the TACTAAC box

in yeast.

A Splicing Model

These recent results, taken together, suggest a plausible splicing pathway. In this pathway (diagrammed in Figure 1), the splicing process in initiated (Step 1) with the recognition of the 5' consensus sequence and the TACTAAC box by one or more trans-acting factors. The demonstration that the U1 containing snRNP will bind to a 5' splice junction *in vitro* is consistent with the suggestion that the recognition of the 5' junction is acheived by the direct base-pairing between the RNAs (Mount et al., 1983). The precursor and the various components of the splicing machinery then assemble into a splicing complex (Step 2). This step may involve some type of linear scanning (see, e.g., Lewin, 1980); in any case, the consequence is presumed to be the spatial juxaposition of the 5' and 3' intron elements.

The third step comprises cleavage at the 5' junction and formation of the 2'-5' phosphodiester bond at the TACTAAC box. While the details of this step are not known, an attractive hypothesis is that the cleavage of the phosphodiester bond at the intron/exon junction and the formation of the branch is a concerted process. In this view, cleavage is catalyzed by the nucleophilic attack of the 2' OH from the TACTAAC box on the phosphate at the 5' intron/exon junction. After formation of this lariat structure, the 3' exon is released by cleavage and ligated to the 5' exon (Step 4). In a


m7Gppp
GTATAT
TACTAAC
AG
AAAA
Step 1
Initial recognition
G TATAT
Step 2
Complex assembly
pppGm7
TACTAAC
Step 3
5' Cleavage and branch formation
Correct cleavage site
Incorrect cleavage site
HO
TACTAAC
Step 4
3' Cleavage and exon ligation
Unable to continue
OH

mechanism similar to cleavage at the 5' junction, the 3' OH generated at the 3' end of the 5' exon could attack the phosphate at the 3' intron/exon junction.

At this point it is important to point out the revelation that has occured recently with regards to the role that RNA can play in the metabolism of the cell. In a result that shattered the dogma of RNA being strictly structural in nature, the intron of the large rRNA of Tetrahymena was found to "self-splice" (Krueger et al., 1983). This, and the more recent observation that the RNA component of RNaseP, by itself, has enzymatic activity (Guerrier-Takeda et al., 1984), has demonstrated that RNAs can possess catalytic activities. In this light it is possible that the cleavage/ligation steps in mRNA splicing are the result of the use of reactive groups within the intron. The role of the splicing machinery would then be to position and activate these groups. From this, it is easy to imagine the primordial intron which retained enough structural information to position these reactive groups in a productive manner in the absence of trans-acting factors.

In the pathway described above (Figure 1), the splicing process is reduced to a series of discrete steps. Within the framework of this model, we can identify at least two roles that the 5' consensus sequence could play, first as part of the initial recognition complex (Step 1), and subsequently as a substrate for the 5' endonuclease (Step 3). As described in Chapter 2, this does appear to be the case. We have identified a point mutation which separates cleavage at the 5' junction

from intitial recognition of the intron. This formally demonstrates that the 5' consensus sequence can play multiple roles and in addition that these roles can be separated by mutation.

The observation that mutations could separate steps within the splicing pathway illustrates an important feature of the yeast splicing machinery; mutations which inactivate a splice junction do not appear to activate cryptic junctions, but instead lead to the accumulation of full-length precursor (Gallwitz, 1982; Chapter 2). Thus, the specific molecular consequences of the attempt to use a mutant splice junction are revealed. For this reason, the detailed characterization of additional point mutations in the yeast splicing signals should prove particularly informative.

However, though we have made considerable progress in understanding the "primary" splicing signals, the TACTAAC box and the 5' consensus sequence, one of the key questions still is unanswered: what completely defines an intron? It is necessary, but not sufficient, that an intron contain a 5' consensus sequence, a TACTAAC box, and a AG. However, these features are not sufficient in themselves to define an intron. For example, a region of the HIS4 transcript contains each of these sequence elements in an arrangement not dissimilar to an intron and yet is not spliced (Donahue et al., 1982; Parker, unpublished observations).

What are the other factors which determine which splicing signals

will be recognized and which will be ignored? At this point it does not appear as if there is an additional key signal at the primary sequence level which remains undiscovered. Instead, the final demarcation of the intron may be explained by understanding how the long-range interactions of the intron and flanking exons influences the three dimensional positioning of the splicing signals. For example, a secondary structure which positions the 5' consensus sequence and the TACTAAC box near each other may result in rapid recognition and subsequent processing by the splicing machinery. Structures which result in the a "poor" arrangement of the splicing signals may not be able to kinetically compete with either a favored arrangement of splicing signals, or with a non-splicing alternative, i.e. transport out of the nucleus or degradation.

In addition to the relationship of the primary splicing signals to each other, the secondary structure or sequence in the immediate vicinity of a splicing signal may modulate the efficiency with which a splicing signal is utilized. In principle this could either prevent the use of a site, perhaps by burying the signal in a strong secondary structure, or could prime a site for rapid recognition by containing multiple binding sites for the splicing machinery. This would be a particularly useful device for the TACTAAC box, in that the TACTAAC box does not directly determine the 3' end of the intron. (In contrast, by its nature of being located <u>at</u> the 5' intron/exon junction the 5' consensus sequence directly determines the point of splicing.) Thus, the inclusion of additional TACTAAC boxes, might increase the rate of

splicing while not demanding alternative splicing.

As discussed in Chapter 3, it is possible to find TACTAAC-like sequences in the vicinity of the TACTAAC box in several yeast introns. We have demonstrated that in the absence of a correct TACTAAC box, a TACTAAC-like sequence in the actin intron can be recognized, albeit inefficiently, by the splicing machinery. While further experiments are required to determine if this sequence plays any role in the wild-type intron, this TACTAAC-like sequence at least has the potential to play some role in the splicing process.

Finally, while progress on the roles these features of RNA secondary structure may play in splicing is forthcoming, the identification of the components of the splicing machinery itself may reveal the mechanism by which the intron is recognized. With the recent development of both powerful genetic and biochemical systems these answers should not be that far away.

REFERENCES

Breathnach, R., Benoist, C., O'Hare, K., Gannon, F., and Chambon, P. (1978). The ovalbumin gene: evidence for a leader sequence in mRNA and DNA sequences at the exon-intron boundaries. Proc. Nat. Acad. Sci. USA 75:10, 4853-4857.

Cech, T. R. (1983). RNA Splicing: three themes with variations. Cell 34 713-716.

Domdey, H., Apostol, B., Lin, R.J., Newman, A., Brody, E., and Abelson, J. (1984) Lariat structures are in vivo intermediates in yeast pre-mRNA splicing. Cell 39 611-621.

Donahue, T.F., Farahaubgh, P.J., and Fink, G.R. (1982). The nucleotide sequence of the HIS4 region of yeast. Gene 18, 47-59.

Gallwitz, D. (1982). Construction of a yeast actin gene intron deletion mutant that is defective in splicing and leads to the accumulation of precursor RNA in transformed cells. Proc. Nat. Acad. Sci. USA 79, 3493-3497.

Gallwitz, D., and Sures, I. (1980). Structure of a split yeast gene: complete nucleotide sequence of the actin gene in Saccharomyces cerevisiae. Proc. Nat. Acad. Sci. USA 77, 2546-2550.

Guerrier-Takada, C., Gardiner, K., Marsh, T., Pace, N., and Altman, S. (1983). The RNA moiety of ribonuclease P is the catalytic subunit of the enzyme. Cell 35, 849-857.

Keller, W. (1984). The RNA Lariat: a new ring to the splicing of mRNA precursors. Cell 39, 423-425.

Keller, E. B., and Noon, W.A. (1984). Intron splicing: a conserved internal signal in introns of animal pre-mRNAs. Proc. Nat. Acad. Sci. USA, in press.

Kruger, K., Grabowski,P.J., Zaug, A.J., Sands, J., Gottschling, D.E., and Cech, T. R. (1982). Self-splicing RNA: autoexcision and autocyclization of the ribosomal RNA intervening sequence of Tetrahymena. Cell 31, 147-157.

Langford, C.J., and Gallwitz, D. (1983). Evidence for an intron-contained sequence required for the splicing of yeast RNA polynerase II transcripts. Cell 33, 519-527.

Lerner, M.R., Boyle, J.A., Mount,S.M., Wolin, S.L., and Steitz, J.A. (1980) Are snRNPs involved in splicing? Nature 283, 220-224.

Lewin, B. (1980). Alternatives for splicing: recognizing the ends of introns. Cell 22, 324-326.

Mount, S. (1982). A catalogue of splice junction sequences. Nucl. Acids. Res. 10, 459-472.

Mount, S.M., Pettersson, I., Hinterberger, M., Karmas, M., and Steitz, J. (1983) The small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. Cell 33, 509-518.

Mount, S., and Steitz, J. (1983). Lessons from mutant globins. Nature 303, 380-381.

Ng, R., and Abelson, J. (1980). Isolation and sequence of the gene for actin in Saccharomyces cerevisiae. Proc. Nat. Acad. Sci. USA 77, 3912-3916.

Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F., and Sharp, P.A. (1984). Lariat RNAs as intermediates and products in the splicing of messenger RNA precursors. Science 225, 898-903.

Pikielny, C.W., Teem, J.L., and Rosbash, M. (1983). Evidence for the biochemical role of an internal sequence in yeast nuclear mRNA introns: implications for U1 RNA and metazoan splicing. Cell 34 , 395-402.

Rodriguez, J.R., Pikielny, C.W., and Rosbash, M. (1984). In vivo characterization of yeast mRNA processing intermediates. Cell 39, 603-610.

Rogers, J. and Wall, R. (1980). A mechanism for mRNA splicing. Proc. Nat. Acad. Sci. USA 77, 1877-1879.

Ruskin, B., Krainer, A.R., Maniatis, T., and Green, M.R. (1984). Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell 38, 317-331.

Teem, J.L., Abovich, N., Kaufer, N.F., Schwindinger, W.F., Warner, J.R., Levy, A., Woolford, J., Leer, R.J., van Raamsdonk-Duin, M.M., Mager, W.H., Planta, R.J., Schultz, L., Friessen, J.D., and Rosbash, M. (1984). A comparison of yeast ribosomal protein gene DNA sequences. Nucl. Acids. Res. 12:22, 8295-8312.

Wallace, J.C. and Edmonds, M. (1983). Polyadenylated nuclear RNA contains branches. Proc. Nat. Acad. Sci. USA 80, 950-954.

Zeitlin, S., and Efstratiadis, A. (1984). In vivo splicing products of the rabbit beta-globin pre-mRNA. Cell 39, 589-602.

CHAPTER ONE

In vivo Isolation of a Splicing Mutation

SUMMARY

To allow the analysis of mRNA splicing in yeast we have constructed a translational fusion between the yeast actin gene and the polyfunctional HIS4 gene such that expression of the HIS4 gene product requires splicing of the actin intron. By taking advantage of the powerful HIS4 genetics we have been sucessful in screening in vivo for point mutations which alter splicing of the fusion transcripts. This demonstrates the potential uses of the chimeric gene in further analysis of mRNA splicing.

GENERAL RATIONALE

The overall goal of this research project was to undertake a two step genetic analysis of mRNA splicing in Saccharomyces cerevisiae. In the first step, point mutations within the splicing substrate - in the intron or flanking exons - which disrupt splicing would be identified. The nature and location of these cis-acting mutations, together with information derived from the characterization of the resultant processing intermediates, are expected to suggest plausible molecular models for substrate recognition. In the second step, these point mutations could be used to define the splicing machinery genetically by the selection of extragenic suppressors.

To isolate cis-acting splicing mutations, two options were considered. First, we could create by in vitro site-specific mutagenesis, mutations predicted to have an effect on splicing. At the initiation of this project the only likely nucleotides for this type of manipulation were the GT/AG at the intron/exon boundaries. In that this approach is severely limited, we decided to undertake a more open-ended strategy. In this strategy, we would screen randomly for any mutation which decreases splicing in a biologically detectable manner. Thus, the experimenters are no longer limited by their prior bias. In addition, any mutations obtained are prepared for the generation of suppressors by the virtue of having a scorable biological phenotype.

We have focused on using point mutations within the substrate to

genetically define components of the splicing machinery by the selection of extragenic suppressors for two reasons. First, the identification of a suppressor of a cis-acting mutation not only genetically defines a component of the splicing machinery, but in addition suggests a likely site of action of the suppressor gene product, i.e. the site of the original mutation. In addition, the identification of specific suppressors, in contrast to screening for general mutations which block splicing, diminishes the likelyhood of identifying general nuclear functions which are required for secondary reasons for efficient splicing.

In order to undertake the approaches described above, it was required that we obtain a spliced yeast gene with a readily scorable genetic phenotype, allowing the easy identification of potential splicing mutations within the gene. The set of known spliced yeast genes consists predominantly of essential genes, i.e. genes for actin and many ribosomal proteins (Teem et al., 1984), which are difficult to experimentally manipulate. However, through the technique of gene fusion we created a chimera between the yeast actin gene, containing a 309 nucleotide intervening sequence (Ng and Abelson, 1980; Gallwitz and Sures, 1980), and the yeast HIS4 gene such that precise and efficient splicing of the actin intron is required for the production of functional HIS4 gene product. We refer to this fusion as the YAH (yeast-actin-HIS4) fusion.

The HIS4 gene was particularly useful for this fusion for the

following reasons. First, the HIS4 gene encodes a trifunctional gene product; thus the fusion has three separable activities. Each of these activities, HIS4A, B, and C, can be assayed independently by complementation tests with the appropriate HIS4 mutant. These three activities allow the rapid identification of potential splicing mutants as polar mutations, i.e. those mutations which cause the loss of all three activities. Finally, HIS4 is useful in that cis-acting mutations which prevent splicing will have a His- phenotype. This allows the selection of suppressors in a straightforward manner.

RESULTS

Construction of the Fusion and Transcript Analysis

The construction of the YAH fusion and the analysis of the transcripts produced are presented in detail in Chapter Two and in Appendix 1 at the end of this chapter. The important features are shown in Figure 1 and are summarized here. The fusion initiates transcription at the actin transcriptional start and terminates at the HIS4 terminator. Translation presumably initiates at the actin AUG, located 10 bases 5' to the 5' intron/exon junction, and terminates at the HIS4 stop codon. The chimeric transcript is processed to a 2.8 kb mature message by the accurate removal of the actin intron.

A useful feature of the YAH fusion is that while we can test all three activities independently by performing complementation tests, we can also assay directly for functional HIS4C gene product. The enzyme encoded by the HIS4C coding domain is histidinol dehydrogenase (Keesey et al., 1979), which catalyzes the last step in the histidine biosynthetic pathway. By growing cells on media lacking histidine but supplemented with histidinol (Hol), we can thus test directly for functional HIS4C activity (without requiring HIS4 A or B activities).

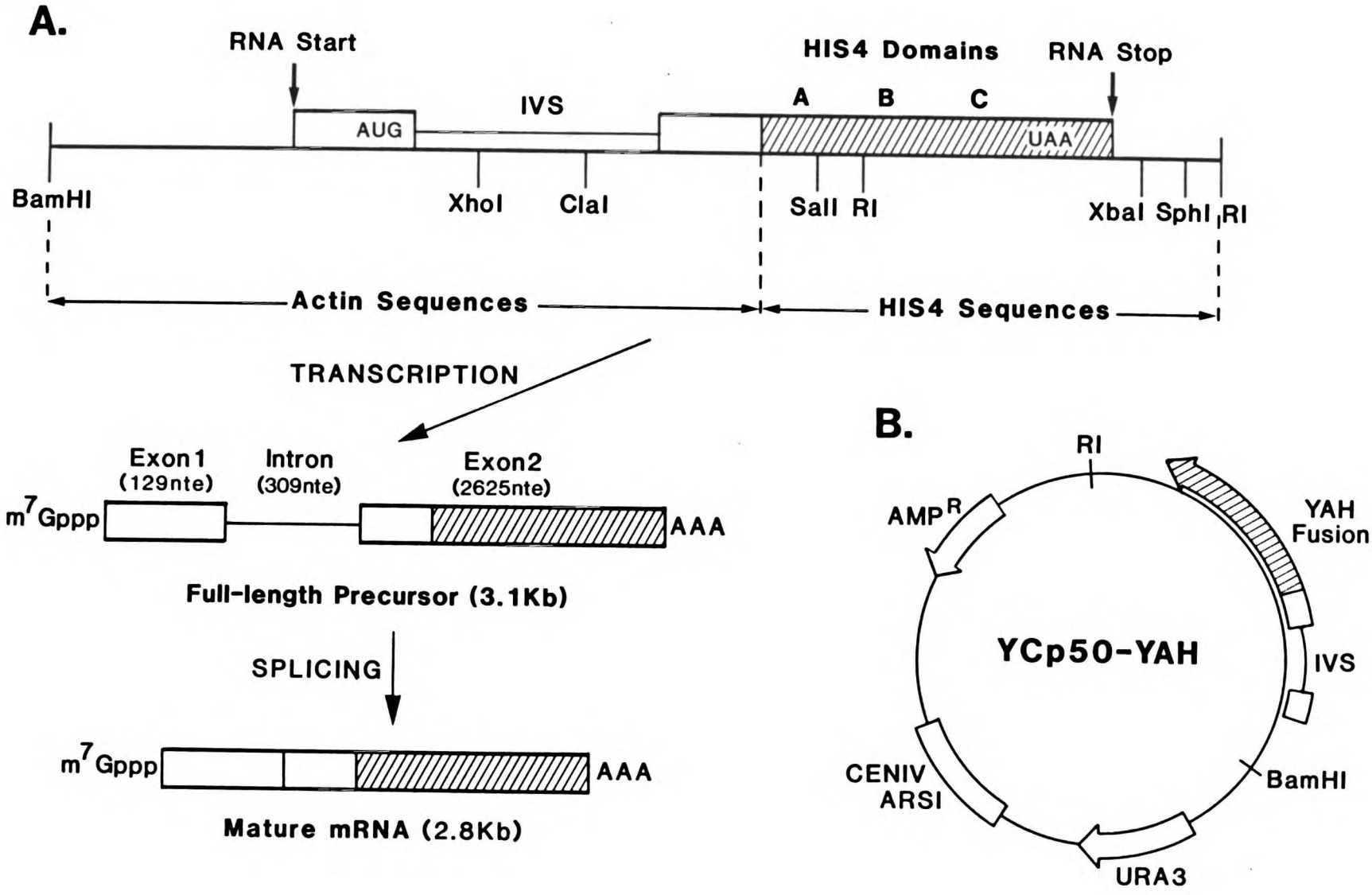

A.
RNA Start
HIS4 Domains
RNA Stop
IVS
A
B
C
AUG
UAA
BamHI
XhoI
ClaI
SalI
RI
XbaI
SphI
RI
Actin Sequences
HIS4 Sequences
TRANSCRIPTION
Exon1
(129nte)
Intron
(309nte)
Exon2
(2625nte)
m7Gppp
AAA
Full-length Precursor (3.1Kb)
SPLICING
m7Gppp
AAA
Mature mRNA (2.8Kb)
B.
RI
AMPR
YAH
Fusion
YCp50-YAH
IVS
BamHI
CENIV
ARSI
URA3

Figure 1 Construction of actin-HIS4 fusion

A) Schematic diagram of the YAH fusion; for details of the construction see Materials and Methods of Chapter 2. Transcribed regions are shown as raised boxes; open boxes correspond to actin sequences, slashed boxes correspond to HIS4 seqeunces. The fusion produces a 3.1 kb precursor which is processed to a 2.8 kb mature mRNA. In addition to the coding sequences, the fusion has 364 bases of 5' flanking sequence and 904 bases of 3' flanking sequence from the HIS4 gene. The fusion presumably initiates translation at the AUG normally used in the actin gene, 10 bases 5' to the 5' intron/exon junction, and terminates at the HIS4 terminator codon. The relative domains of the different HIS4 activities are shown above the fusion. Relevant restriction sites for experiments described are shown schematically. The diagram is not to scale.

B) Map of YCp50-YAH, a centromere plasmid containing the YAH fusion, which is maintained as a freely replicating single copy plasmid. Transformation of a HIS4 deletion strain with this plasmid allows the production of functional HIS4 A,B, and C activities.

Isolation of a Splicing Mutant

The overall scheme for the mutant hunt is diagrammed in Figure 2. In brief, a centromere plasmid containing the fusion was mutagenized *in vitro* with hydroxylamine and transformed into a suitable yeast strain. The resulting transformants were then screened genetically, by a number of criteria, for potential splicing mutants. Finally, mutants which passed all the genetic criteria were subjected to biochemical analysis to determine if there was any alteration in the processing of the fusion transcripts.

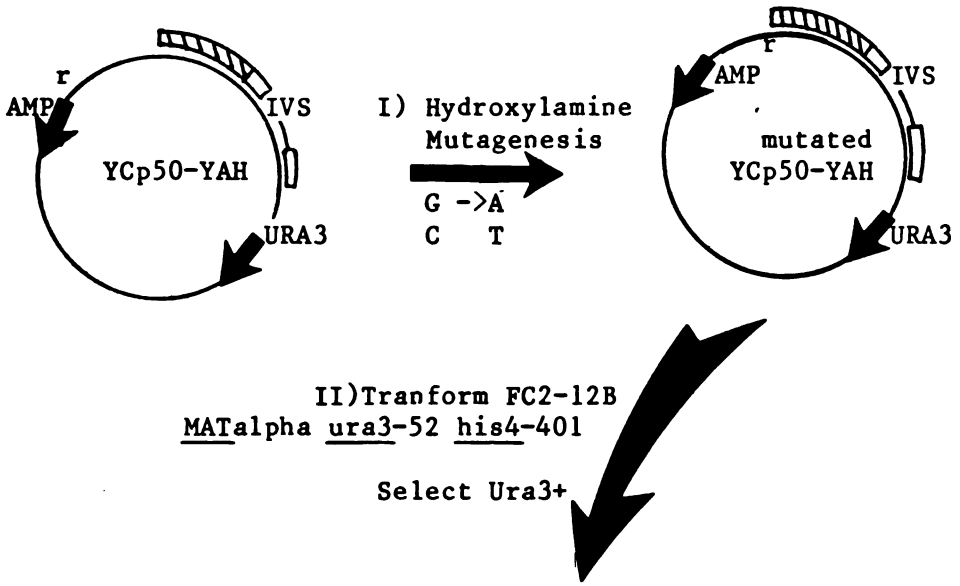
I) GENETIC ANALYSIS

Mutagenesis and Transformation

The YAH fusion was subcloned into the yeast centromere vector YCp50 to yield the plasmid, YCp50-YAH (shown in Figure 1). This plasmid contains the *URA3* gene for selection in yeast and confers ampicillin resistance in *E. coli*. The presence of the centromere maintains the plasmid at a single copy per cell (for review, see Carbon, 1984). Maintaining the fusion in single copy number was desirable in that leaky mutants might not confer a mutant phenotype when present in multiple copies.

The plasmid was mutagenized *in vitro* with hydroxylamine as

FIGURE 2



III) Screen for splicing mutants genetically

1) polar, i.e. HIS4A-B-C-
2) map to actin sequences
3) caused by single lesion
4) not nonsense mutations

IV) Screen potential mutants biochemically

FIGURE 2 Overall view of strategy for isolating splicing mutations in vivo.

described in detail in the experimental procedures of Chapter 2. This mutagen deaminates cytosine residues which, in the absence of error-prone repair, induces G to A and C to T transitions (Phillips and Brown, 1967). The DNA was then used to transform a HIS4 deletion strain, FC2-12B (see strain table for full genotype), selecting for the plasmid marker URA3+. For two of the time points, 60' and 120', a number of individual transformants were picked to master plates to allow a determination of the degree of mutagenesis. As shown in Table 1, after 120' of in vitro mutagenesis, 27% of the transformants were deficient in HIS4C activity, while 8% were lacking all three fusion activities. A useful (though certainly not quanitative) method of following the mutagenesis was evident in that the tranformation of E. coli dropped dramatically after hydroxylamine treatment (see Table 1). In order to obtain large numbers of transformants for screening, the transformants from DNA treated for 90' were resuspended, and replated on -Ura media.

As an initial test for potential splicing mutations, the transformants were screened for the failure to grow on Hol media, thus indicating a loss of HIS4C function. At this point transformants were also screened for the ability to complement HIS4A- and B- mutants. Transformants which failed to grow on Hol, and failed to complement the HIS4A and B- testers, were picked and analyzed further. Of the original ~2500 transformants, 58 HIS4A-B-C- mutants were obtained (see Table 2). The phenotypes of these mutants and their behavior in the subsequent genetic tests is described in detail in Appendix 1 of this chapter.

Time of Mutagenesis	%Transformation in E. coli	%Transformation in yeast	%HIS4C-	%HIS4ABC-
0'	100	100	NT	NT
30'	10	100	NT	NT
60'	3	100	6	1.7
90'	1	100	NT	NT
120'	0	100	27	8.3

TABLE 1 Hydroxylamine mutagenesis of YCp50-YAH

The table shows the result of transformation of E. coli and yeast with the mutagenized DNA. Transformation % in E. coli is based on the comparison of the number of transformants at 0' mutagenesis as 100%, transformants were selected for ampicillin resistance. Transformation into yeast is only a rough estimation of the transformation frequency by comparison of the number of colonies on the transformation plates, thus the transformation frequency could be varying but by no more then 3X. HIS4C activity was scored by growth on Hol media. HIS4A and B activities were scored by complementation with the strains 27/4 and E331 (see strain table).

Original Transformants	2500			
HIS4A-B-C-	58			
Deletion Map to:	Unable to map 11		HIS4 32	actin 15
A and B corevert with C	-		-	11
Not suppressible by SUP4o	-		-	9
Potential Splicing Mutants		Class I (leaky) 11		Class II (tight) 9

TABLE 2 Genetic Screens for Splicing Mutations

The table shows how the HIS4- mutants distributed in the genetic tests applied. Class I mutants are those mutants which were too phenotypically leaky to allow further genetic analysis. The tests are described in detail in the text.

These polar mutations were then analyzed by a set of additional genetic screens to identify those which might be affecting splicing. The criteria for a potential splicing mutant were that: 1) the mutation map to the actin sequences in the fusion; 2) the loss of all three activities be caused by a single event (i.e., be co-revertible); and 3) the mutation not be suppressible by nonsense suppressors.

Deletion Mapping

We reasoned that any mutations which prevent processing should be in the spliced sequences, i.e. the actin sequences. Mutations which are located in the HIS4 sequences in the fusion are not expected to affect splicing. To map the A-B-C- mutations with respect to their relative position in the fusion (i.e., in actin vs. HIS4), we took advantage of the presence of the chromosomal actin gene in the host strain, FC2-12B. That is, mutations within the actin sequences of the fusion should be repairable by gene conversion from the actin gene in the chromosome, whereas mutations within the HIS4 sequences cannnot be repaired due to the deletion of the chromosomal locus (see Figure 3).

The mutants were plated in 1" by 1" patches on -Ura media and grown for two days at 30 C. The patches were then replica-plated to Hol media and incubated for 4-7 days. At the end of the incubation, patches were scored for the number of Hol+ papillae generated. As summarized in Table 2, the mutants separated into two classes, those which can papillate on Hol media (thus mapping to actin sequences), and those

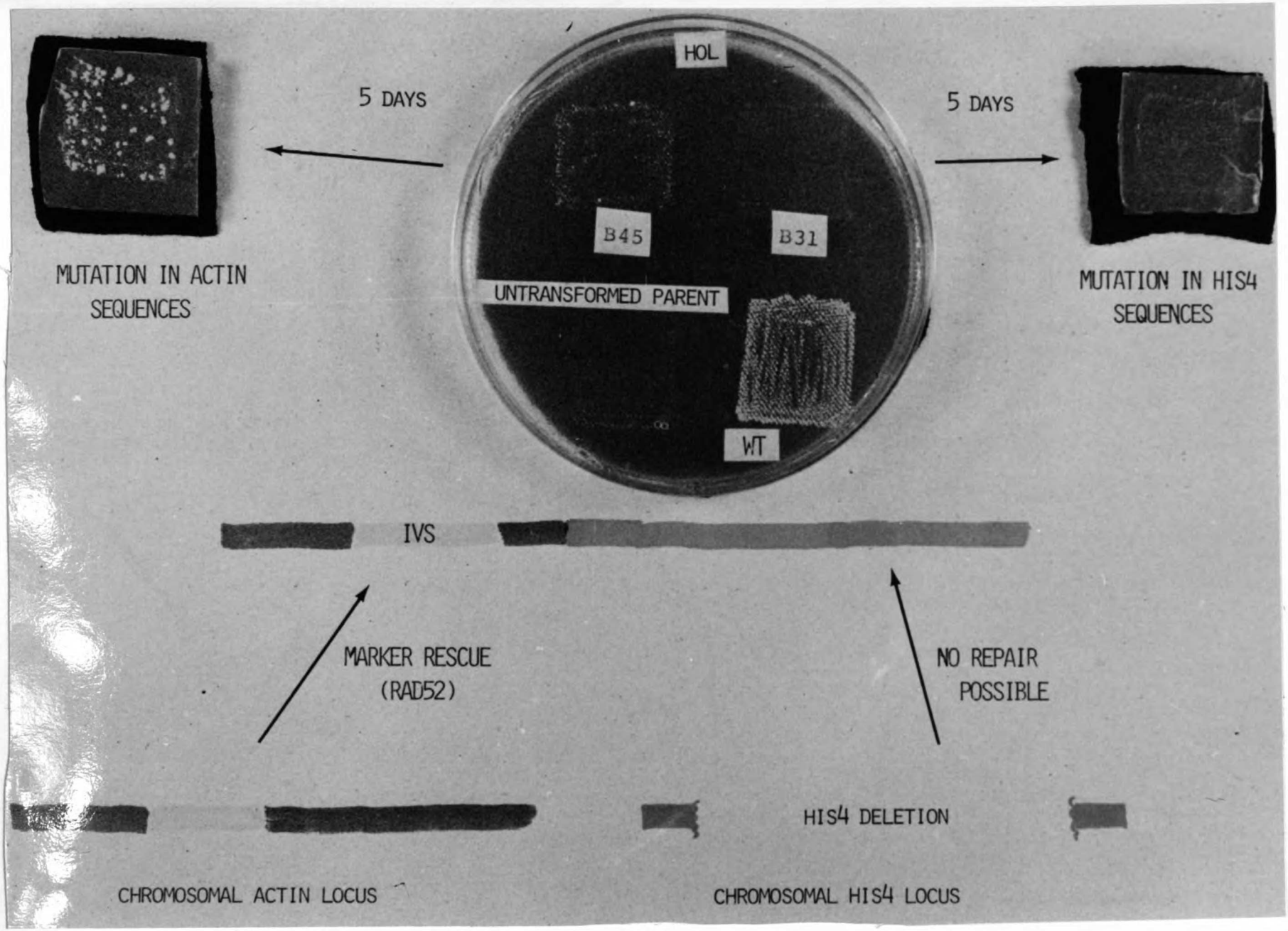

HOL
5 DAYS
5 DAYS
B45
B31
UNTRANSFORMED PARENT
WT
MUTATION IN ACTIN SEQUENCES
MUTATION IN HIS4 SEQUENCES
IVS
MARKER RESCUE (RAD52)
NO REPAIR POSSIBLE
HIS4 DELETION
CHROMOSOMAL ACTIN LOCUS
CHROMOSOMAL HIS4 LOCUS

FIGURE 3

The figure shows the rationale behind the deletion mapping of the polar Hol- mutations. As an example a replica plate of three types of fusion-containing strains on a Hol plate is shown. The plate shown in the center of the figure was incubated for two days at 30C while an identical replica was incubated for five days. From the longer replica, the small patches on the outside of the figure, showing the papillation of B45, and the failure of B31 to papillate, were cut out and photographed. The strain labeled WT is transformed with the wild-type fusion.

which fail to papillate (mapping to HIS4 sequences). In addition, some of the mutants were too leaky for this analysis to work.

In order to convince ourselves that the papillation we were scoring was indeed due to a gene conversion event, we performed two controls. We reasoned that if the papillation is due to a gene conversion event, then mutations which inhibit gene conversion (for example, rad52) should prevent this repair (for review, see Szostak, 1983). To test this, three of the papillating mutants were transferred into a rad52 strain, FC4-17B, (for full genotype see strain table) and scored for the ability to papillate on Hol media. As shown in Table 3, the absence of functional rad52 gene product decreased the frequency of reversion of these mutants.

An additional prediction of the gene conversion explanation was that if we supplied HIS4 coding sequences in trans, mutations which failed to papillate because they map to the HIS4 sequences should now be able to papillate. To test this, we mated the collection of mutants to a HIS4 mutant strain, FC0-7 (see strain table for full genotype), and scored the resulting diploids for the ability to papillate. As expected, all mutants were able to papillate in this diploid.

We concluded from these results that the papillation observed was the result of gene conversions between the resident actin gene and the fusion contained on the plasmid. (Repair events arising by reciprocal

Mutant	λpapilli RAD52+	λpapilli rad52-
B60	11	0
B45	16	1
F120	6	2

TABLE 3 RAD52 Dependence of Mutant Papillation

The original mutant isolates for the mutants shown were crossed to FC4-17B, a rad52- strain, and dissected. Segregants carrying the plasmid and rad52- were compared in their papillation frequency to Hol+ with the original isolate. The numbers are the average of two independent papillation assays.

recombination are not expected to be viable as this will produce a cell unable to produce the essential gene product actin.) This allowed us to very quickly identify 15 mutations which map to the actin sequences amoung the 47 testable polar mutants (Table 2).

Co-Reversion

Because of the heavy in vitro mutagenesis of the fusion, it was possible that the inactivation of the three fusion activities in some of the mutants was the result of multiple mutations. In that this would confuse further analysis, we desired to work with mutants where a single lesion was responsible for the loss of all three activities. If the loss of all three activities is the result of a single lesion, then the reversion of HIS4C, selected as Hol+, should also restore A and B activities. Several Hol+ revertants were isolated for each mutant and tested for the co-reversion of A and B activities. As shown in Table 2, in 11 out of the 15 mutants tested, the HIS4 A and B activities co-reverted with HIS4C.

Nonsense Suppression

In addition to mutations affecting splicing there are other types of mutations which could confer the same genetic behavior. These include mutations preventing or decreasing transcriptional initiation, transport of the mRNA to the cytoplasm, translational initiation, or

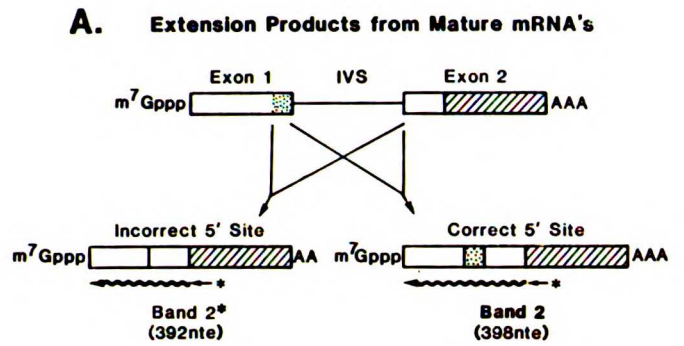
promoting chain termination (i.e. nonsense mutations). Although genetic screens for the former classes are not readily available, it is possible to determine if any of the potential mutants could be suppressed by nonsense suppressors. Consideration of the actin sequence and the mutagen specifity suggests that 9/10 of the possible nonsense mutations would be ochres. To test if any of the mutations were suppressible by nonsense suppressors, each of the mutants was crossed to a strain, FCO-6, with the relevant genotype MATa, SUP4-o, and his4-1176, and the resulting diploids tested for the ability to grow on Hol media. Of the strains tested, two carried mutations suppresible by SUP4-o.

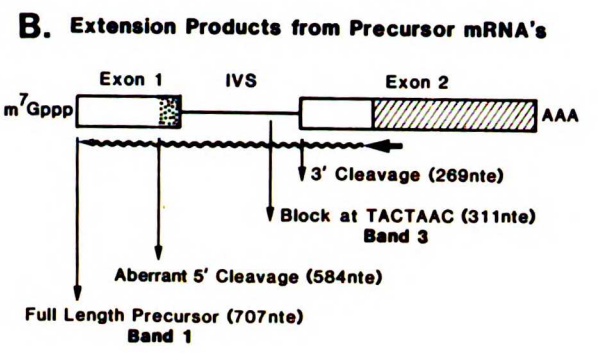
BIOCHEMICAL ANALYSIS

The genetic analyses identified a collection of potential splicing mutants. These mutants fell into two classes; those which had passed all the genetic criteria, i.e. were single, polar mutations which mapped to the actin sequences in the fusion, and a second class which contained mutants which were phenotypically too leaky for the genetic analysis to be carried out. It is important to point out that, though leaky, this class was comprised of mutants in which all three acitivies were leaky to approximately the same degree. To determine if any of these mutations affected splicing, the fusion transcripts produced were analyzed by primer extension analysis to see if there was a change in the ratio of precursor to mature mRNA.

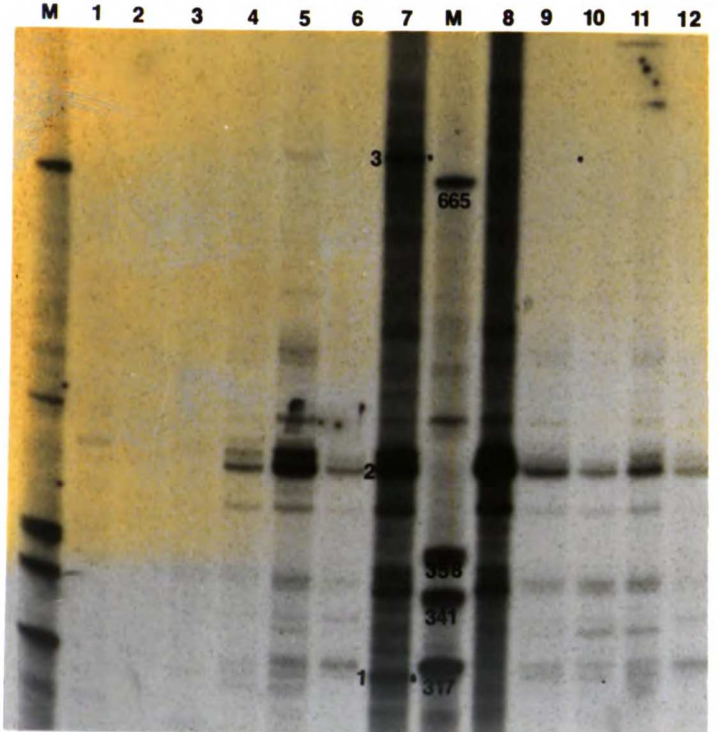
RNA preparations for each mutant were made from 50 ml cultures by the glass bead lysis technique described in the experimental procedures of Chapter 2. 50 micrograms of total RNA from these preps was subjected to primer extension analysis from a primer located in the 3' exon. This primer has the advantage of being specific for the fusion. In addition, because the 3' exon is found in both precursor and mature species, this primer will produce cDNA products from both spliced and unspliced transcripts (see schematic to Figure 4). In this experiment (see Figure 4C) the majority of the mutants gave rise to cDNA extension products which correspond to mature mRNA (Band 2), while a few mutants, B2, B39, B45, produced cDNA extension products the same length as expected for full-length precursor (Band 3, lane 18, lane 3 and data not shown). It is important to point out that because the cDNA product corresponding to the mature mRNA is the same size as in the wild-type construct, the mutants which produce mature RNA are producing accurately spliced mRNA (see Figure 4A).

The results of these analyses identified three mutants which had an increase in precursor RNA and a concominant decrease in mature mRNA. Yeast DNA from these strains was prepared and used to transform E. coli selecting for ampicillin resistance. Examination of the plasmids indicated that in all three cases the unique BamHI site 5' to the transcriptional start (Figure 1) had been inactivated. This and the identical behavior of all three mutants with regards to the accumulation of an aberrant splicing intermediate (see Chapter 2) led to the conclusion that all three mutants arose by the same event. These


A.
Extension Products from Mature mRNA's
Exon 1
IVS
Exon 2
m7Gppp
AAA
Incorrect 5' Site
Correct 5' Site
m7Gppp
AA
m7Gppp
AAA
Band 2*
(392nte)
Band 2
(398nte)


B.
Extension Products from Precursor mRNA's
Exon 1
IVS
Exon 2
m7Gppp
AAA
3' Cleavage (269nte)
Block at TACTAAC (311nte)
Band 3
Aberrant 5' Cleavage (584nte)
Full Length Precursor (707nte)
Band 1

C.


M 1 2 3 4 5 6 7 M 8 9 10 11 12
3
665
2
358
341
1
317

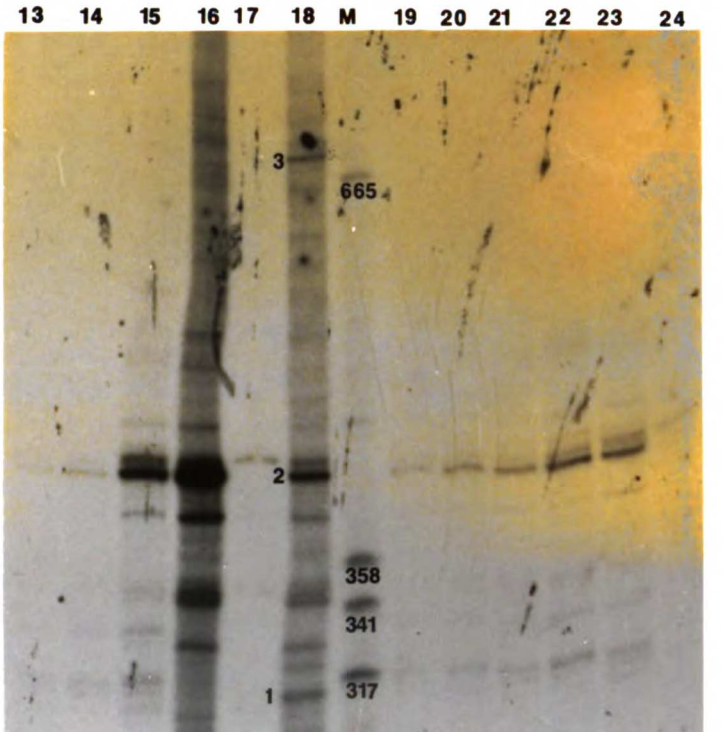

13 14 15 16 17 18 M 19 20 21 22 23 24
3
665
2
358
341
1
317

FIGURE 4

A) The schematic shows the expected extension products from mature mRNA. In addition, in the event of the use of an alternative splice site the expected variation in size of the cDNA is shown. In this example, a site is indicated 6 bases 5' to the correct 5' splice junction.

B) The schematic shows expected extension products from precursor mRNAs.

C)The gel show the results of primer extensions of the various mutants. Lanes are as follows; lane 1, untransformed FC2-12B; lane 2, B1; lane 3, B2; lane 4, B3; lane 5, B6; lane 6, B43; lane 7, B45 polyA+ RNA (band 1, block at TACTAAC box, band 2, mature mRNA, band 3, full-length precursor); lane 8, wild-type polyA+ RNA; lane 9, B9; lane 10, B11; lane 11, B13; lane 12, B14; lane 13, B30; lane 14, B28; lane 15, B36; lane 16, wild-type polyA+ RNA; lane 17, B38; lane 18, B45 polyA+ RNA; lane 23, WT; lane 20, B26; lane 21, B20; lane 22, B19; lane 23, B16; lane 24, untransformed FC2-12B. A repeat of the experiment for B1 and B2 showed that B1 produced mature mRNA and that B2 was identical to B45. While the absolute amounts of extension products vary from mutant to mutant, Northern analysis has indicated that this is the result of varying amounts of RNA in the preparations. PolyA+ RNA was made from wild-type and the earliest identified splicing mutant, B45, to provide markers.

isolates were not independent. Thus, only one of these three, B45, was analyzed further (see Chapter 2). This mutation is referred to as the A5 mutation in later chapters and is phenotypically tighter at high temperatures (36 C).

DISCUSSION

In this chapter I have described the results of an *in vivo* mutant hunt for splicing mutations. This approach was possible because we could use the powerful genetics of the *HIS4* locus to narrow down the class of prospective mutants to a workable number. The analysis of these mutants identified one mutation which affects the splicing of the actin intron. The mutation is a C->A transition of the fifth nucleotide of the intron in the highly conserved hexanucleotide, GTAPyGT, found at all yeast 5' intron/exon junctions (see Chapter 2). The consequences of this mutation are complex and are described in detail in Chapter 2.

In that the mutant hunt only identified a single mutation it raises the question of why additional lesions were not found. A reasonable answer to this question is now possible because of progress in the yeast mRNA splicing field over the last two years. Deletion analysis of intron sequences has revealed two regions of the intron which are required for splicing (Gallwitz, 1982; Langford and Gallwitz, 1983; Pikielny et al., 1983). These are the essentially 100% conserved sequence at the 5' junction, generally called the 5' consensus sequence, and the TACTAAC box , a conserved sequence near but not at the 3' end of the intron (see Introduction). Though other regions of the intron appear to affect the efficiency of the splicing process, these are the only sequences which are absolutely required. In that there are only a few G and C residues in this target (the mutagen used was specific for the deamination of Cs), the number of possible

mutations is limited to 5 (including the AG at the 3' intron/exon junction).

Very recently a number of investigators have constructed point mutations by site specific mutagenesis in vitro in yeast introns within these key sequences (M. Rosbash, pers. comm.; L. Fouzer, pers. comm.; J. Abelson, pers. comm.). While the analysis of these mutants is still in progress two general rules are beginning to take shape: 1) most single point mutations do not have a dramatic affect on splicing, and 2) transitions exert a weaker phenotype than a transversion in the same position. From these results it is possible to predict that the only nucleotides which would have been hit by the mutagen and still given a detectable phenotype are the G at the 5' junction, the G in the 5th position of the intron, and the G at the 3' end of the intron (see Figure 5). In this light it is not so surprising that only a single splicing mutation was isolated in this screen.

If indeed the class of point mutations affecting splicing is fairly small, the question arises as to the defects in the other mutants picked up in this screen. These mutants are polar, map to actin sequences, and are not suppressible by an ochre tRNA suppressor. If one assumes the mutagen specifity is maintained even after in vivo repair then there are a few expected lesions: 1) mutation of the ATG (although in this case the next ATG is in the correct frame); 2) creation of a new ATG upstream of the correct ATG, this would cause translation to initiate at an incorrect position and would terminate prematurely (such

FIGURE 5

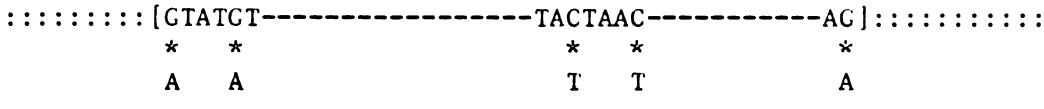


FIGURE 5 Potential splicing mutations induced by hydroxylamine

The essential primary sequences required for intron removal are shown in the schematic intron. The potential mutations induced by hydroxylamine are shown below. The alteration of the first G to an A has been constructed recently in J. Abelson's lab: the alteration prevents efficient splicing. The second G, the fifth nucleotide of the intron, replaced by an A is the lesion isolated (see Chapter 2). Conversion of the first C in the TACTAAC box to a T only still allows 80-90% of wild-type splicing (J. Rossi, pers. comm.), whereas the transversion of the last C in the TACTAAC box to a G still allows 80-90% splicing (L. Fouzer, pers. comm.). In view of this a T in this position may only have a very limited phenotype. There is no information to date about the G at the 3' intron/exon junction.

a mutant has been described for the *cyc1* gene (Stiles et al., 1981)); 3) an amber nonsense mutation (only ochres were tested for) ; and 4) unpredicted mutations, for example a point mutation which prevents nuclear transport of the mRNA would not be able to produce functional gene product.

On the other hand, it may be unfair to assume that mutagen specificity is maintained after transformation. Transformation of mammalian cells with deaminated DNA not only leads to the production of the expected transitions but in addition small insertions and deletions are found (Calos and Miller, 1984). In that either of these alterations could change the reading frame, these are additional explanations for the polar mutations which do not appear to affect splicing.

MATERIALS AND METHODS

The materials and methods for this chapter are all described either within the chapter itself or are found in the experimental procedures of Chapter 2.

STRAIN LIST

MY NAME	ALIAS	GENOTYPE	SOURCE
FC0-1	TD77	MATa ura3-52 his4-401 inol	
FC0-2	HR145-5D	MATalpha ura3-52 leu2-3,leu2-112, trp1 ade5	
FC0-3	7375-2A	MATa leu2 HOL1 his4-delta29	
FC0-4		MATalpha trp1-189 leu2 ura3-52 his3deletion rad52	
FC0-5		MATalpha rna2 trp1 leu2 arg4	
FC0-6	LC1-32B	MATa SUP4o his4-1176 ade2-1 trp5-48	
FC0-7	LC10-16C	MATa leu2-1 ura3 ura4 ade2-1 his4-1176 trp5-48 lys2	
FCT-1	E260	MATa his4-260 (A-B-C-) trp1	
FCT-2	1050C	MATalpha his4-260 (A-B-C-)	
FCT-3	1651/7	MATa his4-25 (A-)	
FCT-4	27/4	MATa his4-588 (A-)	
FCT-5	E588	MATalpha his4-588 (A-) ade2-1	
FCT-6	E331	MATa his4-331 (B-) trp1-1	
FCT-7	E594	MATalpha his4-594 (B-) ade2-1	
FCT-8	5942-1C	MATa his4-864 (C-)	
FCT-9	5942-1D	MATalpha his4-864 (C-)	
FC1-6B leu2		MATalpha trp1 ade5 ura3-52 his4-delta29 HOL1	
FC2-12B		MATalpha ura3-52 his4-401 leu2 trp1 HOL1 can1	
FC4-17B sac1 can1		MATa ura3-52 his4-401 leu2 trp1 rad52 HOL1	

REFERENCES

Carbon, J. (1984). Yeast centromeres: structure and function. Cell, 37, 351-353.

Donahue, T.F., Farabaugh, P.J., and Fink, G.R. (1982). The nucleotide sequence of the HIS4 region of yeast. Gene 18, 47-59.

Gallwitz, D. (1982). Construction of a yeast actin gene intron deletion mutant that is defective in splicing and leads to the accumulation of precursor RNA in transformed cells. Proc. Nat. Acad. Sci. USA 79, 3493-3497.

Gallwitz, D., and Sures, I. (1980). Structure of a split yeast gene: complete nucleotide sequence of the actin gene in Saccharomyces cerevisiae. Proc. Nat. Acad. Sci. USA 77, 2546-2550.

Kessey, Jr., J.K., Bigelis, R., and Fink, G.R. (1979). The product of the HIS4 gene cluster in Saccharomyces cerevisiae. J. Biol. Chem. 254, 7427-7433.

Langford, C., and Gallwitz, D. (1983). Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcripts. Cell 33, 519-527.

Ng, R., and Abelson, J. (1980). Isolation and sequence of the gene for actin in Saccharomyces cerevisiae. Proc. Nat. Acad. Sci. USA 77, 3912-3916.

Phillips and Brown. (1967) Mutagenic action of hydroxylamine. Progr. Nucl. Acid Res. Mol. Biol., 7, 349-368.

Pikielny, C.W., Teem, J.L., and Rosbash, M. (1983). Evidence for the biochemical role of an internal sequence in yeast nuclear mRNA introns: implications for U1 RNA and metazoan splicing. Cell 34 , 395-402.

Stiles, J.I., Szostak, J.W., Young, A.T., Wu, R., Consaul, S., Sherman, F. (1981). DNA sequence of a mutation in the leader region of the yeast Iso-1-cytochrome C mRNA. Cell 25, 277-284.

Szostak, J.W. (1983). Replication and resolution of telomeres in yeast. Cold Spring Harbor Symp. Quant. Biol. 47, 1187-1194.

Teem, J.L., Abovich, N., Kaufer, N.F., Schwindinger, W.F., Warner, J.R., Levy, A., Woolford, J., Leer, R.J., van Raamsdonk-Duin, M.M., Mager, W.H., Planta, R.J., Schultz, L., Friesen, J.D., and Rosbash, M. (1984). A comparison of yeast ribosomal protein gene DNA sequences. Nucl. Acids. Res. 12:22, 8295-8312.

APPENDIX 1

Abundance of Fusion Transcripts

Summary

The examination of the amounts of YAH fusion transcripts as compared to the transcripts from the resident actin gene in a number of constructions demonstrates that the fusion produces ≤10% of the expected amount despite the presence of all the promoter elements. The behavior of these constructs leads to the suggestion that the fusion transcripts produced are inherently unstable.

Results

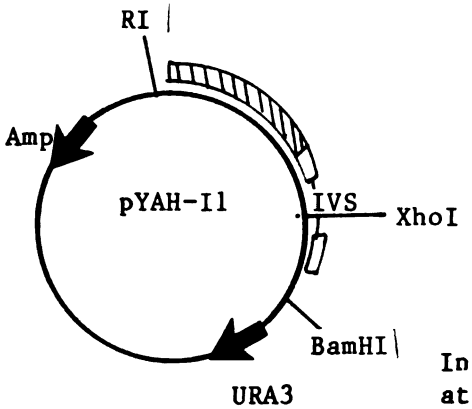
In this appendix three constructs of the YAH fusion will be discussed. The three forms are: (1) integrated at HIS4, pYAH-I2; (2) integrated at the actin locus, pYAH-I1; and (3) maintained on a centromere plasmid, YCp50-YAH (see Figure 1, Chapter 1). The maps of the integrating plasmids and the resulting integrated forms are shown in Figure 1.

For strains with the fusion plasmid integrated at HIS4 or at the actin locus, RNA has been prepared and analyzed on Northern gels. The results of these gels are shown in Figure 2. In summary, for both cases the total amount of fusion transcript is ≤ 10% of the actin transcripts. This suggests that the fusion either is ineffiecntly transcribed, or that the fusion transcript is unstable.

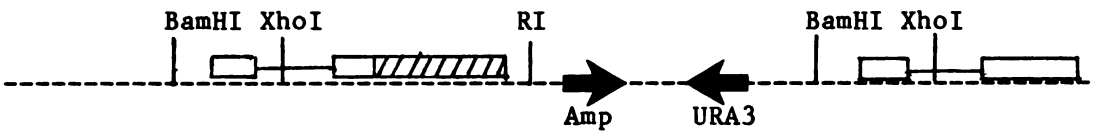
FIGURE 1 Integrated forms of the YAH fusion

The plasmid maps and the resulting integrated forms of pYAH-I1 and pYAH-I2 are shown; A, pYAH-I1 integrated at the actin locus; B, pYAH-I2 integrated at the HIS4 locus. Transcribed regions of interest are shown as either open boxes (actin sequences) or as slashed boxes (HIS4 sequences). The diagrams are not to scale and only relevant restriction sites are shown. The his4-401 allele is a large deletion between the PvuII site 5' of the HIS4 gene and the PvuII site located within the HIS4C coding domain. (Donahue et al., 1982).

A.

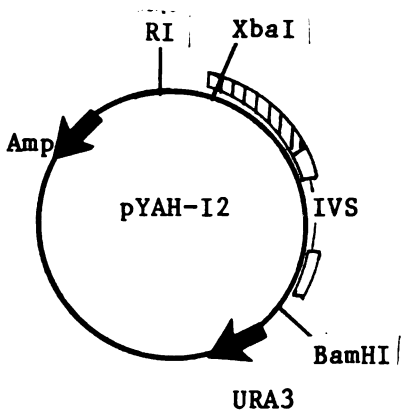


Integrated at actin by cleavage at XhoI site

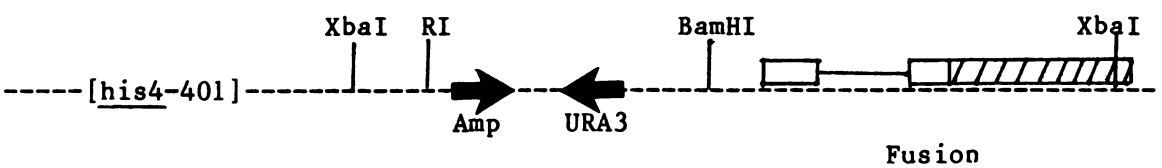


5' Flanking Sequence (actin)

B.



Integrated at HIS4 by cleavage at XbaI site.



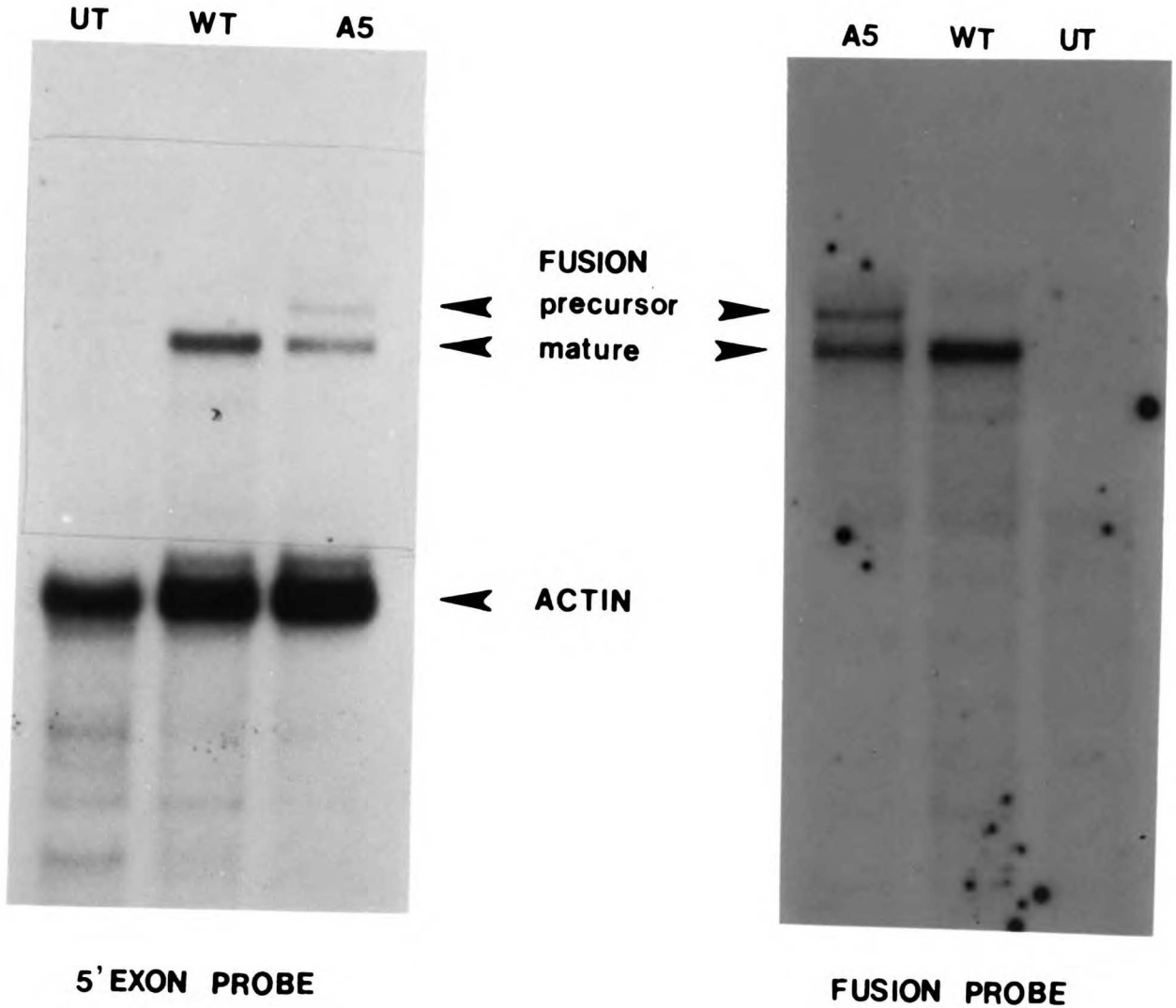
The direct comparison of the resident actin gene transcripts and the centromere form of the fusion has not been done. However, all three forms of the fusion have the same phenotype for growth on Hol plates (data not shown). This suggests that the centromere form of the plasmid allows the same amount of fusion production as the integrated forms.

Discussion

The basis for the reduced amount of fusion transcripts produced could in principle be either a decrease in the transcriptional rate of the gene (as compared to the resident actin gene), or a decrease in the transcript stability.

Elements which have been shown to be important in the proper functioning of yeast promoters include the canonical "TATA" box and the UAS, or upstream activator sequence. These sites for the actin gene have been mapped by John Rossi's group at the City of Hope. As shown in Figure 3 (J. Rossi, pers. comm.) these sites are all included within the BamHI site which delimits the 5' flanking sequences in all these constructs. In addition, the insertion of a tRNA gene at this BamHI site appears to have no effect on the transcription of the actin gene (Ng et al., 1985). Thus, it appears that there are no sequences 5' of the BamHI site flanking the actin gene which are required for transcription. The constructs have all the elements of the actin promoter, yet produce only 10% as much steady state RNA.

A.


UT
WT
A5
FUSION
precursor
mature
ACTIN
5' EXON PROBE
A5
WT
UT
FUSION PROBE

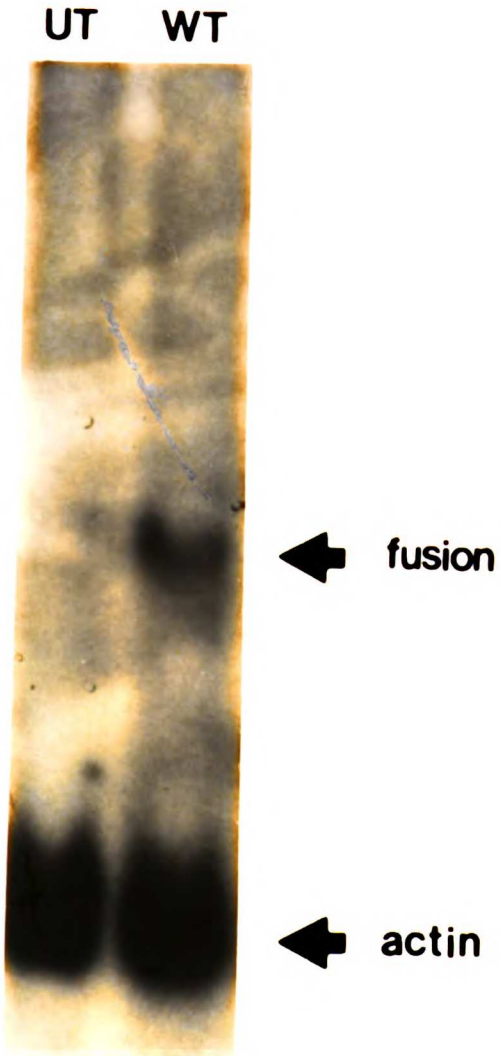

UT
WT
fusion
actin

FIGURE 2 Northern analysis of Fusion Transcripts

A. Northern analysis of polyA+ RNA from the fusion integrated at the HIS4 locus in FC8-24D. 5 micrograms of glyoxalated RNA was electrophoresed on a 1.5% agarose gel. The probe for the 5' exon is the kinased oligonucleotide 5'TGTTAATTCAGTAAATTTTCG3'. The probe for the fusion is the oligonucleotide located at the fusion junction, described in detail in the legend to Figure 4, in Chapter One. Throughout the figure abreviations are: UT, untransformed parent; WT, wild-type fusion; A5, A5 mutant fusion.

B. Northern analysis of 50 micrograms of total RNA from the fusion integrated at the actin locus in the strain TD77. The probe for this blot was the nick-translated actin plasmid, pYact-1 (Ng and Abelson, 1980).

The possible importance of the 5' flanking sequences upstream of the BamHI site is ruled out by examination of the fusion integrated at the actin gene (Figure 1A). As a result of the site of integration, in this construct the fusion now has the entire 5' flanking sequences of the actin gene, while the resident actin gene is limited to 5' sequences extending to the BamHI site. In that this construct again produces less fusion mRNA (see Figure 2b) the lack of sequences 5' of the BamHI site cannot explain the lower levels of fusion transcripts.

We are thus left with a situation where the YAH fusion has all the promoter elements but still produces a decreased level of transcripts. The most likely explanation is that the fusion transcripts produced are inherently unstable. A formal possibility that these experiments do not rule out is the presence of a site required for full levels of expression which is located 3' of the actin gene. Though this would be analogous to enhancer sequences in higher eukaryotes, it would be a novel observation for the transcription of a yeast gene.

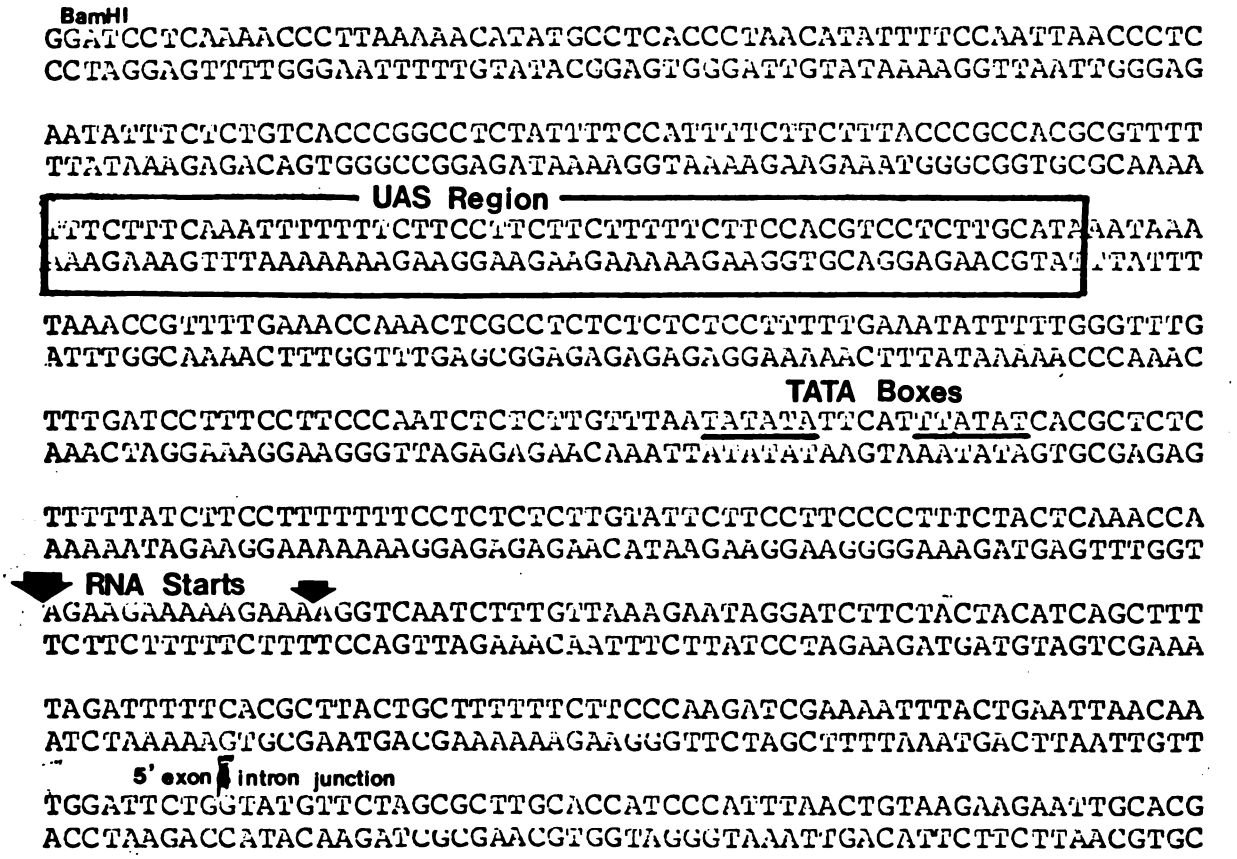

BamHI
UAS Region
TATA Boxes
RNA Starts
5' exon
intron junction

FIGURE 3 Promoter Elements of the Actin Gene

The 5' region of the actin transcript and the 5' flanking sequences to the BamHI site are shown. The two observed RNA starts are shown as vertical arrows. The "TATA" boxes are underlined and the UAS region is shown as a boxed area. The UAS region is defined as a region which when deleted still allows a low level of expression of beta-galactosidase activity in a actin-*lacz* fusion.

APPENDIX 2

Scoring of Potential Splicing Mutations

TABLE I. Scoring of Purified Mutants

In this table the scoring of the potential splicing mutants described in the text is shown. Growth on -Ura and Hol media was scored by replica-plating patches from YEPD master plates and scoring visually for growth; -Ura plates were scored after 24 hours, Hol plates were scored after 48 hours (all manipulations were performed at 30 C). Complementation testing was performed by cross-stamping the master plates containing the mutants with the tester strains, incubating 24 hours at 30C to allow mating, and replica-plating to -His media. The cross-stamped region was then scored for growth at 24 hours. In addition to the testers shown, each mutants was shown to be competent to mate by cross-stamping with an additional mating type tester. Two different HIS4A- testers were used; FCT-3 gave a consistently higher level of complementation. The addition of the scoring of additional 19 mutants from a second mutagenesis gives the numbers in Table 2, Chapter 1. None of these additional mutants passed all the genetic criteria and were therefore never characterized further.

TABLE1

Mutant	-UKA	Hol	xFCT-3 (A-)	xFCT-4 (A-)	xFCT-6 (B-)	xFCT-8 (C-)
B1	+	-	-	-	-	-
B2	+	-/+	+/-	-	-	-/+
B3	+	-	-	-	-	-
B4	+	-	-	-	-	-
B6	+	+/-	++/-	+/-	+/-	+/-
B7	+	-	-	-	-	-
B8	+	-	-	-	-	-
B9	+	+/-	+/-	-/+	-/+	+/-
B10	+	-	-	-	-	-
B11	+	++/-	+/-	-	-	++/-
B13	+	++/-	++/-	+/-	+/-	++/-
B14	+	-	--/+	-	-	-
B15	+	-	-	-	-	-
B16	+	+/-	++/-	+/-	+/-	+/-
B18	+	-	++/-	++/-	+/-	-
B19	+	+/-	-/+	-/+	-/+	-/+
B20	+	-	++/-	+/-	-	-
B21	+	-	+/-	-/+	-	-
B23	+	++/-	--/+	-	-	+
B24	+	-	-	-	-	-
B26	+	-/+	--/+	-	-	-/+
B27	+	-	-	-	-	-
B28	+	-	-	-	-	-
B29	+	-	-	-	-	-
B30	+	-	-	-	-	-
B33	+	-	-	-	-	-
B34	+	-	-	-	-	-
B35	+	-	-	-	-	-
B36	+	++/-	++/-	+/-	+/-	+/-
B38	+	-/+	+/-	-/+	-/+	-/+
B39	+	+/-	-/+	--/+	-	--/+
B40	+	-	-	-	-	-
B41	+	-	-	-	-	-
B42	+	-	-	-	-	-
B43	+	-	-/+	-	-	-
B44	+	+/-	++/-	+/-	-/+	+/-
B45	+	-/+	+/-	-/+	--/+	+/-
B46	+	-	-	-	-	-
WT	+	+	+	+	+	+
UT	-	-	-	-	-	-

TABLE II. Papillation of Mutants and Suppression by SUP4o

This table shows the results of the additional genetic screens described briefly in the text. The number of papilli for each mutant obtained with and without UV irradiation are shown. Mutants scored with a "bg" indicates that the background level of growth was to high to allow the counting of papilli. Irradiation of the Hol plates for 20" at a distance of 12 inches with UV light after replica-plating clearly increased the frequency of papillation for almost all of the mutants.

Mutants in which the A and B activities were found to co-revert with the HIS4C activity are marked with a "*", these constitute the Class II group of mutants. Mutants which were unable to be scored and constitute the Class I mutants are marked with a "$" symbol. Mutants which are considered to suppressed by SUP4 are marked with "!".

TABLE II

Mutant	Papilli no UV	Papilli 20" UV	xFCO-6 (SUP4, his4)
B1*	0	60	-
B2$	hg	hg	-
B3*	0	9	-
B4	0	0	-
B6*	0	17	-
B7	0	0	-
B8	0	0	-
B9$	hg	hg	-
B10	0	0	-
B11$	hg	hg	-
B13$	hg	hg	-/+
B14*	1	16	-
B15	5	25	-
B16$	hg	hg	-
B18	0	0	-
B19$	hg	hg	-/+
B20*	2	16	-
B21	0	0	-
B23$	hg	hg	++/-
B24	0	0	-
B26$	hg	hg	++/-
B27	0	0	-
B28*	0	3	-
B29	0	0	-
B30*	0	6	-
B33	8	45	-
B34	0	0	-
B35	0	1	-
B36$	hg	hg	++/-
B38$	hg	hg	+/-
B39$	hg	hg	+/-
B40!	0	0	+
B41	0	0	-
B42	0	0	-
B43*	9	8	-
B44$	hg	hg	-/+
B45*	16	45	-
B46!	0	0	+

CHAPTER TWO

A Point Mutation in the Conserved Hexanucleotide at the 5' Splice Junction Uncouples Recognition, Cleavage and Ligation.

ABSTRACT

We have constructed a translational fusion between the yeast actin gene and the HIS4 gene such that expression of the HIS4 gene product requires proper splicing of the actin intron. Using this chimeric gene, we have isolated a point mutation which alters splicing of the fusion transcript. The mutation is a G to A transition in the fifth position of the highly conserved sequence /GTAPyGT, found at the 5' junction of all yeast nuclear mRNA introns. The biochemical consequences of this mutation are threefold: 1) a decreased amount of correctly spliced mRNA is produced; 2) full-length precursor accumulates; and 3) a second intron-containing species accumulates, the 5' terminus of which is six bases 5' to the normal 5' junction. Surprisingly, this novel terminus arises via cleavage at a sequence that bears no resemblance to a consensus 5' junction. Moreover, cleavage at this abnormal site does not lead to the production of mature mRNA, although this aberrant intermediate appears to be in a lariat structure. The behavior of this mutant argues that recognition of the 5' junction and subsequent cleavage are separable events and, furthermore, that requirements for 3' endonucleolytic cleavage may be more complex than previously imagined.

INTRODUCTION

Many eukaryotic genes are not colinear with the mature, functional products they encode (for review, see Abelson, 1979). In the yeast Saccharomyces cerevisiae, a subset of nuclear genes encoding mRNAs contain intervening sequences (Teem et al., 1984). The powerful genetic techniques available in yeast can be readily exploited for the generation of mutations affecting splicing. The nature and location of cis-acting mutations, together with information derived from the characterization of the resultant processing intermediates, are expected to suggest plausible molecular models for substrate recognition. Moreover, by describing how the machinery itself operates, such an approach should also provide important clues as to the nature of the splicing components.

Analysis of intron sequences by several methods has revealed two regions that are critical for splicing. In yeast, the conserved sequence TACTAAC, located from 4 to 53 bases upstream of the 3' junction, has been shown by deletion analysis to be required for splicing (Langford and Gallwitz, 1983; Pikielny et al., 1983). A functional analog of this so-called TACTAAC box has recently been shown to exist in mammalian cells (Ruskin et al., 1984; Padgett et al., 1984). A detailed analysis of transcripts in yeast, both in vivo (Domdey et al.,1984; Rodriguez et al., 1984) and in vitro (Newman et al., 1985), and in mammalian extracts (Ruskin et al., 1984; Grabowski et al., 1984), has shown that

splicing of pre-mRNA occurs via an intermediate with a lariat structure, in which the 5' end of the intron is joined by a 2'-5' phosphodiester bond to a site within the intron. The location of the 2'-5' linkage, or branch (cf. Wallace and Edmonds, 1983), is within the TACTAAC box in yeast (Domdey et al., 1984). In mammalian cells, branch points seem to fall in comparable locations (with respect to the 3' junction), but an obvious consensus sequence is less readily identifiable (Padgett et al., 1984; Ruskin et al., 1984; Keller and Noon, 1984). This striking structural resemblance between the intermediates argues that splicing must proceed by fundamentally similar mechanisms in yeast and mammals, despite the stricter adherance to a specified primary sequence at the TACTAAC box in yeast.

A stronger conservation at the level of primary sequence in yeast splicing signals is also seen at the 5' intron/exon junction. The hexanucleotide /GTAPyGT is essentially 100% conserved (Langford and Gallwitz, 1983; Teem et al., 1984). By comparison, the 5' splice junction in metazoans is characterized by the substantial degree of variation from the nine base consensus, $^{C}_{A}$AG/GTPuAGT (Mount, 1982). This apparent flexibility in primary sequence requirements is reflected in the behavior of mutations which inactivate the 5' junction. Almost invariably, one or more cryptic junctions is activated; that is, sequences which resemble the 5' consensus sequence are utilized instead of the mutated junction in the splicing process (for review, see Mount and Steitz, 1983). In contrast, deletions which inactivate the yeast 5' consensus do <u>not</u> appear to activate cryptic junctions, but lead instead

to the accumulation of full-length precursor (Gallwitz, 1982; Pikielny et al., 1983).

The failure of the yeast splicing machinery to use alternative 5' splice sites should prove to be useful, by revealing the specific molecular consequences of the attempt to use a mutant junction. Given this, a genetic approach to the splicing problem should be particularly informative. To this end, we have constructed a gene fusion such that production of an easily scored gene product is dependent on proper splicing. Following random in vitro mutagenesis, we screened in vivo for point mutations which alter the splicing of this gene fusion. The pleiotropic effects of one such mutation, a G to A transition at the fifth position of the 5' consensus sequence, argues that initial recognition of the 5' junction and subsequent cleavage at that site are separable events and, furthermore, that requirements for 3' endonucleolytic cleavage may be more complex then previously imagined.

RESULTS

CONSTRUCTION OF FUSION

To facilitate identification of mutations affecting mRNA splicing, we desired a spliced gene with a readily scored biological phenotype. The set of known spliced yeast genes consists predominantly of essential genes, i.e. genes for actin and many ribosomal proteins (Teem et al., 1984), which are difficult to experimentally manipulate. We thus turned to the powerful technique of gene fusion and created a chimera between the yeast actin gene, which contains a 309 nucleotide intervening sequence (Ng and Abelson, 1980; Gallwitz and Sures, 1980), and the yeast HIS4 gene (see Figure 1). Precise and efficient splicing of the actin intron is required for production of functional HIS4 gene product. Failure to splice should prevent translation of the HIS4 coding sequence because of numerous stop codons within the intron. Inaccurate splicing should, in most cases, prevent functional expression by altering the translational frame. Finally, efficient splicing is required because, for fortuitous reasons, a reduction of 50% in the level of fusion gene product results in a mutant phenotype. HIS4 encodes a polyfunctional protein having three separate enzymatic activities, HIS4 A, B, and C (Keesey et al., 1979). Because the HIS4 coding sequence is downstream of the actin intron in the gene fusion, failure to splice should lead to the loss of expression of all three HIS4 activities.

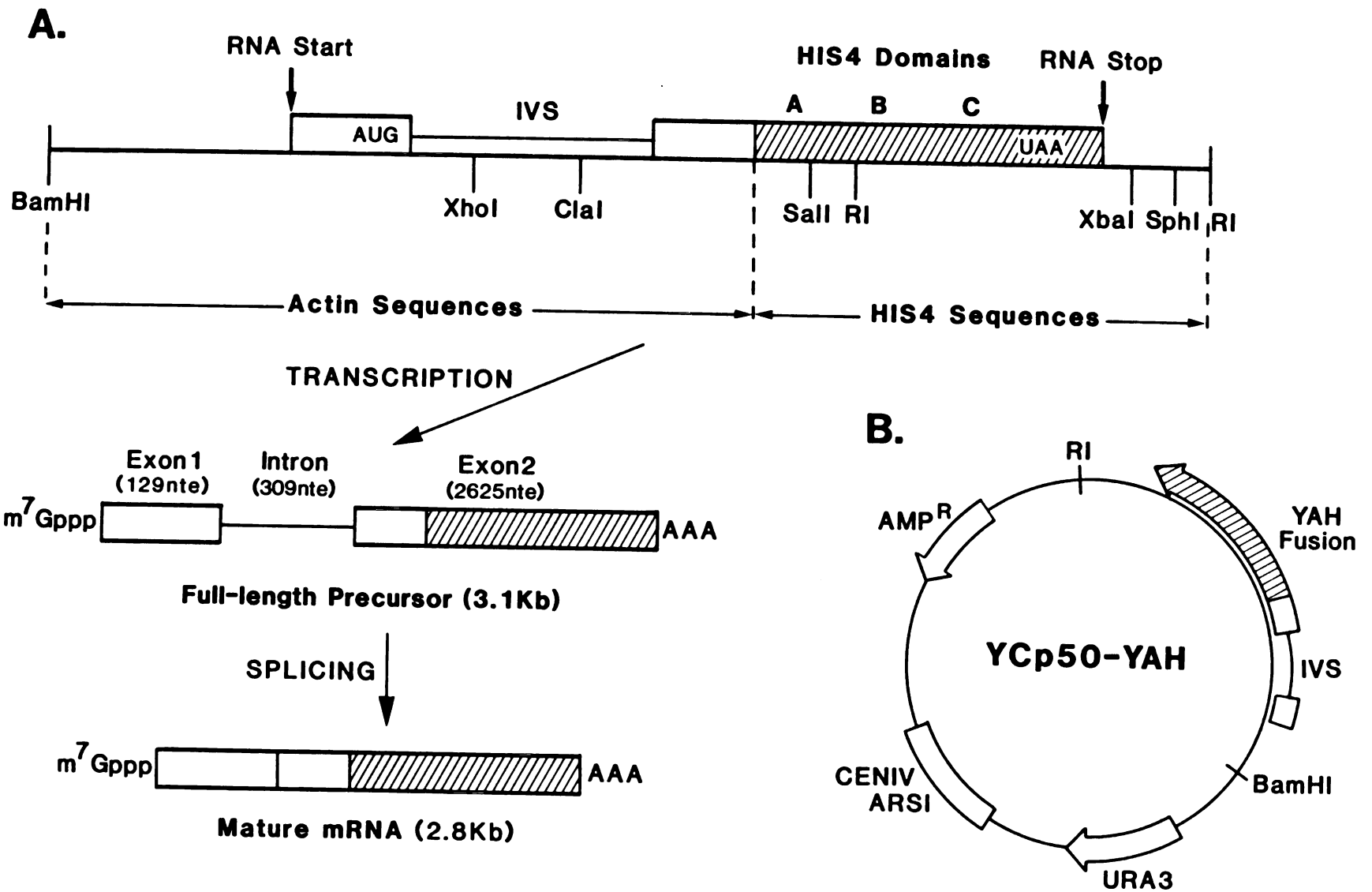

A.
RNA Start
HIS4 Domains
RNA Stop
IVS
A
B
C
AUG
UAA
BamHI
XhoI
ClaI
SalI
RI
XbaI
SphI
RI
Actin Sequences
HIS4 Sequences
TRANSCRIPTION
Exon1
(129nte)
Intron
(309nte)
Exon2
(2625nte)
m7Gppp
AAA
Full-length Precursor (3.1Kb)
SPLICING
m7Gppp
AAA
Mature mRNA (2.8Kb)
B.
RI
AMPR
YAH
Fusion
YCp50-YAH
IVS
CENIV
ARSI
BamHI
URA3

Figure 1 Construction of Actin-HIS4 Fusion

A) Schematic diagram of the YAH fusion (not to scale); for details of construction, see Experimental Procedures. Transcribed regions are shown as boxes; open boxes correspond to actin sequences, slashed boxes to HIS4 sequences. The fusion produces a 3.1 kb precursor which is processed to a 2.8 kb mature mRNA (see Fig. 3). In addition to the coding sequences, the fusion has 364 bases of 5' flanking sequence and 904 bases of 3' flanking sequence. The fusion presumably initiates translation at the AUG normally used in the actin gene, 10 bases 5' to the 5' intron/exon junction, and terminates at the HIS4 terminator codon (Ng and Abelson, 1980; Gallwitz and Sures, 1980; Donahue et al., 1982). The relative domains of the different HIS4 activities (Donahue et al., 1982) are shown above the fusion. Relevent restriction sites are shown.

B) Map of YCp50-YAH, a centromere plasmid containing the YAH fusion, which is maintained as a freely replicating single copy plasmid. Transformation of a HIS4 deletion strain with this plasmid allows the production of functional HIS4 A,B, and C activities (see Figure 2).

The chimeric gene was constructed by the addition of a BamHI linker to the XhoI site in the HIS4 coding sequence, and subsequent ligation to the BglII site in the actin coding sequence, 241 base pairs from the 3' end of the intervening sequence (Figure 1A). The resulting translational fusion retains 364 basepairs of the 5' flanking sequence, the RNA start site, and the translational start of the actin gene. Except for the N- terminal 30 amino acids, the fusion contains the entire HIS4 coding sequence and the HIS4 transcriptional terminator. We refer to this construct as the YAH (yeast-actin HIS4) fusion.

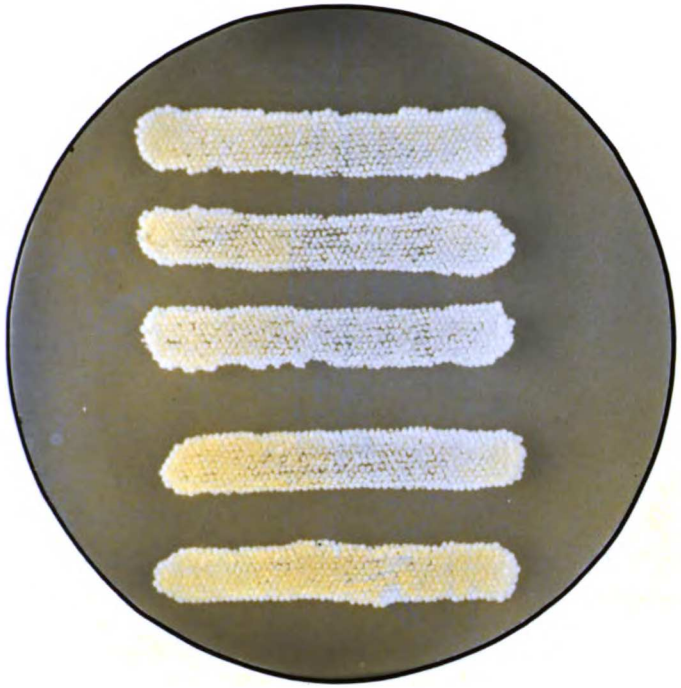
The YAH fusion was inserted into a plasmid carrying a yeast centromere and the URA3 gene, and the resultant product, YCp50-YAH (Figure 1B), was used to transform a ura3-52 strain deleted for HIS4 (his4-401); Ura$^+$ transformants were selected. The presence of the YAH fusion now enables the cell to complement HIS4 A-, B- and C- mutants (data not shown), demonstrating that the fusion produces functional HIS4 A, B, and C activities. While we test all three activities independently by performing complementation tests, we can also assay directly for functional HIS4 C gene product. The enzyme encoded by the HIS4 C coding domain is histidinol dehydrogenase (Kessey et al., 1979), which catalyzes the last step in the histidine biosynthetic pathway. By growing cells on media lacking histidine but supplemented with histidinol (Hol), we can thus test directly for functional HIS4 C activity (without requiring HIS4 A or B activities). As shown in Figure 2 (compare WT and untransformed, UT), the YAH fusion allows the strain carrying the his4-401 deletion to grow on histidinol.

Figure 2 Biological Phenotypes of Fusion-Containing Strains

Transformants of FC2-12B with different forms of YCp50-YAH were grown on rich media (YEPD) and replica-plated to complete synthetic media (Comp) and to Hol media (described in text). Plates were photographed after 3 days at 30°C to exaggerate the leakiness of the phenotype; growth of the A5 mutant is undetectable after 1 day. Abreviations are: UT= untransformed strain, FC2-12B; WT= wild type fusion; A5= A5 mutant fusion; XS1 and XS2 are explained in legend to Figure 4.

Comp

Hol



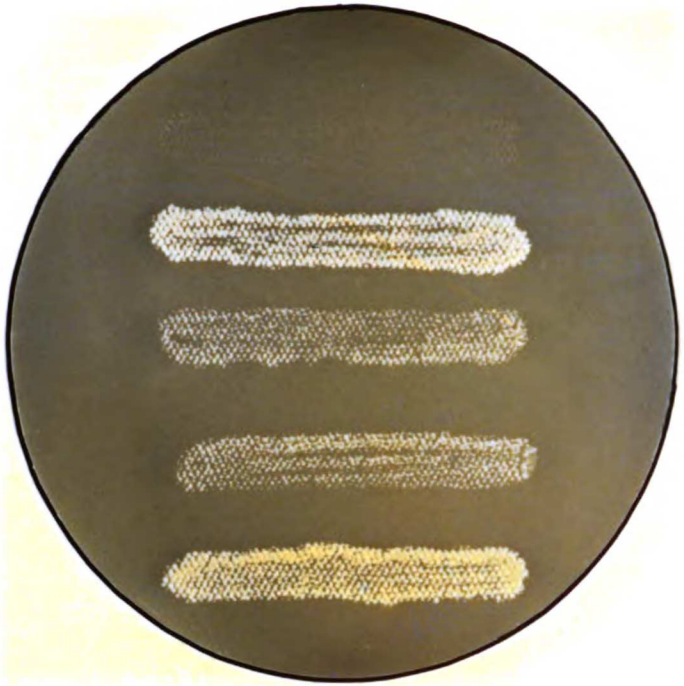
UT

WT

A5

XS1

XS2



Northern analysis of transcripts from strains carrying this fusion, either on the centromere plasmid YCp50-YAH (Figure 1B), or integrated at the actin locus (see Experimental Procedures), reveals a 2.8 kb RNA which hybridizes to both HIS4- and actin-specific probes (Figure 3A, lane 1, and data not shown). This is as expected if the actin RNA start site and the HIS4 terminator are used. To demonstrate that splicing is required for generation of the YAH fusion product, we performed a similar analysis on RNA derived from a strain which also carries the rna2 lesion. Rna2 is a temperature-sensitive lethal mutation which at high temperature causes accumulation of full-length precursors from spliced genes (Rosbash et al., 1981). As shown in Figure 3A (lane 2), at high temperatures in rna2 strains we detect a 3.1 kb RNA, as expected for failure of the fusion to be efficiently spliced.

ISOLATION OF A SPLICING MUTANT

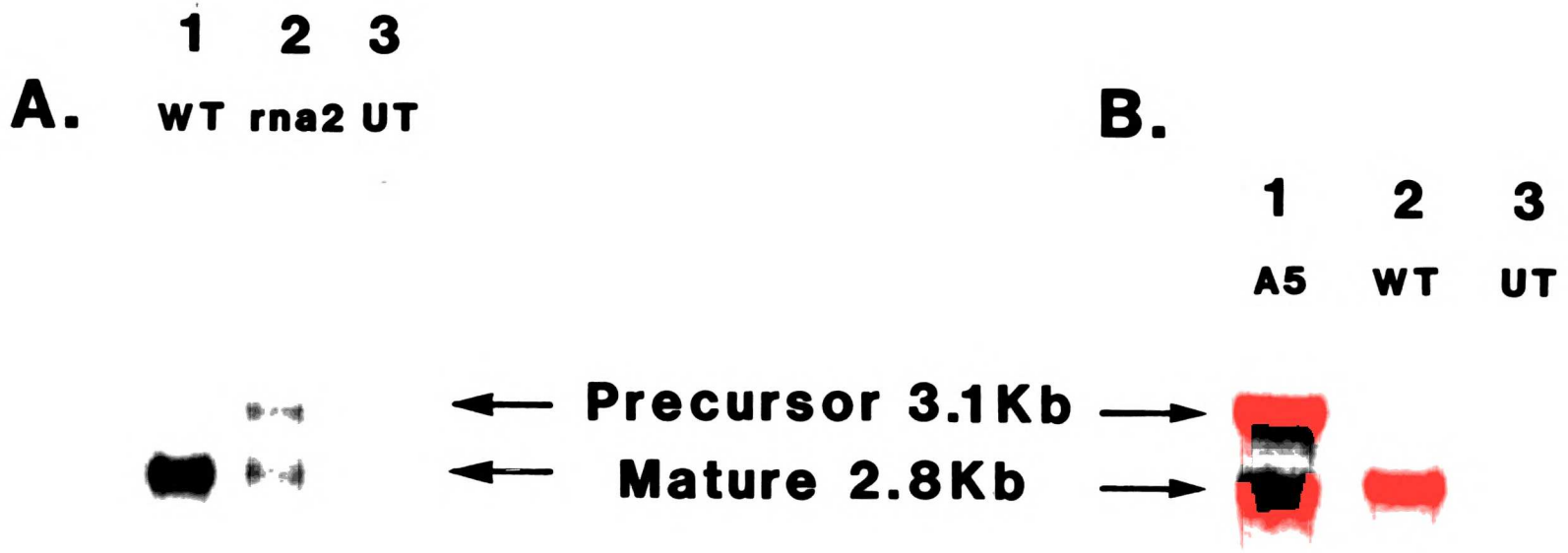
Now that we had a distinct phenotype associated with the removal of a yeast intron, our goal was to isolate cis-acting mutations affecting splicing in an unbiased manner. To introduce random transitions, we mutagenized YCp50-YAH DNA with hydroxylamine in vitro (see Experimental Procedures). The DNA was then used to transform FC2-12B, a strain which is deleted for HIS4 and is ura3-. Ura$^+$ transformants were screened for loss of ability to grow on Hol (i.e., loss of HIS4 C function) as an initial test for splicing mutants.

A set of additional genetic screens were then employed for processing the numerous Hol- mutants obtained by this procedure, to

Figure 3 Northern Analysis of Fusion-Containing Strains

A) Northern gel of total RNA from strains with pYAH-I1 (see Experimental Procedures) integrated at the actin locus. Both Northerns (A and B) were probed with nick translated YIp25 (a HIS4-containing plasmid); lane 1, wild-type fusion in TD77; lane 2, wild-type fusion in FC3-3C, an rna2- strain, RNA prepared after 45'at 36 C; lane 3, untransformed TD77. For both Northern gels, sizes were estimated by comparison with HindIII-cleaved Lambda DNA, and with 18S RNA and 28S RNA visualized by staining with ethidium bromide.

B) Northern gel of total RNA from FC2-12B tranformed with YCp50-YAH. Lane 1; A5 mutant fusion; lane 2, wild-type fusion; lane 3, untransformed FC2-12B. The amount of precursor varies slightly among RNA preparations but is in the range of 50-70% of the total fusion transcripts (estimated by gel tracings). The amount of mature mRNA produced in the A5 mutant and also varies and is between 30-50% of wild-type levels.


A.
1 2 3
WT rna2 UT
B.
1 2 3
A5 WT UT
Precursor 3.1Kb
Mature 2.8Kb

identify those which might be affecting splicing. The criteria for a potential splicing mutant were that: 1) the mutation be polar (i.e., HIS4 A-B-C-); 2) the mutation map to the actin sequences in the fusion; 3) the loss of all three activities be caused by a single event (i.e., be co-revertible); and 4) the mutation not be suppressible by nonsense suppressors. Mutants which passed each of these genetic tests (described in detail in the Experimental Procedures) were subjected to Northern analysis to determine if there was a change in the ratio of precursor to mature mRNA.

Total RNA was prepared, electrophoresed on formaldehyde agarose gels (1.5%), blotted, and probed with DNA prepared by nick-translating a plasmid carrying HIS4 (YIp25; Donahue et al., 1982). One mutant from this screen exhibited a dramatic accumulation of an RNA species with the mobility (3.1 kb) of full-length precursor (Figure 3B, lane 1). Strains carrying this mutation, designated A5, continue to produce mature RNA, but at a noticeably reduced level; by densitometric scanning we estimate this to be 30-50% the amount of mRNA in the wild-type fusion. Consistent with the decreased production of mature mRNA, this mutation confers a slightly "leaky" biological phenotype (Figure 2; compare WT and A5).

SEQUENCE ANALYSIS REVEALS A SINGLE BASE CHANGE IN A HIGHLY CONSERVED HEXANUCLEOTIDE

Determination of the nucleotide change was accomplished by subcloning restriction fragments which span the intron (see

Experimental Procedures) into the M13 sequencing vector MP9, and sequencing by the dideoxy method (Sanger et al., 1977). Sequence analysis of the 5' exon and the 5' portion of the intron revealed a single G to A transition at the fifth nucleotide of the intron (Figure 4A). As described in the legend to Figure 4, we have subcloned a restriction fragment containing this alteration into an unmutated fusion. A strain carrying this construction has the same biological (see Figure 2) and biochemical (primer extension analyses not shown) phenotypes as do strains containing the parental mutant fusion. We can thus conclude that this single base change is sufficient to generate the observed splicing defect.

This mutation is in the hexanucleotide /GTAPyGT located at the 5' junction of all yeast introns. The observation that a single base change within this consensus causes accumulation of mRNA precursor strongly argues that this highly conserved sequence does indeed play an important role in splicing.

TRANSCRIPT ANALYSIS REVEALS AN ABERRANT INTERMEDIATE

In order to rigorously determine the number and types of transcripts produced, we analyzed the RNA from strains carrying the wild-type and A5 mutant fusions by a series of S1 nuclease protection and primer extension experiments.

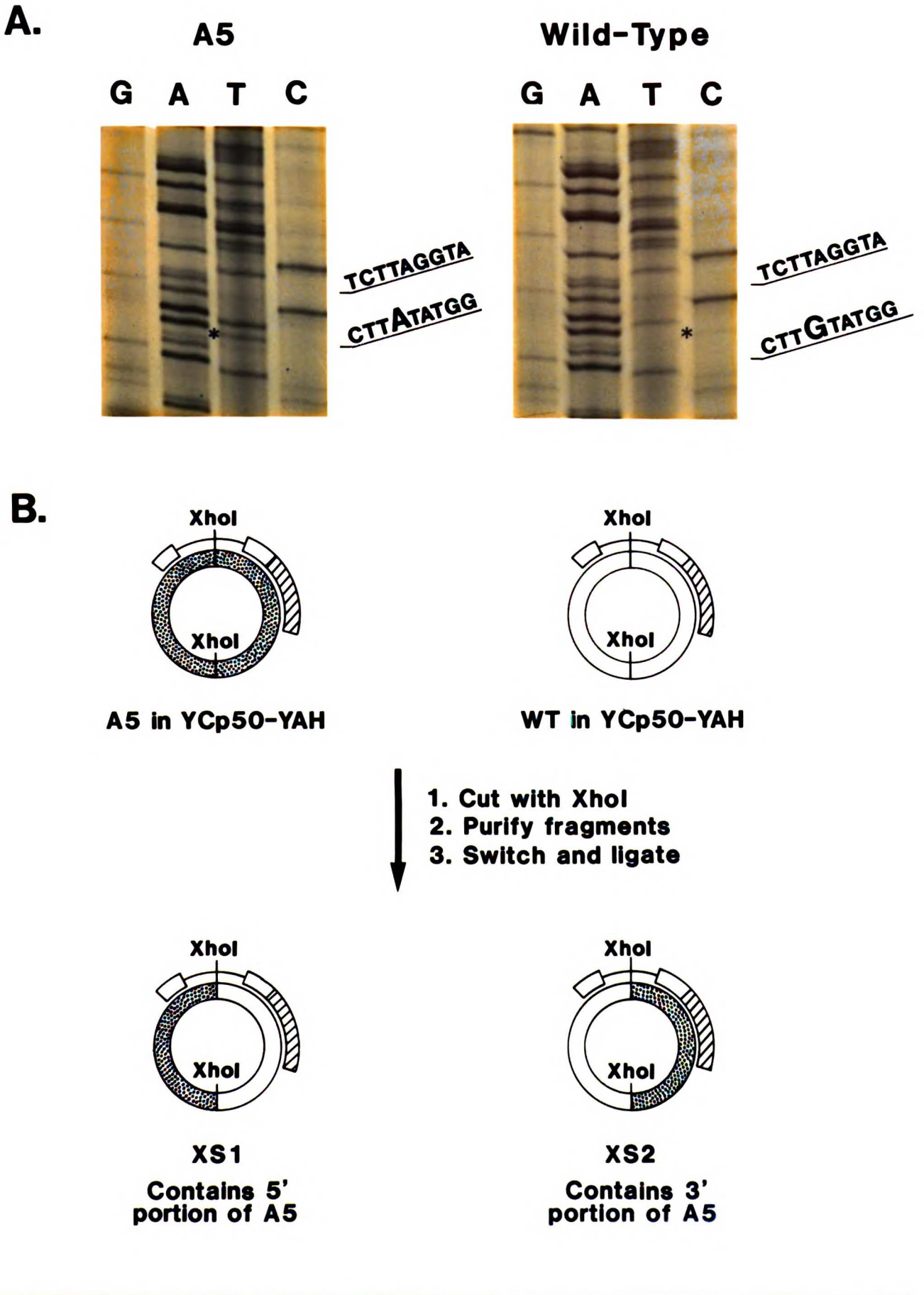

A.
A5
Wild-Type
G A T C
G A T C
TCTTAGGTA
CTTATATGG
TCTTAGGTA
CTTGTATGG
B.
XhoI
XhoI
XhoI
XhoI
A5 in YCp50-YAH
WT in YCp50-YAH
1. Cut with XhoI
2. Purify fragments
3. Switch and ligate
XhoI
XhoI
XhoI
XhoI
XS1
Contains 5' portion of A5
XS2
Contains 3' portion of A5

Figure 4 Sequence Analysis of the A5 Mutation

A) The sequence from MP9 subclones of the XhoI-SalI fragment, which encompasses the 5' portion of the intron and the 5' exon sequences, is shown. The two sequences show the wild-type and A5 mutant sequences in the region of the transition. The letters adjacent to the gel are the same sequence as the RNA and are complementary to the sequence ladder. The mutation is shown in the larger letter.

B) Wild-type and A5 YCp50-YAH plasmids were cleaved with XhoI and the resulting fragments purified. The pieces were then exchanged and ligated to yield XS1, having the 5' portion of the fusion from the A5 mutant, and XS2, having the 3' region of the fusion from the A5 mutant. These plasmids were transformed into FC2-12B and tested for biological (Figure 2) and biochemical phenotypes (data not shown). XS1 behaves identically to the original A5 mutant. In the diagram, speckled plasmid regions refer to sequences derived from the original A5 isolate, open plasmid regions correspond to sequences from the unmutated fusion. Though we have not sequenced the entire 4.6 kb XhoI fragment we have sequenced the entire portion derived from actin sequences (intron, 5' exon and flanking sequences) and find only the single base change corresponding to the A5 mutation.

S1 Analysis From 3' Exon Reveals Two Intron-Containing Species

Our first goal was to confirm that the bands revealed by Northern analysis corresponded to mature and full-length precursor. To do this we performed an S1 nuclease protection experiment using a probe labeled at the SalI site within the HIS4 coding sequence, 625 nucleotides from the 3' end of the intron (Figure 5). This probe has the advantage of detecting any species containing the second exon, i.e. both mature and precursors RNAs.

Total RNA was prepared from strains carrying the mutant and wild-type fusions, annealed to the end-labeled fragment, and treated with S1 nuclease at various concentrations. The protected species were then analyzed on a sequencing gel (Figure 5). In strains carrying the wild-type fusion, the major protected species has a length of 625 bases (lane 1, Band 1), the correct size for mature RNA. In strains carrying the A5 mutant fusion, three major protected species can be seen (lane 2): a band of 625 bases (Band 1), corresponding to mature mRNA; a band of approximately 1063 bases (Band 3), the length expected for full-length precursor; and a new species with a 5' terminus approximately 934 nucleotides from the position of the label (Band 2), i.e. at or near the 5' intron junction. We do not detect this species in untransformed cells (data not shown).

We conclude from this experiment that the A5 mutation causes accumulation of two intron-containing species. One has a 5' terminus at the transcriptional start site and corresponds to full-length precursor. The 5' terminus of the second species is at or near the 5' splice

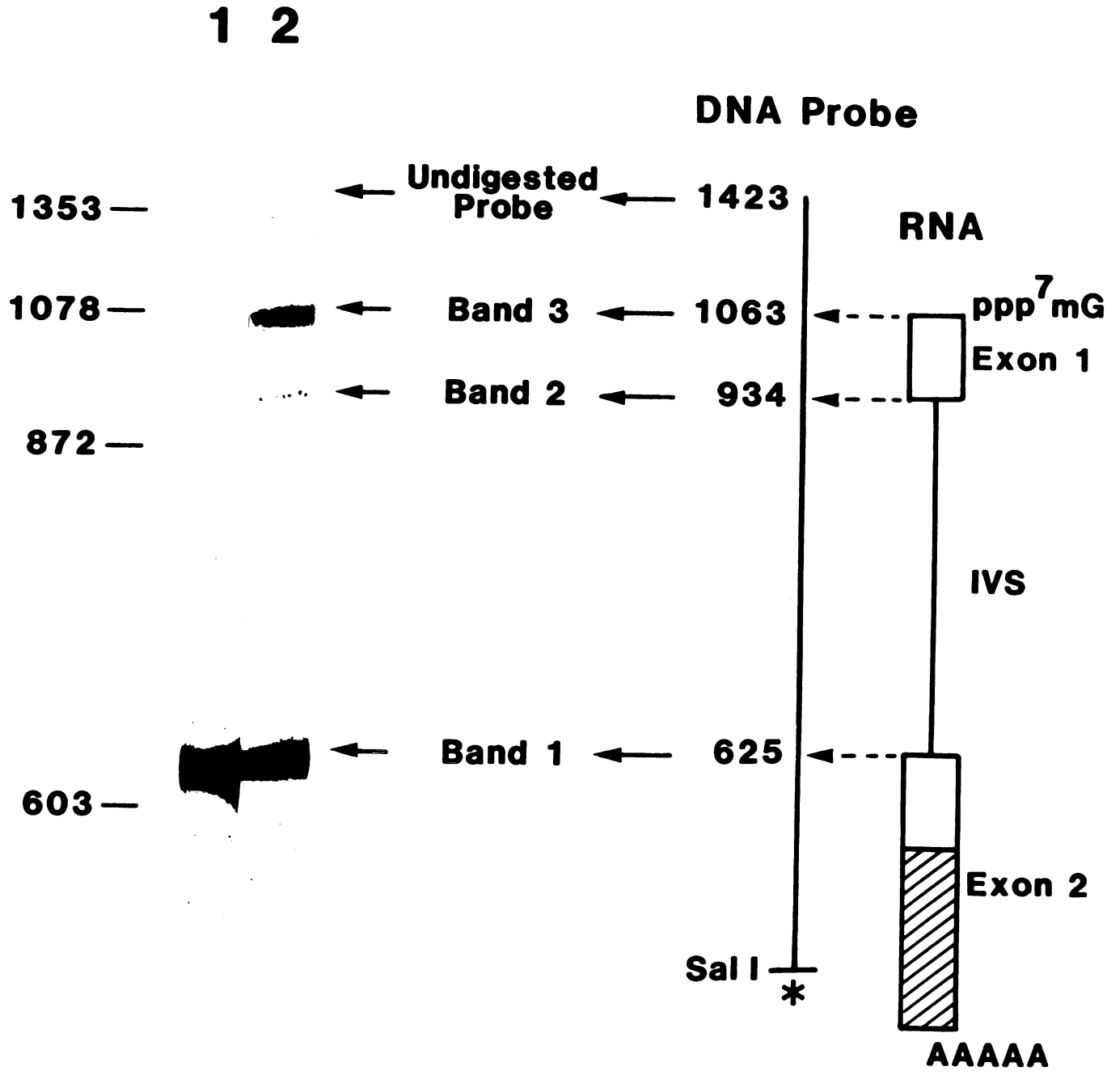

1 2
DNA Probe
Undigested Probe
1423
1353
1078
872
603
Band 3
Band 2
Band 1
1063
934
625
RNA
ppp7mG
Exon 1
IVS
Exon 2
Sal I
*
AAAAA

Figure 5 S1 Nuclease Protection from a Probe in the 3' Exon

Total RNA from FC2-12B, transformed either with the wild-type or the A5 mutant YCp50-YAH plasmids was subjected to S1 analysis as described in the Experimental Procedures and electrophorsed on a 6% sequencing gel. The probe for this experiment is an end-labeled double-stranded restriction fragment extending from the SalI site in the HIS4 sequences to the BamHI site beyond the transcriptional start. Sizes for endpoints are shown schematically. Note the reduction of mature mRNA in the A5 mutant (each lane corresponds to 50 micrograms of RNA, by O.D.). Lane 1, wild-type (900 units S1); lane 2, A5 mutant (900 units S1).

junction, and must contain at least part of the 3' exon. In that the resolution of this experiment was not sufficient to map the terminus of this new RNA species with precision, a second S1 experiment was performed.

The New 5' Terminus Is Near But Not At The 5' Junction

To map the 5' terminus more precisely, a probe was labeled at the XhoI site within the intron, 66 nucleotides from the 5' junction (see Figure 6). Thus only precursor species will be detected in this experiment. Again, total RNA was prepared from strains carrying the mutant and wild-type fusions, annealed to the probe, treated with S1 nuclease, and analyzed on gels.

In strains carrying the wild-type fusion, only a very low level of a species of 195 nucleotides was protected (Figure 6, lanes 4-6; Band 1). This is the correct size for full-length precursor. In strains carrying the A5 mutation we again detected <u>two</u> RNA species (lanes 1-3): one corresponding to full-length precursor (Band 1), and a second species with a length of ca. 72 nucleotides (Band 2). This places the 5' terminus approximately six bases upstream of the normal 5' intron/exon junction. This terminus is not seen in wild-type cells, and is, in any case, located at an unexpected position.

We conclude that the second intron-containing species in the A5 mutant has a 5' terminus <u>near</u> but not <u>at</u> the 5' intron/exon junction. Moreover, it should be noted that no transcripts containing a 5'

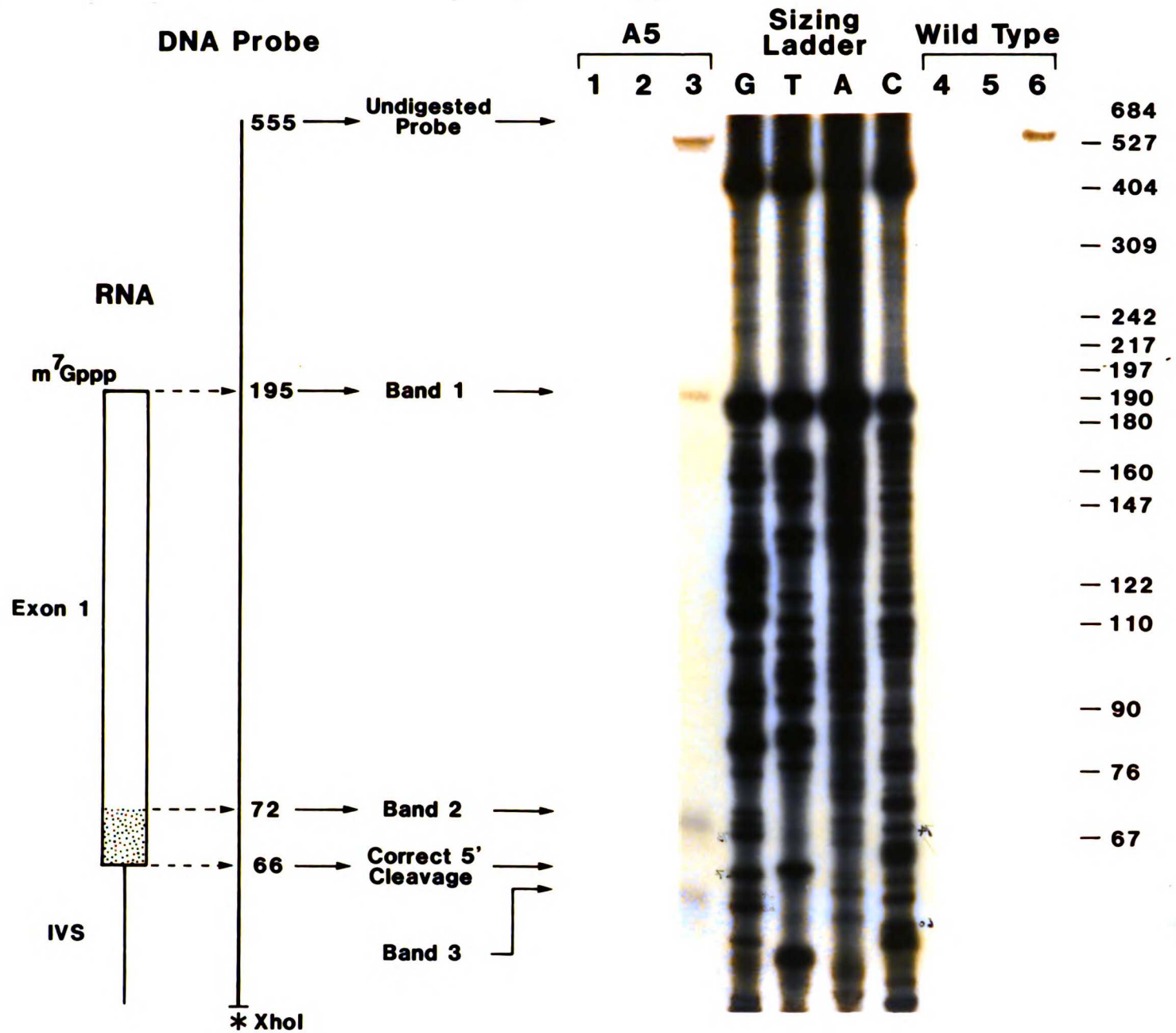

DNA Probe
RNA
m7Gppp
Exon 1
IVS
555
195
72
66
Undigested Probe
Band 1
Band 2
Correct 5' Cleavage
Band 3
*XhoI
A5
Sizing Ladder
Wild Type
1 2 3 G T A C 4 5 6
684
527
404
309
242
217
197
190
180
160
147
122
110
90
76
67

Figure 6 S1 Nuclease Protection from a Probe Within the Intron

Total RNA from FC2-12B transformed either with the wild-type or A5 mutant fusion in YCp50-YAH was treated as described in Experimental Procedures and electrophoresed on an 8% sequencing gel. The probe for this experiment is the double- stranded restriction fragment kinased at the XhoI site and extending to the BamHI site 5' of the transcriptional start. The speckled region in the diagram denotes the bases between the novel terminus and the 5' intron/exon junction. Lane 1, A5 mutant (36 units S1); lane 2, A5 mutant (180 units S1); lane 3, A5 mutant (900 units S1); lane 4, wild-type (900 units S1); lane 5, wild-type (180 units S1); lane 6, wild-type (36 units S1). The sequencing ladder is the sequence of the Herpes TK gene sequenced from the primer 5' GGAGTGCGGGAGTTTCACGCCACC 3' (provided by B. Maler). Band 3 is present in both wild-type and mutant cells; the endpoint is a few bases within the intron and is extremely S1 sensitive. We do not detect this species by any other technique and do not understand its origin.

terminus at the correct exon-intron junction could be detected in this experiment.

PRIMER EXTENSION CONFIRMS S1 ANALYSIS

Because of the surprising location of the novel terminus, we sought to confirm our results using a second technique, primer extension. For these experiments, we integrated the mutant and wild-type fusions at the HIS4 locus (see Experimental Procedures) in a strain (FC8-24D) in which the intron of the chromosomal copy of the actin gene had been precisely deleted (Ng et al., 1984). This allows us to avoid any background due to the hybridization of intron-specific probes to transcripts derived from the chromosomal copy of the actin gene. In addition, we prepared RNA from a strain transformed with a plasmid designed to overproduce actin-derived splicing intermediates (Domdey et al., 1984). This enables us to directly compare the position of the correct 5'cleavage with the postion of this cut in the A5 mutant.

A kinased 22-nucleotide primer was annealed and extended from its 3' end, 24 nucleotides downstream of the 5' intron/exon junction. As shown in Figure 7, comparison of the cDNA products from the two strains clearly demonstrates that this novel terminus is not at the 5' junction (compare lanes 2 and 3). Moreover, by comparison with the ladder generated by dideoxy sequencing of an intron-containing subclone in MP9 using this same primer, we can place this novel terminus at precisely six bases 5' to the correct intron/exon junction.

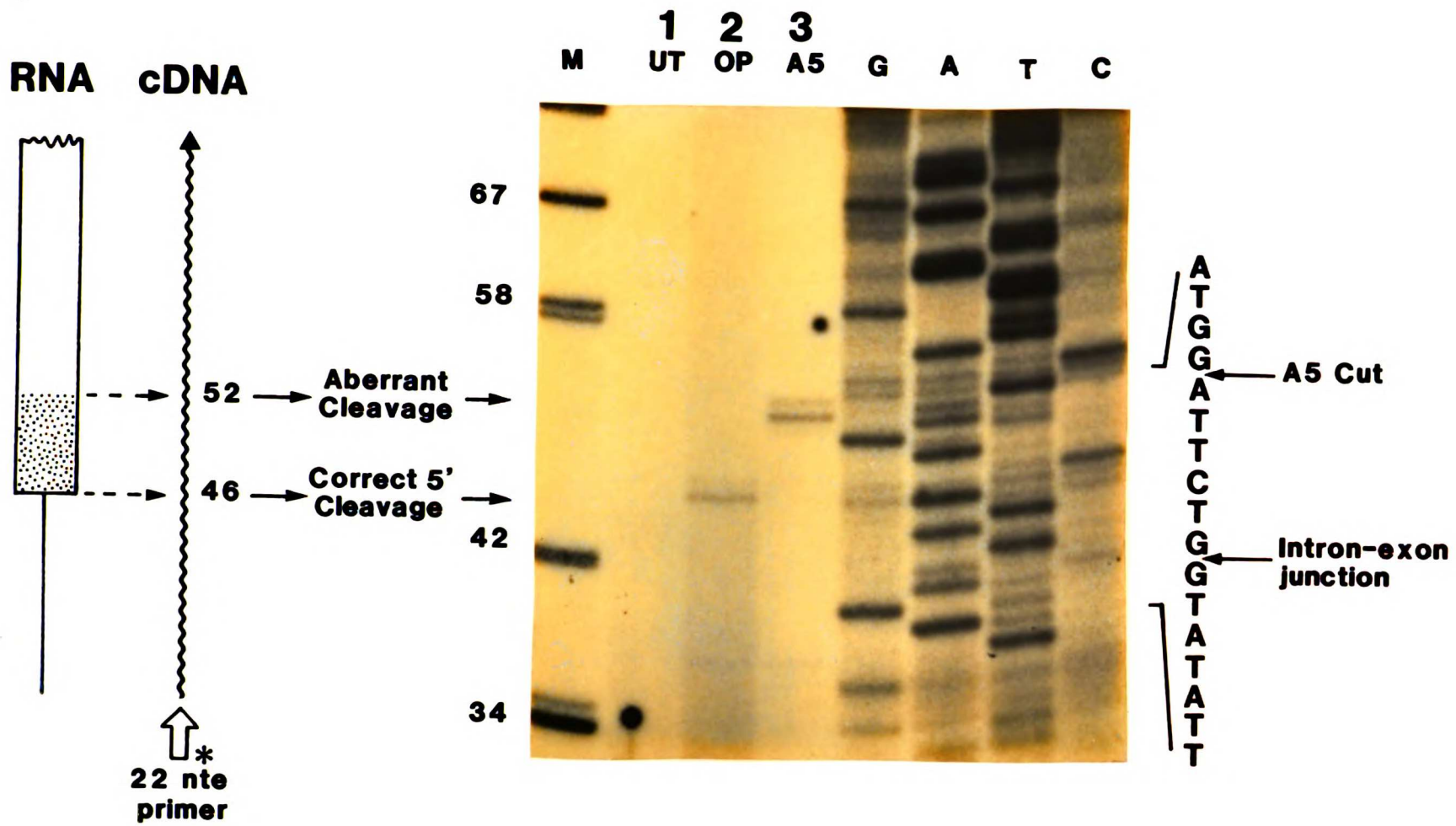

RNA
cDNA
52
Aberrant Cleavage
46
Correct 5' Cleavage
22 nte primer
M
1 UT
2 OP
3 A5
G
A
T
C
67
58
42
34
A5 Cut
Intron-exon junction
ATGGATTCTGGTATATT

Figure 7 Determination of the Position of the Novel Terminus

RNA was prepared from the strain FC8-24D, with the A5 mutant fusion integrated at HIS4, and from FC8-24D transformed with the plasmid pBA3, which by overproducing the actin transcript allows the detection of intermediates in the splicing process (see text). RNA was extended from the oligonucleotide primer, 5'AATTCTTCTTACAGTTAAATGG 3', with reverse transcriptase and the cDNA products run on a 10% sequencing gel. The same primer was used to sequence an MP9 subclone containing the 5' portion of the A5 mutant fusion; the lanes labeled G,A,T, and C are sequencing products with the appropriate dideoxy chain terminator. The letters adjacent to the gel are the same sequence as the RNA, and are complementary to the sequence ladder. Because the sequencing ladder was generated with an unkinased primer the mobilities of the corresponding cDNA products will differ by half a nucleotide (Sollner-Webb and Reeder, 1979). Lane 1, untransformed strain FC8-24D; lane 2, FC8-24D transformed with pBA3; lane 3, A5 mutant. Note that in the A5 mutant we detect this terminus as a doublet, with the longest extension product corresponding to cleavage at /AT. Though it is possible that the 5' endonuclease can also cleave at /TT, the wild-type terminus also appears as a doublet; a plausible explanation is that the last base before the branch structure is incorporated inefficiently by reverse transcriptase.

Because we can detect the species with the novel 5' terminus with a probe in the downstream exon (Figure 5), some (if not all) of these molecules must contain the entire intron as well as at least 625 bases of the second exon (see legend to Figure 8). The surprising location of this novel terminus led us to consider two models for its formation. In the first, the terminus is the result of an aberrant endonucleolytic cleavage by the normal splicing machinery in response to the mutated 5' junction. Alternatively, the novel terminus could simply reflect a degradative process which occurs when mutant full-length precursor accumulates.

GENERATION OF THE NOVEL TERMINUS IS DEPENDENT ON 3' PROCESSING SIGNALS

If the novel terminus is dependent on the function of the splicing machinery, we reasoned that other mutations which block splicing should also prevent the generation of this aberrant intermediate. These include mutations which remove or inactivate the TACTAAC box (Langford and Galwitz, 1983; Pikielny et al., 1983). In collaboration with John Rossi (Cellini et al., 1985)), we have constructed a small deletion which precisely removes the TACTAAC box from the yeast actin intron. This mutation, Δ8, reduces mature mRNA formation to approximately 10% of wild-type levels. As shown in Figure 8 (lane 3), by a primer extension assay, the only intron-containing species which accumulates in this strain is full-length precursor; we do not detect any molecules cleaved near or at the 5' junction. From these results we can conclude that the TACTAAC box must be present for

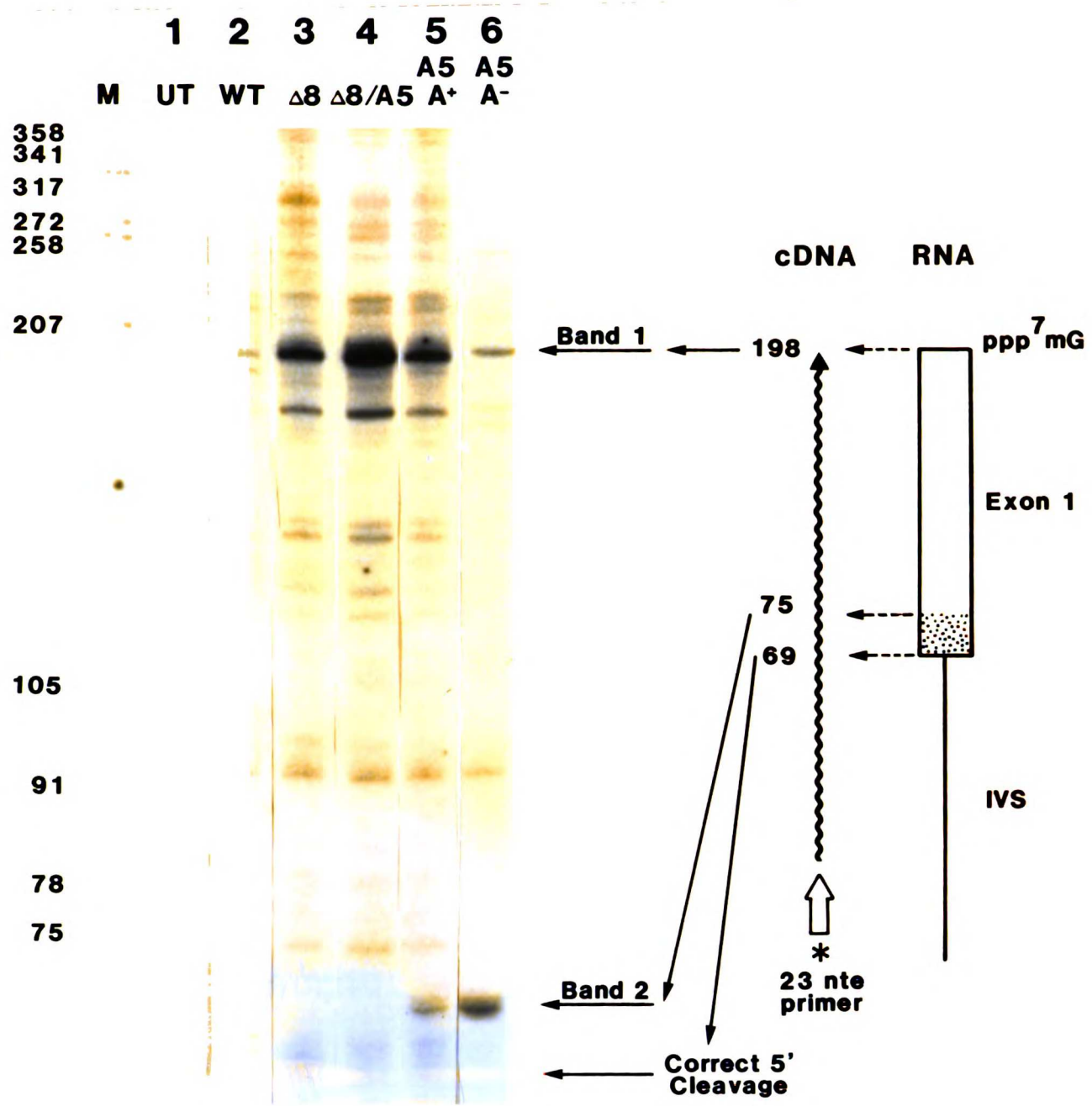

1
2
3
4
5
6
M
UT
WT
Δ8
Δ8/A5
A5 A+
A5 A-
358
341
317
272
258
207
105
91
78
75
Band 1
Band 2
Correct 5' Cleavage
cDNA
RNA
198
75
69
ppp7mG
Exon 1
IVS
*
23 nte primer

Figure 8 Primer Extension from near the 5' Junction

PolyA+ RNA was prepared from FC8-24D transformed with either wild-type or the A5 mutant fusion integrated at HIS4 (see Experimental Procedures). The oligonucleotide 5'CTCTCGAGCAATTGGGACCGTGC3' was annealed and extended with reverse transcriptase and cDNA products run on a 6% sequencing gel. Sizes of the predicted products are shown schematically. Lane 1, untransformed parent; lane 2, wild-type; lane 3, Δ8 fusion; lane 4, Δ8/A5 double mutant; lane 5 A5 mutant; lane 6, A5 mutant polyA- RNA. THis band just below band 1 in lanes 3,4 and 5 corresponds to a second minor transcriptal start which we detect in all our experiments. Note that, because of the low level of transcription from a single copy gene (see Discussion), we are unable to detect any molecules cleaved at the correct 5' junction, either in wild-type (lane 2) or in the A5 mutant (lane 5). In the A5 RNA preparation we detect an enrichment of incorrectly cut molecules in the polyA- fraction (compare lanes 5 and 6). While this could be caused by cleavage at the 3' intron/exon junction we also detect an enrichment with probes in the 3' exon (data not shown). This suggests that this enrichment could be a result of degradation form the 3' end of the RNA. Consistent with this hypothesis, we detect a smaller proportion of this species by Northern analysis as compared to S1 analysis or primer extension. The markers are kinased Sau3A-cut pBR322 DNA.

endonucleolytic cleavage at a wild-type 5' splice junction. We then constructed a double mutant containing the A5 mutation and the Δ8 deletion in cis, to test the prediction that the generation of the aberrant terminus is similarly dependent on the presence of this essential downstream intron signal.

Poly A+ RNA was prepared from strains carrying the A5-Δ8 double mutant fusion and subjected to primer extension from a 23 base primer whose 5' end is 69 bases downstream from the 5' junction. As shown in Figure 8 (lane 4), in strains carrying the double mutant we detect only the 198 nucleotide extension product corresponding to full-length precursor. We can thus conclude that the generation of the aberrant 5' terminus in the A5 mutant is the result of an incorrect cleavage by the normal splicing machinery.

MOLECULES GENERATED BY INCORRECT CLEAVAGE ARE NOT LIGATED INTO mRNA

In light of the foregoing conclusion, it became of interest to know whether molecules generated by cleavage upstream of the 5'intron/exon junction correspond to productive intermediates and lead to the production of mature mRNA. Cleavage at the incorrect position presumably generates a 5' exon shortened by six bases. (We have not yet identified this species directly; to increase the sensitivity necessary to detect this small and possibly unstable fragment, constructions designed to overproduce the mutant transcript are in progress.) If the shortened 5' exon is ligated to the downstream exon, the resulting

mature mRNA should thus be six nucleotides shorter. To test for this, we performed a primer extension assay using a 21 base oligonucleotide primer located in the downstream exon, 269 bases 3' to the 3' end of the intron. As illustrated schematically in Figure 9A, measurement of the length of the mRNA produced establishes the size of the intron which is removed.

In strains carrying the wild-type fusion we detect a single extension product of 398 bases, the expected size for spliced mature RNA (Figure 9B, lane 2, Band 2). In the A5 mutant the only mature RNA observed is the same size as wild-type, 398 nucleotides (lane 1, Band 2); i.e., the size corresponding to cleavage and ligation at the correct 5' junction. The expected position for ligation of the shortened exon is shown as Band 2*. We estimate that we could easily detect the ligation of 5% of the incorrectly cleaved molecules. From these results we conclude that the species generated by aberrant cleavage upstream of the normal junction is not ligated into stable mature mRNA. We cannot rule out the formal possibility that these shortened exons are efficiently ligated into mature RNA, but that this mutant mRNA is rapidly degraded.

MOLECULES GENERATED BY INCORRECT CLEAVAGE ARE BLOCKED TO PRIMER EXTENSION AT THE TACTAAC BOX

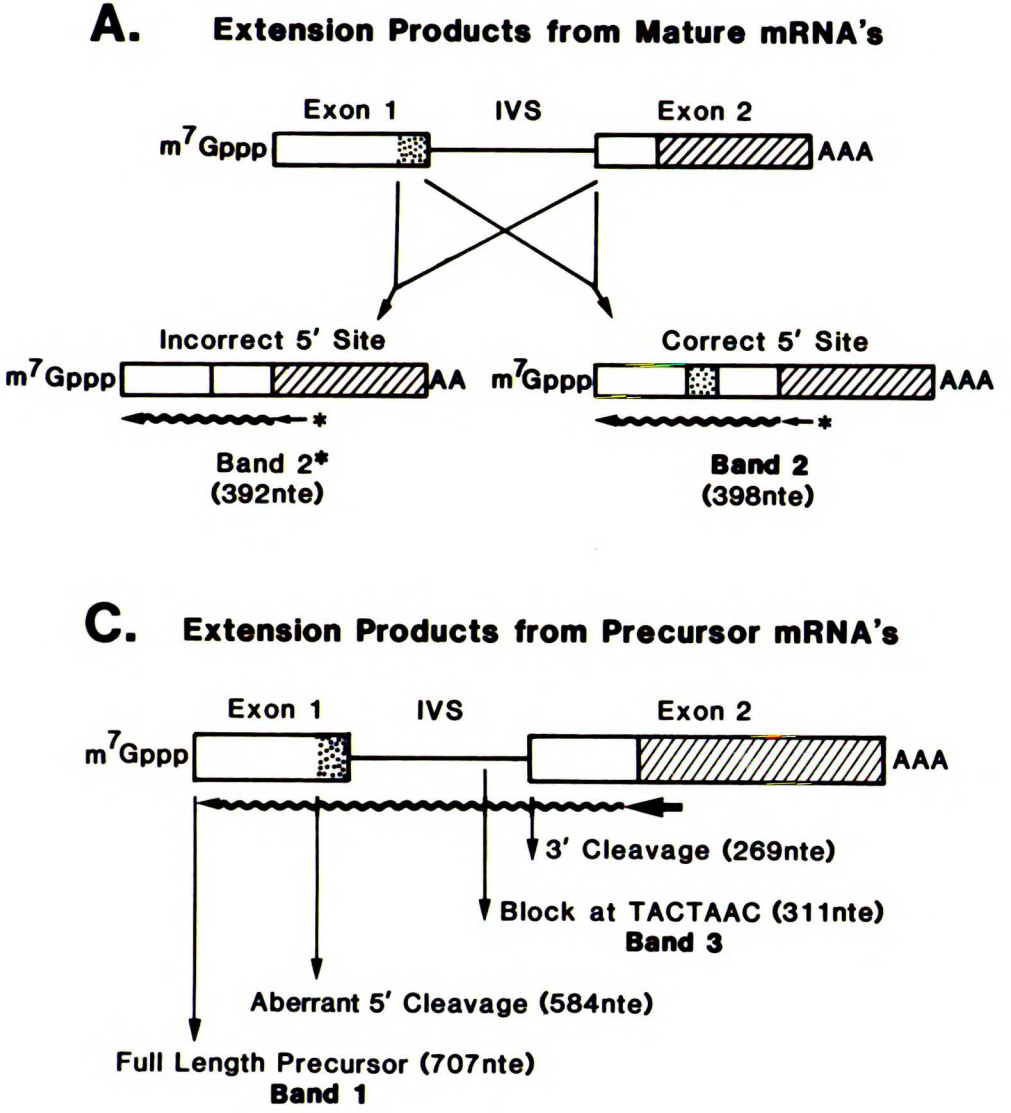
In the primer extension experiment we have just described, we expected to detect products derived from the mutant fusion which correspond to both full-length precursor and to incorrectly cleaved

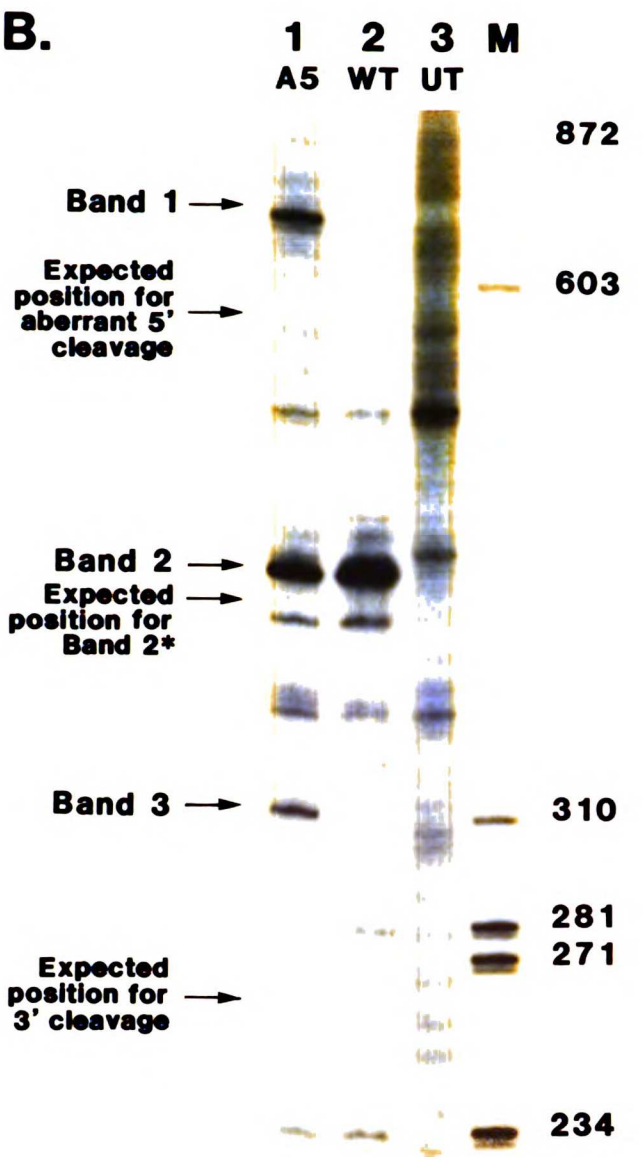
Figure 9 Primer Extension from the 3' Exon

A) Products predicted for extension from mature mRNA species are shown. The correct splice will yield a cDNA of 398 bases. Productive splicing of molecules cleaved at the incorrect position will produce a mature mRNA shortened by 6 nucleotides, leading to a cDNA of 392 nucleotides. As in Figure 1, open boxes represent actin sequences, slashed boxes HIS4 sequences, the speckled region represents the bases between the 5' intron/exon junction and the incorrect cleavage site.

B) PolyA+ RNA was prepared from FC8-24D transfromed with wild-type (lane 2) and A5 versions (lane 1) of the fusion plasmid pYAH-12 and treated as described in the Experimental Procedures. Lane 3 is the untransformed parent. The primer for this experiment is the 21mer, 5'CTCTTCATTACTCAGGCTCGA 3,' which specifically hybridizes to the actin-HIS4 junction. The major mature mRNA is as denoted (Band 2). The slightly smaller band in this size range is a second transcriptional start site which is ca. 14 bases 3' of the predominant start site. We see this band in all our experiments (e.g., see Figure 8). Sizes were estimated by comparison to the markers shown (HaeIII-cleaved ϕX DNA) and to Sau3A-cleaved pBR322 (not shown).

C) Extension products predicted for precursor mRNA species are shown.


A.
Extension Products from Mature mRNA's
Exon 1
IVS
Exon 2
m7Gppp
AAA
Incorrect 5' Site
Correct 5' Site
m7Gppp
AA
m7Gppp
AAA
Band 2*
(392nte)
Band 2
(398nte)
C.
Extension Products from Precursor mRNA's
Exon 1
IVS
Exon 2
m7Gppp
AAA
3' Cleavage (269nte)
Block at TACTAAC (311nte)
Band 3
Aberrant 5' Cleavage (584nte)
Full Length Precursor (707nte)
Band 1


B.
1 2 3 M
A5 WT UT
872
Band 1
Expected position for aberrant 5' cleavage
603
Band 2
Expected position for Band 2*
Band 3
310
281
271
Expected position for 3' cleavage
234

molecules (see Figure 9C). We did indeed detect a band with the correct mobility for the full-length precursor, 707 nucleotides (Figure 9B, lane 1, Band 1). However, even after lengthy exposure of the autoradiogram, we were unable to detect any extension products with a 5' terminus at the site of the aberrant cleavage, 584 bases (Figure 9B, lane 1). Recall that by S1 mapping from a similar position, we could easily detect cleavage at this site (Figure 5, lane 4). Using this downstream exon primer, we instead found a new band which corresponds to a reverse transcriptase stop 311 nucleotides from the primer (Figure 9B, lane 1, Band 3). We believe that this is a block to primer extension, rather than a 5' terminus, since in the latter case we would also detect this species by S1 analyses (cf. Figure 5). Moreover, the complete absence of a cDNA extended to this novel terminus (see expected position in Figure 9B) as compared to the strong signal for full-length precursor (Band 1), indicates that this primer extension block is specific for 5' cleaved molecules. We conclude from these results that <u>all</u> the molecules which possess the novel 5' terminus contain blocks to primer extension.

Interestingly, this primer extension stop in the A5 mutant is at the position of the TACTAAC box. By using a primer located at the 3' end of the intron, we have been able to map this site to the 3'-most AC of this sequence (data not shown). These characteristics -- a strong reverse transcriptase stop which does not correspond to an S1 sensitive-site -- have recently been shown to describe the behavior of the branch point in RNA molecules with a lariat structure (Ruskin et al., 1984; Padgett et al., 1984; Domdey et al., 1984; Rodriguez et al.,

1984). Moreover, the position of the primer extension block in the A5 mutant maps precisely to the site of branch formation identified by Domdey et al. (1984) in splicing intermediates derived from a derivative of the wild-type yeast actin gene in vivo. Taken together, these observations suggest that the primer extension stop in the A5 mutant is likely to be the result of lariat formation.

DISCUSSION

THE A5 MUTATION AFFECTS SPLICING

In conjunction with the construction of defined mutations *in vitro* (Cellini et al., 1985), we have taken advantage of the yeast genetic system to screen *in vivo* for cis-acting mutations which affect splicing. For this purpose, we constructed a translational fusion such that expression of the *HIS4* gene product requires proper splicing of the yeast actin intron. Using this chimeric gene, we have identified a mutation, A5, which causes both a decrease in mature mRNA levels and an increase in intron-containing precursors. Thus splicing has become the rate-limiting step in the production of the A5 mRNA, accounting for the Hol- phenotype of this strain.

The A5 mutation is a G to A transition at the fifth position of the consensus sequence /GTAPyGT found at the 5' intron-exon junction of all yeast mRNA nuclear introns (Teem et al., 1984).

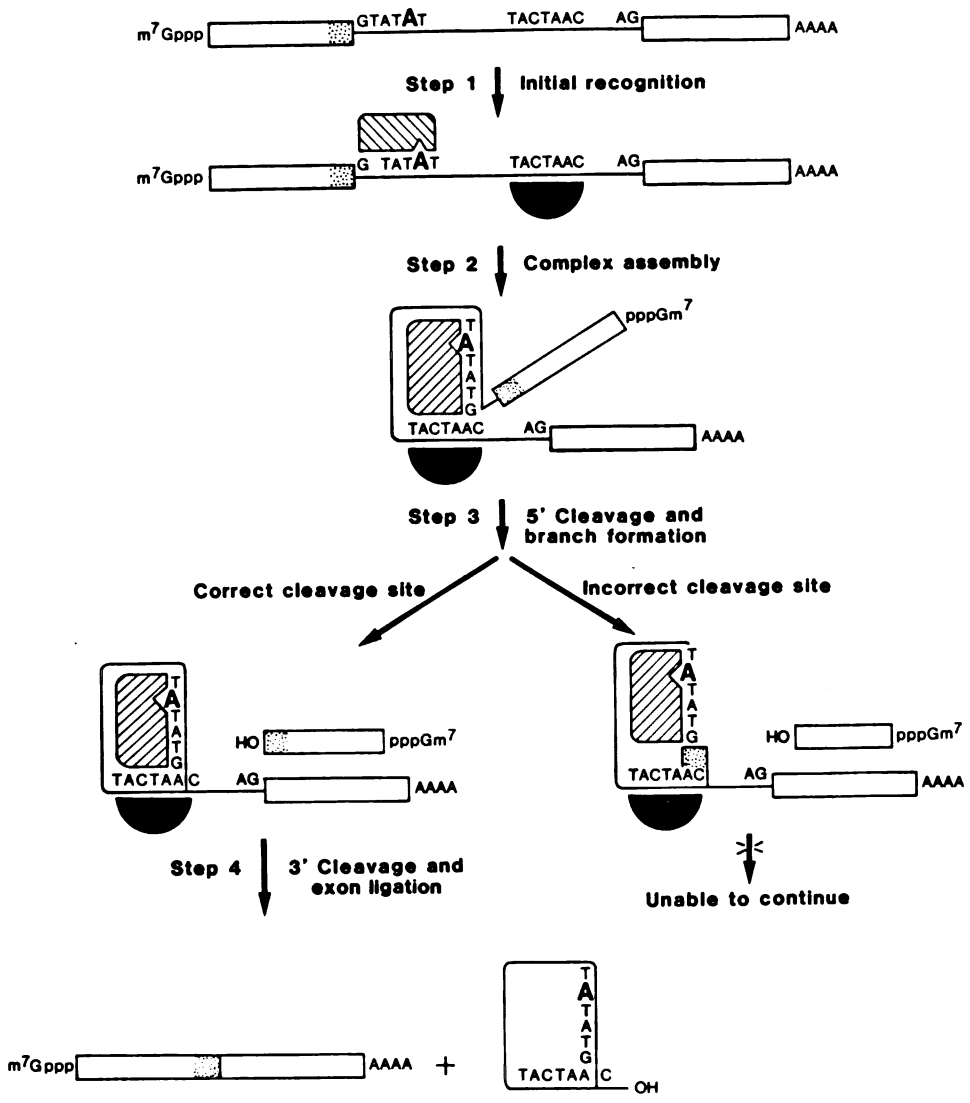
5' Consensus Sequence	/GTAPyGT
	*
A5 Mutant Sequence	/GTAPyAT

While the importance of this hexanucleotide in splicing has been implicated previously by deletion analysis (Gallwitz, 1982; Pikielny et

al., 1983), the observation that a single point mutation in this sequence has a dramatic effect on intron removal confirms and extends the hypothesis that these nucleotides are critical to the splicing process. This may also be a particularly important position within the mammalian consensus sequence; with the exception of the invariant GT, this G is the most highly conserved nucleotide (84%) within the 5' consensus sequence (Mount, 1982).

A PUTATIVE SPLICING PATHWAY

Our ultimate goal is to understand on a molecular level the consequences of the A5 mutation for the splicing process. As a first step, it is helpful to consider a plausible splicing pathway; in Figure 10 we have cartooned one such hypothesis, which draws on recent observations from a number of laboratories (Domdey et al., 1984; Kramer et al., 1984; Mount et al., 1983; Padgett et al., 1984; Rodriguez et al., 1984; Ruskin et al., 1984). In this pathway, the splicing process is initiated (Step 1) with the recognition of the 5' consensus sequence and the TACTAAC box by one or more trans-acting factors, minimally including the U1 snRNP. The precursor and the various components of the splicing machinery are then assembled into a presumptive splicing complex (Step 2), in what may be an ATP-dependent process. In variations on this theme, this step is specifically viewed as a linear scanning process (see, e.g., Lewin, 1980); in either case, the final consequence is presumed to be the physical juxtaposition of the 5' and 3' intron elements. Step 3 comprises cleavage at the 5' junction and formation of the 2'-5'


m7Gppp
GTATAT
TACTAAC
AG
AAAA
Step 1
Initial recognition
G TATAT
Step 2
Complex assembly
pppGm7
TACTAAC
Step 3
5' Cleavage and branch formation
Correct cleavage site
Incorrect cleavage site
HO
TACTAA C
Step 4
3' Cleavage and exon ligation
Unable to continue
OH

phosphodiester branch at the TACTAAC box; whether or not 5'cleavage and lariat formation are mechanistically coupled is not yet known. After formation of this branched intermediate, the 3' exon must be released by cleavage at the 3' intron-exon junction and ligated to the 5' exon (Step 4). Here again there is no information as to whether these events are concerted.

THE A5 MUTATION AFFECTS AN EARLY STEP IN THE SPLICING PATHWAY

Since one of the biochemical consequences of the A5 mutation is an accumulation of full-length precursor, this G to A transition must be affecting one or more steps required for the first endonucleolytic event in the splicing pathway to occur. According to the pathway depicted in Figure 10, we can identify at least two roles that the 5' consensus sequence could play, first as part of the initial recognition process (Step 1), and subsequently as a substrate for the 5' endonuclease (Step 3). While the further refinement of in vitro splicing systems (Newman et al., 1985) will be required to operationally distinguish between these reactions, it seems to us likely that at least one effect of the A5 mutation is exerted at the level of initial recognition (Step 1), as we will now discuss.

Two models have been proposed for the selection of the 5' junction of introns. It has been suggested that, in yeast, the 5' junction is recognized via intramolecular base-pairing with the nucleotides comprising the TACTAAC box (Pikielny et al., 1983; Rodriguez et al.;

1984). The effect of the A5 mutation on this putative helix would be to substitute an A-C for a G-C base pair. In that this would destroy one of only four base pairs in a stem initially containing two internal mismatches, the observation that the A5 mutant still produces 30-50% of correctly spliced mRNA demonstrates that this specific intramolecular interaction cannot be absolutely required, as initially predicted by Pikielny et al. (1983). To determine whether this potential helix might play some more subtle role in the splicing process will require measuring splicing efficiency in a mutant construction containing a compensatory change in the TACTAAC box.

While these experiments are in progress, we nonetheless favor the initial hypothesis that primary recognition of intron junctions occurs via intermolecular complementarity between the substrate and an snRNA (Lerner et al., 1980; Rogers and Wall, 1980). Experimental support for an essential role of U1 RNA in the splicing reaction involving direct physical interaction with the 5' junction is now strong (Kramer et al., 1984; Mount et al., 1983; Padgett et al., 1983). Yeast has been shown to possess snRNAs (Wise et al., 1983), although the analog of U1 has not yet been identified. According to this view, then, the accumulation of full-length precursor in the A5 mutant would be due to the destabilization of this intermolecular helix resulting from the replacement of a G-C base pair by an A-C juxtaposition.

THE A5 MUTATION UNCOUPLES RECOGNITION AND CLEAVAGE AT THE 5' JUNCTION

Fidelity of Cleavage is Decreased

Strains carrying the A5 mutant fusion also accumulate a second intron-containing species, the terminus of which is located six bases 5' to the normal 5' junction. Since the generation of this species is dependent on the presence of a downstream 3' processing signal (the TACTAAC box) that is itself required for cleavage at the wild-type 5' junction (see Cellini et al., 1985), we can conclude that this novel terminus arises as a result of endonucleolytic cleavage by the normal splicing machinery at an abnormal site. The important question then becomes whether the effect of the A5 mutation is exerted by decreasing the fidelity of the endonuclease reaction directly (i.e., at Step 3), or indirectly (at Step 1).

In the latter case, one imagines that the A5 mutation has so decreased the efficiency with which the primary recognition element binds at the correct position that an alternative binding site can now compete sucessfully. According to this model, endonucleolytic cleavage is thus directed to the inappropriate site by the prior binding of the U1 snRNP, e.g., at this location. We consider this possibility unlikely for a number of reasons.

Perhaps the most compelling argument is that, while rigid sequence conservation at the 5' junctions appears to be one of the most marked

characteristics of yeast introns (Teem et al., 1984), the sequence in the region of the aberrant cleavage exhibits virtually no homology to the 5' consensus sequence.

5' CONSENSUS	/GTAPyGT
	* *
ABERRANT CUT SITE	/ATTC TG

Indeed if the splicing machinery were to select an alternative site based on homology to the 5' consensus, a much better match, /GTAAGA, can be found 35 bases within the intron. Our data show no evidence that this site is used in the A5 mutant. Moreover, a mutant of the yeast actin gene has been constructed in which all but the first G of the 5' consensus sequence has been deleted, while leaving intact the sequence in the region of the aberrant cleavage; this mutant has been shown to accumulate full-length precursor only (Gallwitz, 1981). Taken together, these observations argue quite persuasively that 1) cleavage at the incorrect site is dependent on and separable from recognition of the 5' consensus sequence; and 2) that the fidelity of the 5' cleavage event (Step 3) is altered by the A5 mutation.

5' Endonuclease Can Also Cleave at /AT

In trying to understand the molecular basis of the decreased fidelity of the 5' endonuclease on this mutant substrate, it may be useful to consider the behavior of two cryptic junctions in rabbit

β-globin (Wieringa et al., 1983), where the site of 5' cleavage (and subsequent ligation) is not at the expected position within the consensus sequence. In their effort to interpret these results, Mount and Steitz (1983) also suggested that the apparently constant positions of the consensus sequence and the 5'cut site relative to one another may reflect two separable aspects of the splicing mechanism; for example, the consensus match may reflect the requirement for formation of a double helix with the U1 RNA, while the 5' endonuclease may be flexible enough to act at sites other than the preferred /GT at its normal position within the intermolecular helix.

That we detect inappropriate cleavage at only a single site suggests the likely possibility that the endonuclease exhibits some degree of local sequence specificity, perhaps the dinucleotide /PuT. Consistent with this hypothesis, low level cleavage at the sequence /AT may also occur when the A5 mutation is present in another spliced yeast gene (A. Jacquier and M. Rosbash, pers. comm.). Note that in our experiments, cleavage at an alternative location occurs despite the fact the the normal site of cleavage, GT, is <u>not</u> altered by the A5 mutation. Coupled with the observation that the 5' endonuclease can cleave at /AT, this leads us to conclude that the extraordinary conservation of the G at the 5' intron/exon junction is almost certainly reflecting a primary role for this residue in <u>other</u> aspects of the splicing process (see below).

THE ABERRANT CLEAVAGE PRODUCES A DEAD-END INTERMEDIATE

In the mammalian cases we have just described, endonucleolytic cleavages at inappropriate sites with respect to the consensus sequence are nonetheless productive splicing events in that the corresponding mRNAs are produced (Wieringa et al., 1983). In the case of the A5 mutant, however, we have been unable to detect transcripts of the predicted size; i.e., as described in Figure 9, primer extension assays fail to reveal mRNA molecules shortened by six bases. Results completely consistent with this observation have also been provided by an independent method, Northern analysis using a 17 base deoxyoligonucleotide which spans the correct splice junction and, under the hybridization conditions employed, is specific for correctly spliced mature mRNA (data not shown). We conclude from these experiments that molecules arising by cleavage at the inappropriate site cannot be ligated into stable mature RNA. In other words, incorrect cleavage produces aberrant intermediates which are unable to undergo the final step in the splicing pathway, 3' endonucleolytic cleavage and exon ligation (Step 4, Figure 10).

Interestingly, molecules with the aberrant 5' terminus appear to be in a lariat structure. This inference is derived from a number of experimental observations. S1 analysis (Figure 5) demonstrates that these molecules contain (at least a portion of) the 3' exon, yet all molecules cleaved at this 5' site are blocked to primer extension from a primer located downstream of the TACTAAC Box (Figure 9). The block is specific for cleaved molecules, and the position of the block

corresponds precisely to the branch point mapped in the yeast actin intron by Domdey et al. (1984). A more direct test for the lariat structure of these molecules will come from further experiments with a shortened form of the fusion, which would then enable resolution of molecules with the predicted anomalous electrophoretic mobility (Ruskin et al., 1984; Grabowski et al., 1984; Domdey et al., 1984; Rodriguez et al., 1984). In the meantime, it is worth noting that the sequence of the novel 5' terminus (/AT) suggests that, as for endonucleolytic cleavage per se, branch formation cannot have an absolute requirement for a G residue in 2'-5' linkage to the A in the TACTAAC box; recent experiments by Abelson and his colleagues (Newman et al., 1985) have demonstrated that 5' endonucleolytic cleavage at an A residue can indeed lead to lariated intermediates in an in vitro splicing system.

This leads us to speculate as to why this species appears to be, at best, an inefficient substrate for the 3' endonuclease. There are several alterations produced by 5' cleavage at the incorrect site, any one of which may be sufficient to affect 3' cleavage. In the simplest case, the presence of an A at the branch point, a position normally occupied by a G, might cause this species to be a poor substrate for the 3' endonuclease; as we have discussed above, the strong conservation of this G may reflect a more important role in 3' cleavage than in 5' cleavage. In addition, it seems important to consider the possible consequences of six additional nucleotides in the lariat, and of the shortened 5' exon. As we have illustrated schematically in Figure 10, the splicing process is almost certainly taking place in some type of complex. Conceivably, then, the spatial arrangement of these sequence

elements is also a critical parameter in determining the efficiency of this substrate in the splicing reaction. Such displacements within the complex (Step 3, Incorrect Cut), may disturb crucial interactions between substrate elements and components of the splicing machinery.

It is, in any case, important to note that the inability of this aberrant lariat to undergo 3' cleavage and ligation is not a direct result of the A5 mutation, but is instead a consequence of the incorrect 5' cleavage. That is, molecules which contain the A5 mutation but are cleaved at the correct intron/exon junction (Step 3, Correct Cut) appear to be efficient substrates for Step 4, in that we never observe transient accumulation of these lariated intermediates (cf. Figures 5-9). Thus we can conclude that while the G to A transition appears to affect the efficiency of initial recognition (Step 1), and the fidelity of 5' endonucleolytic cleavage (Step 3), this substitution does not itself affect 3' endonucleolytic cleavage (Step 4).

With this information in hand, we can now use the yeast genetic system to probe the specific interactions between substrate and splicing components, by selecting second-site suppressors of the A5 mutation. In particular, these suppressors should include trans-acting elements of the splicing machinery directly involved in initial recognition and 5' endonucleolytic cleavage.

EXPERIMENTAL PROCEDURES

Materials

Restriction enzymes were purchased from New England Biolabs. γ - ^{32}P-ATP was purchased from ICN; α^{32}-P-dCTP from Amersham. Reverse transcriptase was obtained from Life Sciences. S1 Nuclease was obtained from Sigma. Oligonucleotides were provided by John Rossi. Yeast media was prepared as described by Sherman et al. (1979).

Construction of Actin-HIS4 Fusion Plasmids

The yeast actin-HIS4 (YAH) gene fusion was constructed by cleaving YIp25 (Donahue et al., 1982) with XhoI, filling in the resulting overhang with Klenow polymerase and blunt ligation of the BamHI linker, 5'GGGGATCCCG 3'. The DNA was then partially digested with EcoRI, and the 3.3 kb XhoI-RI partial product containing the HIS4 coding sequence was purified, and ligated into pBR322 which had been digested to completion with EcoRI and BamHI. The resulting plasmid, pYAH0, was cut with BamHI, and the 1047 base BamHI-BglII fragment from pYact1 (Ng and Abelson, 1980) which contains the 5' region of the actin gene (including the intron), was inserted in the correct orientation to allow translational fusion; this plasmid is designated pYAH1.

To produce an integrating plasmid containing the fusion, pYAH1 was digested with BamHI, to completion, and partially with EcoRI; the 4.3 kb fragment containing the fusion was then cloned into the BamHI to EcoRI sites of YIp5 (Struhl et al., 1979). This plasmid, pYAH-I1, was targeted (Rothstein, 1983) to the yeast actin locus by cleavage at the XhoI site within the actin intron (Figure 1). A second integrating plasmind, pYAH-I2, was constructed by subcloning a 4.4 kb SphI fragment containing the fusion and 187 bp of pBR322 from YCp50-YAH into the SphI site of YIp5. This plasmid was targeted to the HIS4 locus by cleavage at the XbaI site in the 3' flanking sequence which remains in the his4-401 deletion strain. The centromere plasmid, YCp50-YAH, was constructed by subcloning the 4.3 kb BamHI-EcoRI partial product into YCp50 (C. Mann and R. Davis, pers. comm.).

Yeast Strain Constructions

FC2-12B was constructed by first crossing 7375-2A (from G. Fink) to HR125-15D (obtained from R. Jensen) to generate FC1-6B, mat α, ura3-52, leu2, trp1-1, ade5. FC1-6B was then crossed to TD77, carrying the HIS4 deletion his4-401 (provided by T. Donahue), to generate FC2-12B, mat α, trp1-1, ura3-52, leu2, his4-401, HOL1. FC8-24D was constructed by first inactivating the SUP6 suppressor in 513V by selection on KCl media (Singh et al., 1979); 513V is a strain in which the intron within the chromosomal copy of the actin gene has been precisely deleted (Ng et al., 1985). This derivative of 513V was then crossed to FC4-17B, a derivative of FC2-12B, to yield FC8-24D, which is mata, leu2, trp1-1, ura3-52, his4-401, HOL1, actinIVSΔ. The

deletion of the actin intron was followed in the tetrads by Southern (1975) analysis.

The rna2 strain carrying the fusion was constructed by crossing FC2-12B, with pYAH-I1 integrated at the actin locus, to an rna2-containing strain provided by J. Haber (mat α, rna2, trp1, leu2, arg4). Sporulation and dissection were performed as previously described (Sherman et al., 1979).

In Vitro Mutagenesis and Yeast Transformation

100 ug of YCp50-YAH was mixed with an equal volume of a solution containing 2 M NH_2OH-HCl, 4 mM EDTA, 100 mM pyrophosphate, 200 mM NaCl adjusted to pH 7.0 with solid NaOH pellets, heated at 75°C for 0-90 minutes, and cooled on ice 10 minutes. The DNA was recovered by spin dialysis twice through P10, followed by ethanol precipitation. Yeast transformations were performed by the glusulase method of Hinnen et al. (1978). Approximately 2500 transformants were replated on -Ura media and screened as described below.

Genetic Screens

Biological Tests for HIS4 Function

Complementation tests were performed by cross-stamping the test strain with the appropriate his4 mutant, incubating overnight at 30°C on YEPD, replica-plating to -His media, and scoring for growth after

24-76 hours. Strains used were obtained from G. Fink: his4 A-B-C-, E260; his4 A-, 1651/7; his4 B-, E331; his4 C-, 5942-1D. Our lab strain WF100 (mata, his3) was used as a positive control in all tests. Histidinol (Hol) media was prepared by adding 3 ml of filter-sterilized 1 M histidinol to a liter of synthetic media (lacking histidine), after autoclaving. From the 2500 original transformants screened, 58 His A-B-C- mutants were obtained.

Deletion Mapping

To map these A-B-C- mutations with respect to their relative position in the fusion (i.e., in actin vs. HIS4), we took advantage of the presence of the chromosomal actin gene in FC2-12B. That is, mutations within the actin sequences of the fusion should be repairable by gene conversion from the actin gene in the chromosome, whereas mutations within the HIS4 sequences cannnot be repaired due to the deletion of the chromosomal locus. As expected, two classes of Hol- mutants were obtained, those which can papillate on Hol media (thus mapping to actin sequences), and those which fail to papillate (mapping to HIS4 sequences). We have demonstrated that papillation is dependent on the presence of a functional rad52 gene product and is stimulated by UV radiation (data not shown). Mutants which fail to papillate could be induced to do so by supplying HIS4 sequences in trans. Of the 47 mutations analyzed, 15 mapped to the actin sequences. The other 11 mutants were too leaky for this analysis.

Co-reversion

To determine if the loss of all three HIS4 activities was due to a single lesion, Hol$^+$ papillae (i.e., HIS4 C$^+$) were picked and then tested for co-reversion of HIS4 A and B activities by complementation. 11 out of 15 mutants tested co-reverted; only these mutants were analyzed further.

Nonsense Suppression

To determine if any of these 11 mutations were nonsense, each of these strains was crossed to FC0-1A, with the relevant genotype mata, SUP4-o, and his4-1176. In light of the actin sequence and the mutagen specifity, 9/10 of the possible nonsense mutations should be ochres. Of the 11 strains tested, two carried mutations suppresible by SUP4-o.

DNA Sequence Analysis

MP9 subclones of the fusion from the XhoI site within the actin intron extending to the SalI sites in either pBR322 (5' portion of the gene) or in the HIS4 coding sequence (3' portion of the gene) were sequenced by the dideoxy chain termination method of Sanger et al. (1977).

Preparation of RNA

For experiments with total RNA, RNA was prepared by the glass bead lysis technique as described by Jensen et al. (1983). For the preparation of polyA+ RNA, RNA was purified as described by Domdey et al. (1984).

Northern Analysis

50 µg of total RNA was electrophoresed in each lane on formaldehyde gels by the method of Derman et al. (1981), as modified by Jensen et al. (1984).

S1 Nuclease Protection

For S1 protection experiments using probes within the 3' exon, YCp50-YAH was cut with SalI, kinased, cleaved again with BamHI, and the resulting 1422 base end-labeled fragment purified. For S1 protection experiments within the intron, DNA was end- labled at the XhoI site, cleaved with BamHI, and the 550 bp fragment purified. DNA fragments were annealed to total RNA in 80% formamide, 50 mM NaCl, 10 mM PIPES pH 6.4, 1mM EDTA, at 42 C for 3-12 hours. Samples were treated with different amounts of S1 (36 → 900 units) in 280 mM NaCl, 30mM NaOAc, 4.5mM $Zn(OAc)_2$, 20 ug/ml single stranded DNA for 30' at 30°C. 2.5M NH_4OAc, 50 mM EDTA was added to stop the reaction. The protected species were ethanol precipitated and run on sequencing gels as described in the Figure legends.

Primer Extension

Primer extension reactions were performed as described by Domdey et al. (1984). 5 µg of poly A+ RNA was used in each experiment.

Acknowledgments

We are grateful to Gerry Fink and Tom Donahue for providing several HIS4 strains and the HIS4 clone; to Ray Ng and John Abelson for providing the actin clone; to Rob Jensen and Jim Haber for providing additional yeast strains. We are indebted to John Rossi for helpful discussions, as well as providing strains, and oligonucleotide primers. C.G. also thanks John Abelson and the members of his laboratory for their generous help and hospitality during a sabbatical visit. Lucita Esperas provided excellent technical assistance. We gratefully acknowledge Michael Rosbash, John Abelson and their colleagues for providing information prior to publication. We thank Bruce Patterson, Trey Simmons, Harold Swerdlow, Joe Tamm, and Ken Zaret for critical discussions and valuble suggestions on the manuscript. This work was supported by NIH grant GM21119.

REFERENCES

Abelson, J. (1979). RNA processing and the intervening sequence problem. Ann. Rev. Biochem. 48, 1035-1069.

Cellini, A., Parker, R., McMahon, J., Guthrie, C., and Rossi, J. (1985). Activation of a cryptic TACTAAC box in the yeast actin intron. Mol. Cell. Bio., submitted.

Derman, E., Drauter, K., Walling, L., Weinberger, C., Ray, M., and Darnell, Jr., J.E. (1981). Transcriptional control in the production of liver-specific mRNAs. Cell 23, 731-739.

Domdey, H., Apostol, B., Lin, R.J., Newman, A., Brody, E., and Abelson, J. (1984) Lariat structures are in vivo intermediates in yeast pre-mRNA splicing. Cell, in press.

Donahue, T.F., Farabaugh, P.J., and Fink, G.R. (1982). The nucleotide sequence of the HIS4 region of yeast. Gene 18, 47-59.

Gallwitz, D. (1982). Construction of a yeast actin gene intron deletion mutant that is defective in splicing and leads to the accumulation of precursor RNA in transformed yeast cells. Proc. Nat. Acad. Sci USA 79, 3493-3497.

Gallwitz, D., and Sures, I. (1980). Structure of a split yeast gene: complete nucleotide sequence of the actin gene in Saccharomyces cerevisiae. Proc. Nat. Acad. Sci. USA 77, 2546-2550.

Grabowski, P.J., Padgett, R.A., and Sharp, P.A. (1984). Messenger RNA splicing in vitro: an excised intervening sequence and a potential intermediate. Cell 37, 415-427.

Hinnen,A., Hicks, J.B., and Fink G.R., (1978). Transformation of yeast. Proc. Nat. Sci. USA 75, 1929-1933.

Jensen, R., Sprague, Jr., J.F., and Herskowitz, I. (1983). Regulation of the yeast mating-type interconversion: feedback control of the HO gene expresssion by the mating type locus. Proc. Nat. Acad. Sci USA 80, 3035-3039.

Keesey, Jr., J.K., Bigelis, R., and Fink, G.R. (1979). The product of the HIS4 gene cluster in Saccharomyces cerevisiae. J. Biol. Chem. 254, 7427-7433.

Keller, E.B., and Noon, W.A. (1984). Intron splicing: a conserved internal signal in introns of animal pre-mRNAs. Proc. Nat. Acad. Sci. USA, in press.

Kramer, A., Keller, W., Appel, B., and Luhrmann, R. (1984). The 5' terminus of the RNA moiety of U1 small nuclear ribonucleoprotein particles is required for the splicing of messenger RNA precursors. Cell 38, 299-307.

Langford, C.J., and Gallwitz, D. (1983). Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcripts. Cell 33, 519-527.

Lerner, M.R., Boyle, J.A., Mount, S.M., Wolin, S.L., and Steitz, J.A. (1980). Are snRNPs involved in splicing? Nature 283, 220-224.

Lewin, B. (1980). Alternatives for splicing: recognizing the ends of introns. Cell 22, 324-326.

Mount, S. (1982). A catalogue of splice junction sequences. Nucl. Acids. Res. 10, 459-472.

Mount, S., and Steitz, J. (1983). Lessons from mutant globins. Nature 303, 380-381.

Mount, S.M., Pettersson, I., Hinterberger, M., Karmas, M., and Steitz, J. (1983). The U1 small nuclear RNA-protein complex selectively binds a 5' splice site in vitro. Cell 33, 509-518.

Newman, A., Lin, R.J., Cheng, S-C and Abelson, J.A. In vitro mRNA splicing in yeast. Cell, submitted.

Ng, R., and Abelson, J. (1980). Isolation and sequence of the gene for actin in Saccharomyces cerevisiae. Proc. Nat. Acad. Sci. USA 77, 3912-3916.

Ng, R., Domeday, H., Larson, G., Rossi, J., and Abelson, J. (1985) A test for intron function in the yeast actin gene. Nature, in press.

Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F., and Sharp, P.A. (1984). Lariat RNAs as intermediates and products in the splicing of messenger RNA precursors. Science 225, 898-903.

Pikielny, C.W., Teem, J.L., and Rosbash, M. (1983). Evidence for the biochemical role of an internal sequence in yeast nuclear mRNA introns: implications for U1 RNA and metazoan mRNA splicing. Cell 34, 395-402.

Rodriguez, J.R., Pikielny, C.W., and Rosbash, M. (1984). In vivo characterization of yeast mRNA processing intermediates. Cell, in press.

Rogers, J., and Wall, R. (1980). A mechanism for RNA splicing. Proc. Nat. Acad. Sci. USA 77, 1877-1879.

Rosbash, M., Harris, P.K., Woolford, Jr., P.K., and Teem, J.L. (1981). The effect of temperature-sensitive RNA mutants on the transcription products from cloned ribosomal protein genes of yeast. Cell 24, 679-686.

Rothstein, R.J. (1983). One-step gene disruption in yeast. Methods Enzymol. 101, 202-211.

Ruskin, B., Krainer, A.R., Maniatis, T., and Green, M.R. (1984). Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell 38, 317-331.

Sanger, F., Nicklen, S., and Coulson, A.R. (1977). DNA sequencing with chain-terminating inhibitors. Proc. Nat. Acad. Sci. USA 74, 5463-5467

Singh, A., Helms, C., and Sherman, F. (1979). Mutation of the non-Mendelian suppressor, Ψ^+ in yeast by hypertonic media. Proc. Nat. Acad. Sci. USA 76 1952-1956.

Sherman, F., Fink, G., and Lawrence, C. (1979). Methods in yeast genetics. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Sollner-Webb, B., and Reeder, R.H. (1979). The nucleotide sequence of the initiation and termination sites for ribosomal RNA transcription in X. laevis. Cell 18: 488-499.

Southern, E.M. (1975). Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98, 503-517.

Struhl, K., Stinchomb, D.T., Scherer, S. and Davis, R.W. (1979). High-frequency transformation of yeast: autonomous replication of hybrid DNA molecules. Proc. Nat. Acad. Sci. USA 76, 1035-1039.

Teem, J.L., Abovich, N., Kaufer, N.F., Schwindinger, W.F., Warner, J.R., Levy, A., Woolford, J., Leer, R.J., van Raamsdonk-Duin, M.M., Mager, W.H., Planta, R.J., Schultz, L., Friesen, J.D., and Rosbash, M. (1984). A comparison of yeast ribosomal protein gene DNA sequences. Nucl. Acids. Res., in press.

Wallace, J.C. and Edmonds, M. (1983). Polyadenylylated nuclear RNA contains branches. Proc. Nat. Acad. Sci. USA 80, 950-954.

Wieringa, B., Meyer, F., Reiser, J., and Weissmann, C. (1983). Unusual splice sites revealed by mutagenic inactivation of an authentic splice site of the rabbit β-globin gene. Nature 301, 38-43.

Wise, J.A., Tollervey, D., Maloney, D., Swerdlow, H., Dunn, E.J., and Guthrie, C. (1983). Yeast contains small nuclear RNAs encoded by single copy genes. Cell 35, 743-751.

CHAPTER THREE

Activation of a Cryptic TACTAAC box

in the Actin Intron

SUMMARY

We have constructed a translational fusion between the yeast actin gene and the E. coli beta-galactosidase structural gene such that expression of beta-galactosidase activity requires accurate splicing of the actin intron. Using this chimeric gene, we have generated a series of internal deletions which remove the TACTAAC box, or in addition, TACTAAC-like sequences within the intron. Analysis of the fusion transcripts produced in these deletions has allowed to us to conclude that the TACTAAC-like sequence TACTAAG, can substitute, albeit inefficiently, for the correct TACTAAC box in the splicing process. This demonstrates that, at least in this instance, an alternative TACTAAC box can be used. This finding is in sharp contrast to failures to observe the use of cryptic 5' junctions in yeast. It also raises the possibility that TACTAAC-like sequences in introns have the potential to play a physiological role.

INTRODUCTION

A key step in the splicing process is the definition of the precise boundaries of the intron by the splicing machinery. In the yeast Saccharomyces cerevisiae, a subset of nuclear genes encoding mRNAs contain intervening sequences (Teem et al, 1984). Previous work has identified two regions within the intron which are required for splicing. The sequence at the 5'intron/exon junction, though subject to substantial variation about a consensus sequence in metazoans, in yeast is an essentially 100% conserved hexanucleotide (Langford and Gallwitz, 1983). This apparent lack of variance in yeast splicing signals allowed the identification of an additional conserved sequence within the intron, the so-called TACTAAC box, a heptanucleotide found in all yeast introns 4-53 nucleotides upstream of the AG at the 3' splice junction. The importance of these conserved sequences in the splicing process was demonstrated by the observation that deletion of either the 5' consensus sequence (Gallwitz, 1982) or the TACTAAC box (Pikielny et al., 1983; Langford and Gallwitz, 1983) abolishes splicing and leads to the accumulation of full-length precursor (op. cit.)

While consensus sequences at the 5'(GTAATG) and 3' ($(Py)_nAG$) intron/exon junctions had been recognized in metazoan introns (Mount, 1982), there was no evidence to suggest an analog of the TACTAAC box. This apparent difference has recently been resolved with the development of extracts which carry out the splicing reaction in vitro. A detailed analysis of the transcripts in both mammalian and yeast

extracts (Ruskin et al., 1984; Padgett et al., 1984), and more recently the characterization of intermediates in the splicing process in vivo (Domdey et al., 1984; Rodriguez et al., 1984; Zeitlin and Efstratiadis, 1984) has shown that splicing of pre-mRNA in both systems occurs via a common intermediate. This intermediate has been termed a lariat, because the 5' end of the intron is joined by a 2'-5' phosphodiester bond to a site within the intron. The location of this 2'-5' linkage, or branch (cf. Wallace and Edmonds, 1983), is within the TACTAAC box in yeast (Domdey et al., 1984). In mammalian cells, branch points seem to fall in comparable locations (with respect to the 3' junction), but an obvious consensus sequence is less readily identifiable (Padgett et al., 1984; Ruskin et al., 1984; Keller and Noon, 1984).

The novel nature of these intermediates argues that splicing must proceed by fundamentally similar mechanisms in yeast and mammals. Yet a significant distinction between these two systems is revealed by the response of the splicing machinery to mutation of the 5' splice junction. In striking contrast to mammalian cells, where inactivation of a 5' splice site almost invariably leads to the activation of an alternative or cryptic site (for review, see Mount and Steitz, 1983), deletions or point mutations in the 5' consensus sequence in yeast introns have so far failed to activate cryptic sites (Gallwitz, 1982; Pikielny et al, 1983; Parker and Guthrie, 1985). The failure of the yeast splicing machinery to use alternative 5' splice sites has proven to be extremely useful, by revealing the molecular consequences of the attempt to use a mutant junction (Parker and Guthrie, 1985). It

suggests, moreover, that yeast may be more stringent not only in the requirement for primary sequence of the splicing signals, but perhaps also with regard to the location of these signals within the intron.

An important distinction between the 5' consensus and the TACTAAC box is that the former, by virtue of its position at the intron/exon junction, directly determines the site of splicing. The latter is, rather, a "signal" (cf. Keller and Noon, 1984) which appears to specify the particular choice of a 3' junction (perhaps analogous to the role of the TATA box in directing the choice of a particular start-site of transcription). The apparent flexibility of the location of the TACTAAC box with respect to the 3' splice site (among different introns) is consistent with this notion. We might thus imagine that the TACTAAC box provides several functions in splicing, including 1) a recognition sequence for components of the splicing machinery; 2) a "branch acceptor" site for the 5' end of the lariat; and 3) a recognition sequence for the 3'endonuclease/ligase.

One of the intriguing features of some yeast introns is the presence of TACTAAC-like sequences in the vicinity of the TACTAAC box. This led Langford and Gallwitz (1983) to propose that these sequences might serve as additional primary binding sites for factors which could then diffuse along the RNA to the correct TACTAAC box. As we have just pointed out, the distinctive function of the 3' consensus signal (in contrast to the 5' consensus) is such that the inclusion of additional TACTAAC boxes might increase the rate of splicing, while not demanding

alternative 3' splice site selection.

If these TACTAAC-like sequences do play any biologically important role, the splicing machinery must able to recognize them at some level. As a first step, we have asked if a cryptic sequence in the yeast actin intron can function as a TACTAAC box. Our approach to this problem has been to construct a translational fusion which allows a simple quanitative assay for the efficiency of a splicing event. Using this fusion we have constructed a series of deletions which remove the TACTAAC box or, in addition, TACTAAC-like sequences nearby. The results of these analyses allow us to conclude that a TACTAAC-like sequence, TACTAAG, can be utilized by the yeast splicing machinery, albeit inefficiently. This demonstrates that, at least in this instance, an alternative TACTAAC box can be used. This finding sharply contrasts the failure to observe the use of cryptic 5' junctions, and suggests that TACTAAC-like sequences found in some yeast introns at least have the potential to play a physiological role.

RESULTS

CONSTRUCTION OF A SPLICED LACZ FUSION

In order to investigate the contributions of various intron features to the efficiency of the splicing process we desired a simple, quanitative assay for a splicing event. For this reason, we constructed a translational fusion between the yeast actin gene, which contains a 309 nucleotide intervening sequence (Ng and Abelson, 1980; Gallwitz and Sures, 1980), and the easily assayed E. coli beta-galactosidase structural gene. The fusion is designed such that precise and efficient splicing of the actin IVS is required for production of full levels of beta-galactosidase activity. Failure to splice should prevent translation because of numerous stop codons within the intron. Inaccurate splicing should, in most cases, prevent functional expression by altering the reading frame. Finally, inefficient splicing will be reflected as a decrease in the level of beta-galactosidase activity produced.

This fusion was constructed by a series of steps which are described in detail in the Materials and Methods. The resulting fusion (Figure 1) contains 364 base pairs of 5' flanking sequence, the RNA start site, and the translational start of the actin gene. In addition to the lacz coding sequence the chimeric gene also includes 454 codons from the yeast HIS4 gene (see Materials and Methods). We refer to this construct as the YAHB (yeast-actin-HIS4-beta-galactosidase) fusion. For

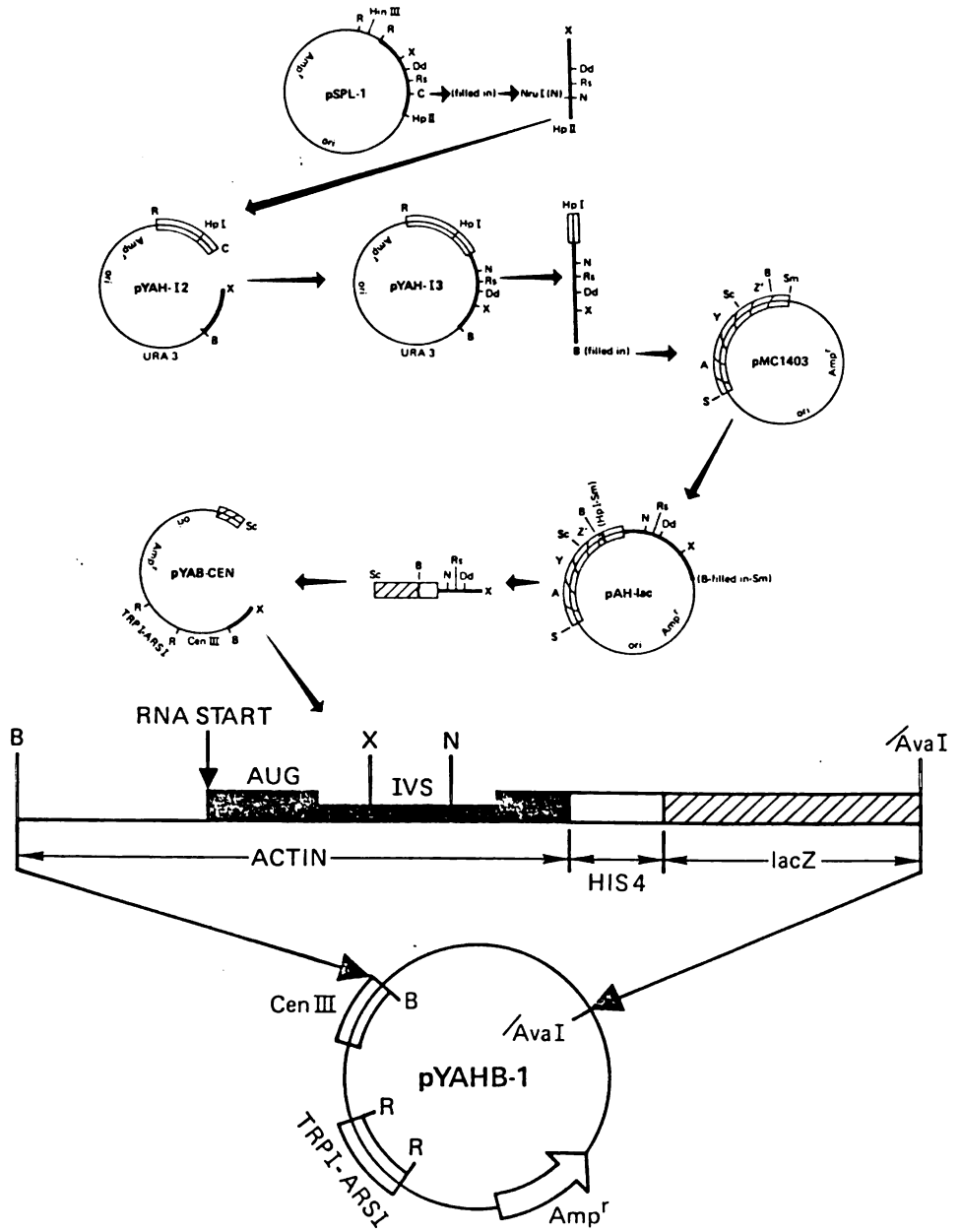

pSPL-1
pYAH-I2
URA 3
pYAH-I3
URA 3
pMC1403
pAH-lac
pYAB-CEN
RNA START
AUG
IVS
B
X
N
AvaI
ACTIN
HIS4
lacZ
Cen III
pYAHB-1
TRP1-ARS1
Amp^r

Figure 1

Schematic diagram of YAHB fusion and the yeast centromere plasmid, pYAHB-1 (not to scale); for details of the construction see Materials and Methods. The features of the fusion are shown in greater detail above the plasmid map. Transcribed regions are shown as boxes; open boxes correspond to actin sequences, slashed boxes to HIS4 sequences and shaded boxes to lacz coding sequence. The fusion presumably initiates translation at the AUG normally used in the actin gene, 10 bases 5' to the 5' intron/exon junction (Ng and Abelson, 1980; Gallwitz and Sures, 1980). In this construction there is no defined transcriptional terminator, the transcripts terminate at a number of positions within the plasmid sequences (data not shown).

all the experiments described in this paper the YAHB fusion or its derivatives are maintained in yeast on the centromere plasmid, pYAHB-1 (Figure 1). A useful feature of this plasmid is the presence of two unique restriction sites within the intron (see Figure 1). These sites, XhoI and NruI, are located such that cleavage at both sites removes the internal portion of the intron without altering the 5' and 3' portions. Thus the cassette nature of this plasmid allows us to easily remove the internal region of the intron and replace it with a modified region for analysis.

Transformation of a yeast strain, FC8-24D, with this plasmid, pYAHB-1, results in the production of 108 units of beta-galactosidase activity (Table 1). As we will show below this activity is dependent on precise and efficient splicing of the actin intron.

A NEARBY SEQUENCE CAN SUBSTITUTE FOR THE TACTAAC BOX IN THE SPLICING EVENT

The actin intron has two TACTAAC-like sequences in addition to the TACTAAC box (see Figure 2). To test the role of these sequences we constructed a series of deletions which first remove only the TACTAAC box, and then either one, or both of the TACTAAC-like sequences. These deletions were made by inserting different restriction fragments from within the actin intron into the XhoI-NruI sites of pYAHB-1 (see

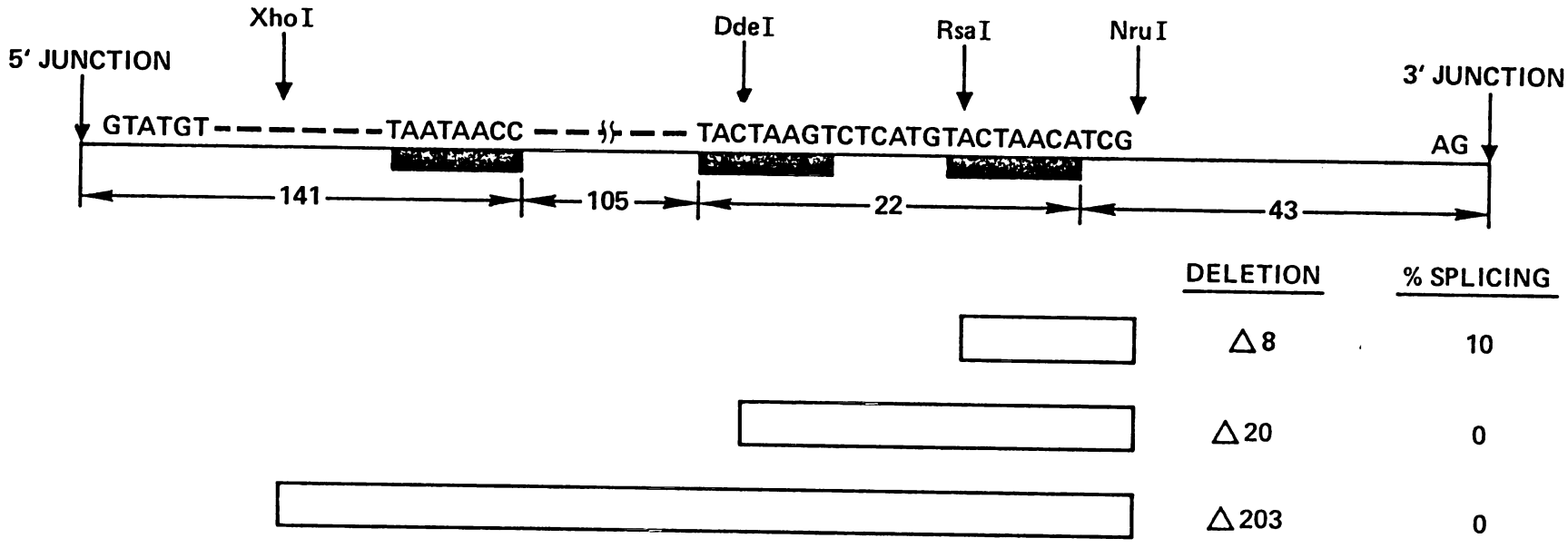

XhoI
DdeI
RsaI
NruI
5' JUNCTION
3' JUNCTION
GTATGT
TAATAACC
TACTAAGTCTCATGTACTAACATCG
AG
141
105
22
43
DELETION
% SPLICING
Δ8
10
Δ20
0
Δ203
0

Figure 2

The actin intron is shown with the relevant sequences. The 5' and 3' junctions are marked by the 5' consensus sequence, GTAPyGT, and the 3' AG. The TACTAAC box is underlined with the heavy black box; TACTAAC-like sequences are underlined with the fainter boxes. The region deleted in each construct is shown as the open box beneath the intron. Restriction sites utilizied for the construction of these deletions are shown above the intron. Distances, where appropriate, are shown beneath the intron.

Materials and Methods), and are illustrated in Figure 2. We refer to a deletion by the number of bases removed, e.g. 8 indicates eight nucleotides have been deleted; pYAHB- 8 is the corresponding plasmid name.

Plasmids carrying these deletions were transformed into the strain, FC8-24D, and biological and biochemical phenotypes were assayed. The results of beta-galactosidase assays are shown in Table 1. Surprisingly, the deletion which removed only the correct TACTAAC box, 8, still allowed the expression of approximately 10% of wild-type beta-galactosidase activity. In contrast, a slightly larger deletion, 20, which removes 12 more nucleotides 5' to the TACTAAC box, including the TACTAAC-like sequence TACTAAG, virtually abolished beta-galactosidase activity. Similarly, the removal of the correct TACTAAC box and both TACTAAC-like sequences, 203, gave only very low levels of beta-galactosidase activity.

The observation that deletion of the correct TACTAAC box did not abolish beta-galactosidase activity suggested that a low level of splicing could occur which was dependent on the sequences just 5' to the TACTAAC box. However, an alternative possibilty was that some of the beta-galactosidase activity in these constructs was not dependent on splicing. We reasoned that if the enzymatic activity, both in wild-type and in 8, was dependent on the function of the splicing machinery, then other mutations which block splicing should also prevent expression of this beta-galactosidase activity.

TABLE I. β-galactosidase measurements of yeast cells harboring intact, deleted or modified TACTAAC sequences.

VECTOR	UNITS	PERCENT pYAHB-1
pYAHB-1	108(+/-34)	100
pYAHB-Δ8	9.6(+/-6)	9
pYAHB-Δ20	0.2	0.2
pYAHB-Δ203	0.1	0.1
pYAHB-Rep	106(+/-33)	98

All measurements were carried out in transformants of the strain NNY1. For pYAHB-1, Δ8, Δ20 and Rep, the means and standard deviations from 7 independent measurements are presented. For the other construct, the value is derived from 5 independent measurements.

Rna2 is a locus in which temperature-sensitive alleles have been identified which lead to the accumulation of precursors to spliced nuclear genes at 36 C (Rosbash et al., 1981). To test if the residual beta-galactosidase activity is affected by the rna2 lesion we transformed an rna2- strain, NNY1-rna2, with both pYAHB-1 and pYAHB-8 and followed the beta-galactosidase activity after a shift to 36 C. As a critical control we also transformed the same strain with a plasmid, pYAB- IVS, which contains an actin-lacz fusion with the IVS precisely removed (Larsen et al., 1983).

As shown in Figure 3, while the activity of the fusion without the IVS is relatively unaffected, the beta-galactosidase activity from strains transformed with either pYAHB-1 or pYAHB-delta8 declines with time. The disappearance of the residual activity produced in delta8 mutant with time demonstrates that this activity is sensitive to the loss of the rna2 gene product and is thus dependent on splicing.

As an additional confirmation that the beta-galactosidase activities were an accurate reflection of the amount of correctly spliced mRNA being produced, we analyzed the fusion transcripts by primer extension from a primer located in the 3' exon. This primer has the advantage of being specific for the fusion. In addition, because the 3' exon is found in precursor and mature mRNA, this primer will produce cDNA products from both spliced and unspliced transcripts (see schematic to Figure 4).

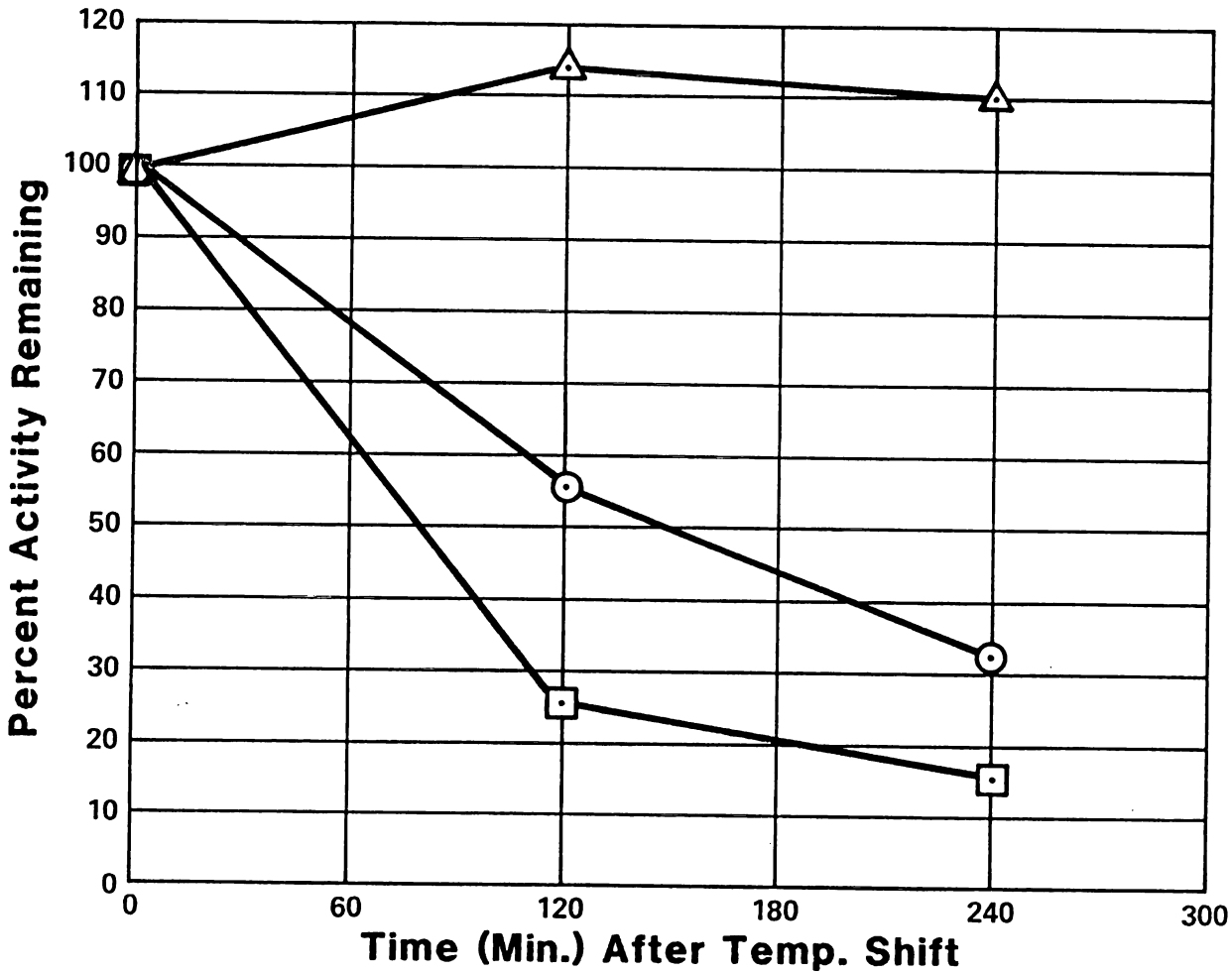

Percent Activity Remaining
120
110
100
90
80
70
60
50
40
30
20
10
0
0
60
120
180
240
300
Time (Min.) After Temp. Shift

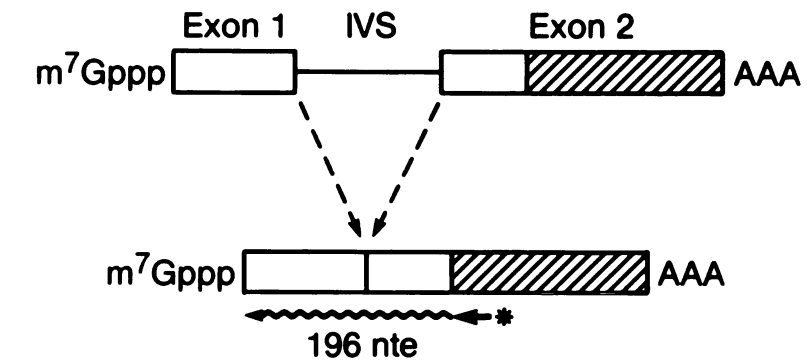
Figure 3

The strain, NNY-rna2, was transformed with pYAHB-1, pYAHB-delta8, and pYAB-deltaIVS (see text). Transformants were maintained at 23 C and then shifted to 37 C and assayed for beta-galactosidase activity at various times. Boxes, pYAHB- IVS; circles, pYAHB-1; triangles, pYAHB- 8.

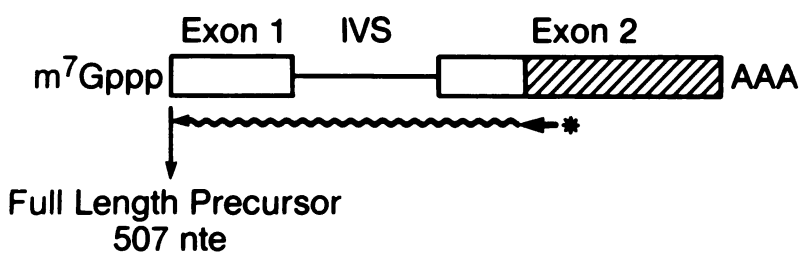
25 micrograms of total RNA from the appropriate strains was annealed to a 5' end-labeled 21 nucleotide primer. The primer was then extended with reverse transcriptase and the cDNA products seperated on a 6% sequencing gel. In strains carrying the wild-type plasmid, pYAHB-1, the major cDNA products are the correct length, 196 nucleotides, to correspond to correctly spliced mature mRNA (Figure 4, lane 2, band 1), slightly larger extension products are the result of multiple 5' ends further upstream. In the smallest deletion, 8, which removes only the TACTAAC box, a low level of spliced product is still detected (lane 3, band 1) but the predominate extension products correspond to full-length precursor (band 2). However, the larger deletions, 20 and 203, only have cDNA products corresponding to full-length precursor (lane 4, band2; and lane 5, band 2*).

We conclude from these experiments that removal of the TACTAAC box reduces splicing to approximately 10% of wild-type levels. This residual splicing is dependent on the sequences located just upstream of the TACTAAC box. The most likely explanation for this residual splicing is that the TACTAAC-like sequence, TACTAAG, in this region can substitute for the TACTAAC box in the splicing process, though at a reduced efficiency. In this model, the reduction in splicing efficiency could be because of a difference in primary sequence, TACTAAG vs. TACTAAC, and/or because the position of the TACTAAG sequence is five nucleotides further 5' from the AG at the 3' end of the intron.

Extension Products from Mature mRNA



Extension Products from Precursor mRNA



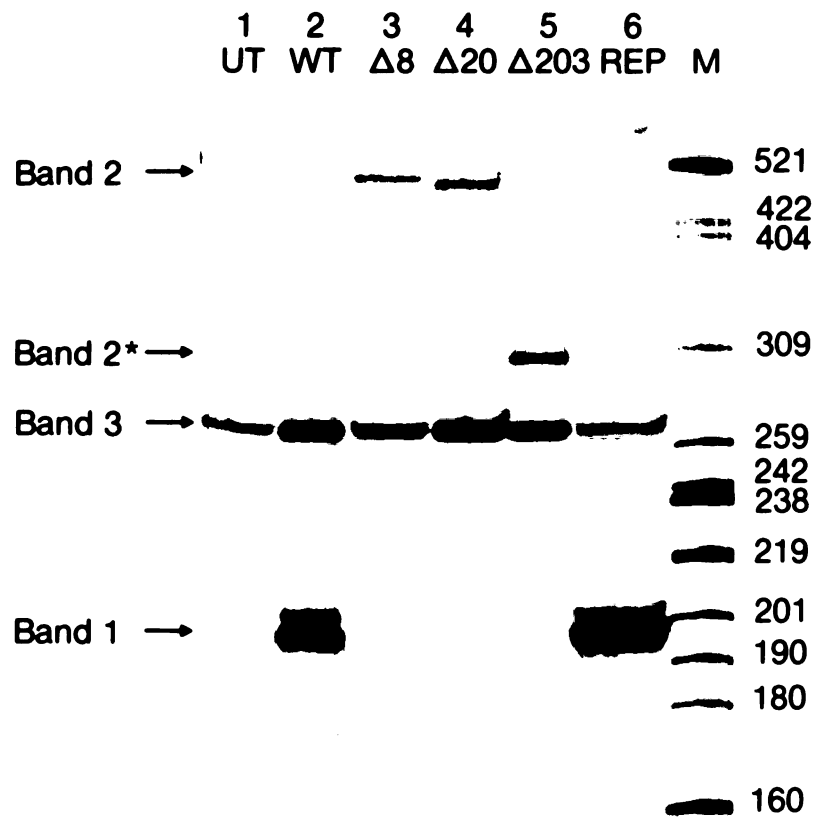


Figure 4

Total RNA from the strain, FC8-24D, transformed with the following plasmids was subjected to primer extension analysis as described (see Materials and Methods). Sizes of the cDNA products are shown schematicaly for unspliced precursor and for mature mRNA. Lane 1, untransformed parent; lane 2, pYAHB-1; lane 3, pYAHB- 8; lane 4, pYAHB- 20; lane 5, pYAHB- 203; lane 6, pYAHB-Rep. The difference in the size of the full-length precursor (band 2) results from the different sizes of the deletions in each case. In the case of 203, the size of the full-length precursor is noted as band 2*. The primer for this experiment is the 21mer, 5' TAAACGCCACGACCCAAATCGGC 3', which hybridizes specifically to the actin-HIS4 junction. Band 3 is a background band, not related to the fusion, and is also found in the untransformed parent (lane 1).

THE EFFICENCY OF TACTAAG CAN BE IMPROVED

If the residual splicing is a result of the use of the cryptic TACTAAC box, TACTAAG, then we should be able to restore splicing to an efficient rate by altering this sequence to more closely resemble the consensus. For this purpose, we inserted the DdeI(filled in)-XhoI restriction fragment from within the actin intron into the XhoI-NruI sites of pYAHB-1. The result of this construction, pYAHB-Rep, is the conversion of the cryptic TACTAAG sequence to TACTAAC (see Figure 5). In addition to this alteration in primary sequence, the new TACTAAC box is also positioned nine nucleotides closer to the 3' end of the intron than the cryptic TACTAAG sequence was (see Figure 5).

Transformation of this construct into the strain, FC8-24D, now allows full expression of beta-galactosidase activity (Table 1). In addition, analysis of the fusion transcripts reveals the same spectrum of cDNA products corresponding to mature mRNA as in wild-type, demonstrating that splicing has been restored to high efficiency (Figure 4, lane 6).

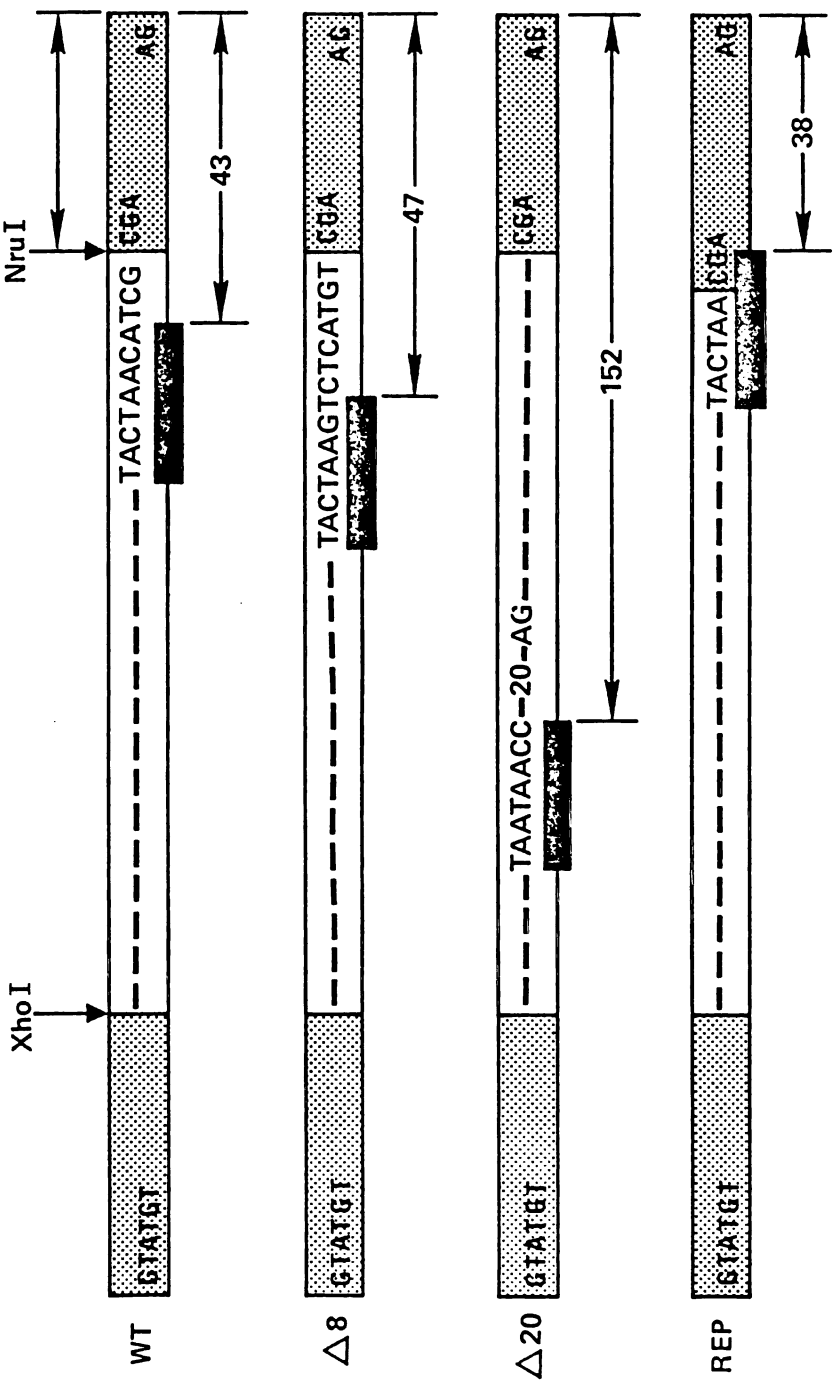

XhoI
NruI
WT
GTATGT
TACTAACATCG
CGA
AG
43
Δ8
GTATGT
TACTAAGTCTCATGT
CGA
AG
47
Δ20
GTATGT
TAATAACC–20–AG
CGA
AG
152
REP
GTATGT
TACTAA
CGA
AG
38

Figure 5

The intron is illustrated showing the positioning of the TACTAAC signal in each case within the intron. Shaded regions represent the regions outside the XhoI and NruI sites which are not altered in any of the constructs. The TACTAAC signal in each case is underlined for clarity.

DISCUSSION

In constructing the spliced beta-galactosidase fusion we had hoped to develop a simple quantitative assay for the efficiency of the removal of an intron. This hope has been realized in that not only does the fusion require splicing for expression of beta-galactosidase activity, as demonstrated by the dependence on functional rna2 gene product, but quantitative differences between constructs are paralleled in the level of mature mRNA produced. Thus, we can use beta-galactosidase activity as an initial measure of splicing efficiency in constructs designed to test the role of intron features in splicing.

While the importance of the TACTAAC box in the splicing of yeast nuclear introns is well documented (Langford and Gallwitz, 1983; Pikielny et al., 1983), the smallest deletion of the TACTAAC box which has been previously reported is 29 nucleotides (Pikielny et al., 1983); moreover, point mutations in the TACTAAC box have only been described in constructs in which the 3' end of the intron has also been altered (Langford et al., 1984). We report here the construction of an eight-nucleotide deletion (8) which precisely removes the TACTAAC box. Consistent with an important role for the TACTAAC box, this deletion reduces splicing to approximately 10% of wild-type levels. We have demonstrated that this residual splicing is dependent on a TACTAAC-like sequence located just upstream of the correct TACTAAC box. In that the

deletion of this TACTAAC-like sequence together with the TACTAAC box (20) leads to no detectable mature mRNA, we can make the rigorous conclusion that a TACTAAC box is an absolute requirement for splicing in yeast.

The first question which arises is why the cryptic signal is recognized with such poor efficiency, despite its close spatial proximity relative to the correct TACTAAC box, and its strong homology at the primary sequence level. While our experiments do not allow us to make a direct distinction between these two alternatives, it is useful to consider the most plausible contributions of each of these factors based on available information.

The importance of the spacing between the TACTAAC box and the AG (much less the potential effects of particular nucleotide sequences within this distance) has been only cursorily examined. It is quite clear that for <u>different</u> introns, the spacing varies widely, the known range being from 4 to 53 nucleotides (Teem et al., 1984). Nonetheless, it also appears that variations in the spacing/sequence for a <u>given</u> intron can significantly affect the efficiency of intron removal (Langford and Gallwitz, 1983; Cellini and Rossi, unpublished observations). In the case under examination here, however, the alteration in spacing is very subtle; the cryptic TACTAAC box is further than the wild-type from the 3' splice site by only five nucleotides (see Figure 5). Moreover, the construction which "repairs" the cryptic TACTAAC box at the primary sequence level also alters the

spacing relative to the wild-type situation by five nucleotides (in this case bringing the TACTAAC box closer to the AG). In that the latter configuration promotes wild-type splicing efficiency, we consider it most likely that the primary defect in delta eight is due to the alteration in primary sequence.

To evaluate the potential significance of an apparent requirement for a particular residue in a consensus sequence, one looks for information from genetic studies and from quantitative comparative sequence data. In Saccharomyces cerevisiae, every single position of this element shows absolute conservation (and is thus quite unilluminating in this regard). This is not the case in other fungi, however, where three of the seven positions show an equal preference for one of two alternative nucleotides. Three positions are absolutely conserved. In the final case, the terminal residue of the TACTAAC box is a C in thirteen of fifteen instances; in the two exceptions the residue is a T (Toda et al., 1984; Hiraoka et al., 1984; Woudt et al., 1983; Nunberg et al., 1984; Dons et al., 1984). Thus there would appear to be an absolute requirement for a pyrimidine at this position. In that this nucleotide occupies the position immediately adjacent to the branch site in the lariat intermediate, steric constraints due to accomodation of the 2'-5'phosphodiester linkage provide one plausible rationalization for the requirement of a pyrimidine at this location.

Genetic analyses to date have focused on alterations of other nucleotide positions. In only one instance that we are aware of has

this base been altered. Interestingly, however, it is a case in which the terminal C of the wild-type TACTAAC box in the yeast actin intron was also mutated to a G. To our surprise, splicing in this construction (a different fusion than our own, thus difficult to compare directly) occurs with close to wild-type (80-90%) efficiency (Lynette Fouser, pers. comm.). That is, this result seems in direct conflict with our analysis of the effect of this same "mutation" at the cryptic site, which allows only approximately 10% of wild-type splicing efficiency, and as we have just argued that it should if a pyrimidine is absolutely required at the position adjacent to the branch-acceptor nucleotide.

These results are only in contradiction with one another if lariat formation in the constructions with altered TACTAAC boxes still occurs at its normal position, i.e. 5' to the terminal nucleotide (in these cases, a G). In fact, in the case of 8 our preliminary data suggest the likelihood that the branch-site is six nucleotides downstream of the TACTAAG sequence, at the position of an AT dinucleotide. The same sequence (i.e., AT) occurs immediately adjacent to the normal TACTAAC box. According to this hypothesis, branch formation in each instance would occur at the closest AT 3' to the TACTAAC-like sequence TACTAAG. The dramatic difference in splicing efficiencies would then reflect two separable roles of the TACTAAC box, one in recognition by the splicing machinery, another in serving as the site of branch formation. The more inefficient splicing in 8 would derive from the larger physical distance between the two sequence elements in this instance than that demanded in the other construction. Clearly, definitive information on

the structure of the lariats in each of these cases is called for in order to evaluate this intriguing hypothesis. In the meantime, however, we note that an analogous case for the functional dissection of the 5' consensus sequence into separable requirements for recognition and for 5' endonucleolytic cleavage has recently been demonstrated (Parker and Guthrie, 1985).

The establishment that this TACTAAC-like sequence can be recognized and utilized as a TACTAAC box allows the possibilty that these cryptic elements might play some physiological role. The presence of TACTAAC-like sequences in the vicinity of the TACTAAC box is not uncommon (see Table 3). In an attempt to directly test the role of the TACTAAG sequence in the actin intron, we have construted a deletion which removes only this sequence. Preliminary results suggest that this deletion has no detectable effect on splicing as determined by levels of beta-galactosidase activity. Of course, if this TACTAAC-like sequence exerts only a subtle effect on the efficiency of the splicing process, the analysis of steady state levels of gene product would probably not be sensitive enough to distinguish this contribution. Fortunately, a senstive kinetic comparison can be made using the <u>in vitro</u> splicing system recently reported by Newman et al., (1985). These experiments should determine if these sequences do contribute to the efficiency of splicing.

TABLE II. TACTAAC-like sequences in other yeast introns.

Yeast Gene	TACTAAC-like sequence	Locations (relative to TACTAAC)
Actin (6-24)	TACTAAGT	14 bases 5'
RP51A (34)	TACAACT	22 bases 3'
S10-1 (16)	TTATAACA	19 bases 3'
L17A (15)	TGCTACT	23 bases 5'
L25 (15	TACCAACA	135 bases 5'
RP28A (20)	TTCTAAT	62 bases 5'
RP28B (20)	TACCAAT	23 bases 5'
RP29 (19)	TGTTGAC	36 bases 5'
L29 (9)	TACCACG	41 bases 5'
TACTAAC consensus	TACTAACPu	

MATERIALS AND METHODS

Enzymes and Biochemicals

Restriction endonucleases, T4 DNA ligase and DNA polymerase I(Klenow fragment) were purchased from Bethesda Research Laboratories. 32p- labeled nucleotides were purchased from either ICN or New England Nuclear. AMV reverse transcriptase was obtained from Life Sciences, Inc.. Oligonucleotides were prepared by triester synthesis on a solid support followed by purification by high performance liquid chromotography.

Strains

E. coli strain MC1061 was used for routine manipulations and was provided by M. Casadaban. Yeast strains and their sources were as follows; NNY1, MATalpha trp1 gal2 gal10 ura3-52 his3-; NNY1-rna2, MATa trp1 ura3-52 his7 lys2 tyr1 gal1 rna2 (derived from a cross of NNY1 and rna2-1 (obtained from Berkeley Stock Center); FC8-24D, MATa trp1-1 leu2 his4-401 ura3-52 HOL actin IVS (Parker and Guthrie, 1985).

Bacterial and Yeast Transformations

E. coli cells were transformed by the method of Kushner (1978). Saccharomyces cerevisiae were transformed either by the spheroplasting method using glusulase (Hinnen et al., 1978) or the lithium acetate

procedure.

Betagalactosidase Assays

Cells were grown in liquid minimal medium to an O.D. 600 of 0.5 to 1.8. 0.1 to 1.0 mls of cells were pelleted, washed in 1.0 mls of Z buffer and assayed as described previously (Larsen et al., 1983).

RNA Analysis

Preparation of RNA, kinasing of oligonucleotides, and primer extension experiments were all done following the methods as described by Domdey et al. (1985).

Construction of pYAHB-1

An outline for the construction of this vector is presented in Figure 1. To construct pYAHB, pSPL-1 was cleaved with ClaI and the overhang filled in (this modification of the intron was shown to have no effect on splicing by a number of criteria). The resulting plasmid, pSPL-2 was cleaved with XhoI and HpaII and the 287 basepair fragment containing most of the intron and part of the 3' exon was ligated into XhoI/ClaI cut pYAH-I2 (Parker and Guthrie, 1985) to yield pYAH-I3. From this plasmid the BamHI-HpaI restriction fragment, containing the entire actin region of the fusion and parts of the HIS4 gene was inserted into the SmaI site of pMC1403 to create a lacz fusion. Finally, this

plasmid, pAH-lac, was cleaved with XhoI and ScaI, and the fragment containing the fusion was inserted in to pYAB-cen to yeild the final plasmid, pYAHB-1 (Figure 1).

Deletion constructions

Deletions were constructed by first cleaving pYAHB-1 with XhoI and NruI followed by insertion of various restriction fragments purified from pSPL-1. All fragments are internal portions of the actin intron. The following fragments were inserted for the following constructs; delta8, XhoI/RsaI fragment; delta20, XhoI/DdeI (DdeI overhang removed with S1 nuclease as described (Maniatis et al., 1982); delta203, no fragment inserted; rep, XhoI/DdeI (DdeI overhang filled in with Klenow polymerase). The location of these sites is shown in Figure 2. Deletions constructions were all sequenced by labeling the DNA at the HindIII site ca. 100 base pairs downstream of the modified region, followed by cleavage at a second site and purification of the end-labeled fragment. Sequences were then determined by the chemical method of Maxam and Gilbert (1980).

REFERENCES

Abelson, J. (1979). RNA processing and the intervening sequence problem. Ann. Rev. Biochem. 48, 1035-1069.

Casadaban, N.J., Cohen, D. (1980). In vitro gene fusions that join an enzymatically active betagalactosidase segment to aminoterminal fragments of exogenous proteins: E. coli plasmind vectors for the detection and cloning of translational signals. J. Bacteriol. 143, 971-980.

Domdey, H., Apostol, B., Lin, R.J., Newman, A., Brody, E., and Abelson, J. (1984) Lariat structures are in vivo intermediates in yeast pre-mRNA splicing. Cell, in press.

Dons, J.J.M., Mulder, G.H., Rouwendal, G.J.A., Springer, J., Bremer, W., and Wessels, J.G.H. (1984). Sequence analysis of a split gene involved in fruiting from the fungus Schizophyllum commune. EMBO Journal 3:9, 2101-2106.

Gallwitz, D. (1982). Construction of a yeast actin gene intron deletion mutant that is defective in splicing and leads to the accumulation of precursor RNA in transformed cells. Proc. Nat. Acad. Sci. USA 79, 3493-3497.

Gallwitz, D., and Sures, I. (1980). Structure of a split yeast gene: complete nucleotide sequence of the actin gene in Saccharomyces cerevisiae. Proc. Nat. Acad. Sci. USA 77, 2546-2550.

Hinnen, A., Hicks, J.B., and Fink, G. R. (1978). Transformation of yeast. Proc. Nat. Acad. Sci. USA 75 , 1929-1933.

Hiraoka, Y., Toda, T., and Yanagida, M. (1984). The NDA3 gene of fission yeast encodes tubulin: a cold-sensitive nda3 mutation reversibly blocks spindle formation and chromosome movement in mitosis. Cell 39, 349-358.

Kaufer. N.F., Fried, H.M., Schwindinger, W.F., Jasin, M., Warner, J.R. (1983) Cyclohexamide resistance in yeast: the gene and its protein. Nucl. Acids Res. 11:10, 3123-3135.

Keller, E. B., and Noon, W.A. (1984). Intron splicing: a conserved internal signal in introns of animal pre-mRNAs. Proc. Nat. Acad. Sci. USA, in press.

Kushner, S.R. (1978). An improved method for the transformation of E. coli with ColE1 derived plasmids. Proceedings of the International Symposium on Genetic Engineering. Elsevier/North Holland Biomedical Press.

Langford, C.J., and Gallwitz, D. (1983). Evidence for an intron-contained sequence required for the splicing of yeast RNA polynerase II transcripts. Cell 33, 519-527.

Langford, C.J., Klinz, F-J., Donath, C. and Gallwitz, D. (1984). Point mutations identify the conserved intron-contained TACTAAC box as an essential splicing signal sequence in yeast. Cell 36, 645-653.

Larson, G.P., Itakura, K., Ito, H., Rossi, J.J. (1983). Saccharomyces cerevisiae actinE. coli lacz gene fusions: synthetic oligonucleotide mediated deletion of the 309 base pair intervening sequence in the actin gene. Gene 22, 3139.

Leer. R.J., van Raamsdonk-Duin, M.M.C., Hagendoorn, M.J.M., Mager, W.H., and Planta, R.J. (1984) Structural comparison of yeast ribosomal protein genes. Nucl. Acids Res. 12:17, 6685-6700.

Leer. R.J., van Raamsdonk-Duin, M.M.C., Molenaar, C.M.T., Cohen, L.H., Mager, W.H., and Planta, R.J. (1982) The structure of the gene coding for the phosphorylated ribosomal protein S10 in yeast. Nucl. Acids Res. 10:19, 5869-5878.

Maniatis, T., Fritsch, E.F., Sambrook, J. (1982). Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, NY. pp. 9091.

Maxam, A.M., and Gilbert, W. (1980). Sequencing end-labeled DNA with base specific cleavages. Methods Enzymol. 65, 499-560.

Mitra, G., and Warner, J.A. (1984). A yeast ribosomal protein whose intron is in the 5' leader. J. Biol. Chem. 259:14, 9218-9224.

Molenaar, C.M.T., Woudt, L.P., Jansen, A.E.M., Mager, W.H., and Planta, R.J. (1984). Structure and organization of two linked ribosomal protein genes in yeast. Nuc. Acids Res. 12:19, 7345-7358.

Mount, S., and Steitz, J. (1983). Lessons from mutant globins. Nature 303, 380-381.

Newman, A., Lin, R.J., Cheng, S-C and Abelson, J.A. In vitro mRNA splicing in yeast. Cell, submitted.

Ng, R., and Abelson, J. (1980). Isolation and sequence of the gene for actin in Saccharomyces cerevisiae. Proc. Nat. Acad. Sci. USA 77, 3912-3916.

Nunberg, J.H., Meade, J., Cole, G., Lawyer, F.C., McCabe, P., Schweickart, V., Tal, R., Wittman, V.P., Flatgaard, J.E., and Innis, M.A. (1984). Molecular cloning and characterization of the gluocoamylase gene of Aspergillus awamori. Mol. Cell. Biol. 4 2306-2315.

Padgett, R.A., Konarska, M.M., Grabowski, P.J., Hardy, S.F., and Sharp, P.A. (1984). Lariat RNAs as intermediates and products in the splicing of messenger RNA precursors. Science 225, 898-903.

Parker, R., and Guthrie, C. (1985). A point mutation in the conserved hexanucleotide at the 5' junction of the yeast actin intron uncouples recognition, cleavage, and ligation. Cell, in press.

Pikielny, C.W., Teem, J.L., and Rosbash, M. (1983). Evidence for the biochemical role of an internal sequence in yeast nuclear mRNA introns: implications for U1 RNA and metazoan splicing. Cell 34 , 395-402.

Rodriguez, J.R., Pikielny, C.W., and Rosbash, M. (1984). In vivo characterization of yeast mRNA processing intermediates. Cell, in press.

Rosbash, M., Harris, P.K., Woolford, Jr., P.K., and Teem, J. L. (1981). The effect of temperature-sensitive RNA mutants on the transcription products from cloned ribosomal protein genes of yeast. Cell 24, 679-686.

Ruskin, B., Krainer, A.R., Maniatis, T., and Green, M.R. (1984). Excision of an intact intron as a novel lariat structure during pre-mRNA splicing in vitro. Cell 38, 317-331.

Sherman, F., Fink, G., and Lawrence, C. (1979). Methods in yeast genetics. (Cold Spring Harbor, New York: Cold Spring Harbor Laboratory).

Teem, J.L., Abovich, N., Kaufer, N.F., Schwindinger, W.F., Warner, J.R., Levy, A., Woolford, J., Leer, R.J., van Raamsdonk-Duin, M.M., Mager, W.H., Planta, R.J., Schultz, L., Friessen, J.D., and Rosbash, M. (1984). A comparison of yeast ribosomal protein gene DNA sequences. Nucl. Acids. Res. 12:22, 8295-8312.

Teem, J.L. and Rosbash, M. (1983). Expression of a betagalactosidase gene containing the ribosomal protein 51 intron is sensitive to the rna2 mutation of yeast. Proc. Natl. Acad. Sci. USA 80, 44034407.

Toda, T., Adachi, Y., Hiraoka, Y., and Yanagida, M. (1984). Identification of the pleitropic cell division cycle gene NDA2 as one of two different -tubulin genes in Schizosaccharomyces pombe. Cell, 37 233-242.

Woudt, L., Pastink., A., Kempers-Veenstra, A. E., Jansen, A., Mager, W.H. and Planta, R. (1983). The coding sequence for histone H3 and H4 in Neurospora crassa are unique and contain intervening sequences. Nucl. Acids Res. 11, 5347-5360.

CHAPTER FOUR

Isolation and Characterization of Extragenic Suppressors of Cis-acting Splicing Mutations

SUMMARY

We have isolated extragenic suppressors to two cis-acting splicing mutations in an attempt to identify components of the splicing machinery. However, by focusing on suppressors which confer a temperature sensitive lethal phenotype, we appear to have restricted ourselves to a class of recessive suppressors whose gene products may be involved in the degradation of unstable transcripts. These suppressors define five complementation groups. Interestingly, the conditional lethal phenotype of these suppressors suggest that proper and efficient degradation of RNA is required for cell growth.

INTRODUCTION

The powerful yeast genetic system offers a unique opportunity to functionally define the splicing machinery by the isolation of trans-acting mutations affecting splicing. As described in the previous chapters, we have constructed a translational fusion between the yeast-actin gene and the HIS4 structural gene such that expression of the HIS4 gene product is dependent on proper splicing of the actin IVS. We have isolated or constructed mutations within this chimeric gene which inhibit the splicing of the actin intron (see Chapters 2 and 3). Though the intimate molecular details responsible for the inhibition of splicing in these mutations is not yet known, part of the mutant phenotype is likely to be the result of a decreased interaction between one or more components of the splicing machinery and the mutated recognition signal, in this case either the TACTAAC box or the 5' consensus sequence.

We have attempted to identify genes encoding the splicing machinery as those which can mutate to restore the proper interaction with the mutated intron by the selection of extragenic suppressors. This approach has proven useful in a number of systems for the identification of unknown components in a partially defined pathway (reviewed in Botstein and Maurer, 1983; Hartman and Roth, 1973). In the initial experiments, we have focused on suppressors which also confer a conditional lethal phenotype, in the expectation that normal

components of the splicing machinery will be required for cell growth. However, to our initial disappointment, the class of suppressors obtained in this manner appears to define a set of gene products involved in another aspect of RNA metabolism. In particular, these gene products appear to be involved in the degradation of unstable transcripts. Interestingly, the conditional lethal phenotype of these suppressors suggests that proper and efficient degradation of RNA is essential for cell viability.

RESULTS

Isolation of suppressors

We desired to isolate suppressors of mutations located within both regions of critical primary sequence in the intron, the TACTAAC box (Pikielny et al., 1983; Langford and Gallwitz, 1983) and the 5' consensus sequence (Gallwitz, 1982; Parker and Guthrie, 1985). In principle, this approach could allow us to identify components which interact with one or both of the splicing signals. For example, if suppressors of a point mutation in the 5' consensus sequence map to a particular locus, then the gene product of that locus is presumed to directly interact with the 5' junction. Loci which can give rise to suppressors of mutations in either the 5' consensus sequence or the TACTAAC box will define gene products which directly interact with both splicing signals, or alternatively cause a general increase in the efficiency of splicing.

For analyzing the 5' consensus sequence, we have selected suppressors to the A5 mutation, described in detail in Chapter 2. One consequence of this single base change is to reduce the production of mature RNA to approximately 30% of wild-type levels. As discussed in Chapter 2, this mutation appears to affect not only initial recognition of the intron, but in addition the 5' endonucleolytic cleavage step. In this light, the spectrum of suppressors obtained might be expected to include mutations in the 5' endonucleolytic component(s), as well as in

functions required for initial recognition. A final class of suppressors which might be obtained would be components of the splicing machinery which allow the defective lariat produced by the A5 mutation (see Chapter 2) to be sucessfully ligated into mature mRNA. This class of suppressors could conceivably define components of the 3' endonuclease (see Chapter 2).

In the case of the TACTAAC box, we have isolated suppressors to a small deletion, delta8, which precisely removes the TACTAAC box (described in Chapter 3). However, due to the presence of a TACTAAC-like sequence TACTAAG just upstream of the TACTAAC box, the deletion still produces approximately 10% of wild-type mature mRNA. In that residual splicing is dependent on the use of this TACTAAC-like sequence, we can in principle treat this small deletion as a "point mutation" within the TACTAAC box, TACTAAG. Suppressors of this mutation would identify components of the machinery which interact in some manner with the TACTAAC box.

Genetic Manipulations

The strategy for the isolation of conditional lethal suppressors from these starting mutants is operationally simple. In brief, the fusion containing either starting mutation was integrated at the HIS4 locus in the yeast strain, FC2-12B. By integrating the fusion we signifigantly lowered the frequency of gene conversion events with the resident actin gene (Szotack, 1983). For delta8, strains carrying the

fusion were plated as single colonies on YEPD plates and then replicated to Hol media (all at 25C). After 7-14 days, Hol+ papilli were picked. In order that all the mutants would be of independent origin, no more then one revertant was picked per original colony. These revertants are designated as Eλ, where λ is the number of the exact revertant.

The A5 mutation confers a slightly leaky phenotype (see Chapter 2), such that replica-plated colonies eventually exhibit low level growth. This background growth is enough to hinder the selction of mutants from replica-plated colonies. Because suspensions of yeast cultures spread on plates have a lower level of background growth, 30 single colonies were picked and grown for 2 days in YEPD liquid cultures at 25C. The cultures were then plated on Hol media and allowed to grow for 5-14 days. Revertants were picked as colonies above a level of background growth (again all manipulations were carried out at 25C). The original cultures were plated on separate plates to maintain the independence of isolates. These revertants are refered to as Tλ, again where λ is the number of the exact revertant.

In that we were particularly interested in mutations which conferred a ts lethal phenotype, revertants were picked to master plates at 25C and tested for the ability to grow on YEPD at 30 and 36 C. For each starting mutation, six independent revertants were obtained which also had a ts lethal phenotype. As shown in Table 1, these suppressors varied both in their degree of suppression, as determined by growth on

TEMPERATURE	25			30			36		
ISOLATE	YEPD	-URA	HOL	YEPD	-URA	HOL	YEPD	-URA	HOL
E46	+	+	+	+	+	+	-	-	-
E87	++/-	++/-	-/+	+/-	+/-	-/+	-	-	-
E124	+	+	++/-	+	+	++/-	-	-	-
E170	+	+	++/-	+	+	++/-	-/+	-	-
E183	++/-	++/-	+/-	+/-	+/-	+/-	-/+	-	-
E202	+	+	+/-	+	++/-	++/-	-/+	-	-
T6	+	+/-	+/-	+	+	++/-	-	nt	nt
T19	+	+/-	+	+	+	+	-	nt	nt
T20	+	+	+	+	+	+	-	nt	nt
T26	+	+	+	+	+	+	-	nt	nt
T27	+	+/-	+	+	++/-	+	-	nt	nt
T211	+	+	++/-	+	+	+/-	-	nt	nt
controls									
FC2-12B.WT	+	+	+	+	+	+	+	+	+
FC2-12B.A5	+	+	-	+	+	-	+	+	-
FC2-12B. 8	+	+	-	+	+	-	+	+	-
FC2-12B	+	-	-	+	-	-	+	-	-

TABLE 1 Phenotypes of TS Lethal Suppressors

The table shows the results of scoring the ts lethal suppressors at three temperatures. A + indicates growth, +/- indicates intermediate growth, and a - indicates failure to grow. Each suppressor is the purified original isolate, controls are FC2-12B transformed with various forms of the fusion. All plates were scored at 24 hours after replica-plating except the 25C Hol plate which was scored after 48 hours. nt stands for not tested.

Hol media, and in their sensitivity to high temperatures.

TS LETHALITY AND SUPPRESSION CO-SEGREGATE

In order to demonstrate that the lethality at high temperatures and the suppressor phenotype were caused by the same mutation the revertants were outcrossed to a strain, FC2-16D (see strain table for full genotype), and the segregation of the suppressor and the ts lethality analyzed by tetrad analysis. The results of these dissections are shown in detail in the appendix at the end of this chapter. In summary, this analysis demonstrated that the conditional lethality and the suppressor phenotype in all the revertants tested are the result of a single or two closely linked mutation(s). Outcrosses of the E202 suppressor gave essentially zero spore viability and were therefore resistant to this analysis.

In addition, in two suppressors, E46 and T25, a secondary cis-acting reversion event linked to the fusion was uncovered. The nature of the linked reversion in E46 is a gene conversion event from the resident actin gene which repairs the eight base deletion (see Appendix). Surprisingly, the E46 suppressor also increases the level of expression of this "wild-type" fusion (see below). In T25 the cis-acting suppressors has a slightly weaker Hol+ phenotype than the wild-type fusion, suggesting that there is a linked psuedo-reversion event. Finally, the E87 suppressor is a nuclear petite.

THE SUPPRESSORS DEFINE FIVE COMPLEMENTATION GROUPS

In order to determine the number of genes identified in this selection the suppressors were sorted into complementation groups based on the ts lethal phenotype. Segregants from the outcross of FC2-16D with the suppressors and of the MATa genotype were backcrossed to the original isolates by cross-stamping at 25C. The plates were grown at 25 C for 24 hours to allow mating and then replica-plated to YEPD and incubated at 36 C. The diploids were then scored for growth after 24-48 hours. The results of this analysis are shown in Figure 1a.

In summary, the mutants define five complementation groups. Three of the groups are defined by suppressors of the delta8 mutation (groups I, II, and III). One complementation group (group IV) consists of suppressors isolated from both delta8 and from A5, while the final group (group V) is defined by only a single suppressor of A5 (see Figure 1b).

At this point we concluded that we had identified phenotypic suppressors of two splicing mutations. These suppressors also confer a ts lethal phenotype at 36 C and define 5 complementation groups. The critical question in the understanding of these suppressors is the mechanism for the suppression of the Hol- phenotype. We have approached this in two ways. First, we have asked genetically if the suppressors have any effect on an actin-lacz fusion which is not defective for splicing; i.e. are the suppressors affecting another aspect of

A)

	E46	E87	E124	E170	E183	E202	T6	T19	T20	T26	T27	T211
E46	-											
E87	+	-										
E124	+	+	-									
E170	+	+	+	-								
E183	+	+	-	+	-							
E202	-	+	+	+	+	nt						
T6	+	+	-	+	-	+	-					
T19	+	+	-	+	-	+	-	-				
T20	+	+	-	+	-	+	-	-	-			
T26	+	+	-	+	-	+	-	-	-	-		
T27	+	+	-	+	-	+	-	-	-	-	-	
T211	+	+	+	+	+	+	+	+	+	+	+	-

B)

```
:::::[GTATGT--------------------------TACTAAC------AG]::::::::
          *                               *
          A                               G
```

A5 Suppressors	delta8 Suppressors	Complementation Group
	E46, E202	I
	E87	II
	E170	III
T6,T19,T20,T26,T27	E124, E183	IV
T211		V

FIGURE 1 Complementation Analysis of TS Lethal Suppressors

A. The figure shows the scoring of a cross stamped grid of supprssors of MATa and MATalpha mating type. The suppressors with MATalpha mating type are the original isolates. The MATa testers are segregants from the outcrossing of each suppressor (see Apendix). These strains were as follows; E46, EC1-1A; E87, EC3-1C; E124, EC4-6C: E170, EC5-3B; E183, EC6-9C; T211, TC1-2B; T6, TC2-1B; T19, TC3-3A; T20, TC4-13B; T26, TC5-9B; T27, TC6-4D. A segregant of the E202 suppressor was not obtained because of the zero spore viability of this revertant (see text).

B. A summary of the results of the complementation analysis.

expression of the fusion which is not related to splicing?. Second, we have asked if the suppressors have a detectable effect of the splicing or metabolism of the fusion transcripts, by directly analyzing the RNA produced.

THE SUPPRESSORS ARE GENERAL AND RECESSIVE IN NATURE

To determine the effect of the suppressors on a gene without a splicing defect, each suppressor was crossed to a strain transformed with pYAHB-1 (see Chapter 3). This plasmid carries an actin-lacz fusion which is spliced at wild-type levels. The resulting diploids were first examined to determine if the suppressors were dominant by assaying beta-galactosidase activity on plates containing the chromatic indicator X-Gal. In no case did the suppressor increase the level of beta-galactosidase activity over the diploid control. We conclude from this that the suppressors either have no effect on this construct or are recessive in nature.

To distinguish between these two possibilities we dissected the diploids to obtain haploid segregants carrying the suppressor and the pYAHB-1 plasmid. Segregants carrying the fusion, with and without the suppressor, were then assayed for beta-galactosidase activity on plates containing the chromatic indicator, X-Gal. In each case tested the presence of the suppressor strikingly increased the level of beta-galactosidase activity. We conclude from this that at least the suppressors in complementation groups I, II, and IV are recessive and

are affecting some aspect of expression of the fusion gene product which is independent of the requirement for splicing. Unfortunately, the single isolates in groups III and V proved difficult to examine because of plasmid instability and poor spore viability. However, the identical behavior of these groups with the other classes in all other tests (see below) leads us to the working hypothesis that the mechanism of suppression in these two groups is similar to the other three groups.

THE SUPPRESSORS ACCUMULATE MORE FUSION TRANSCRIPTS

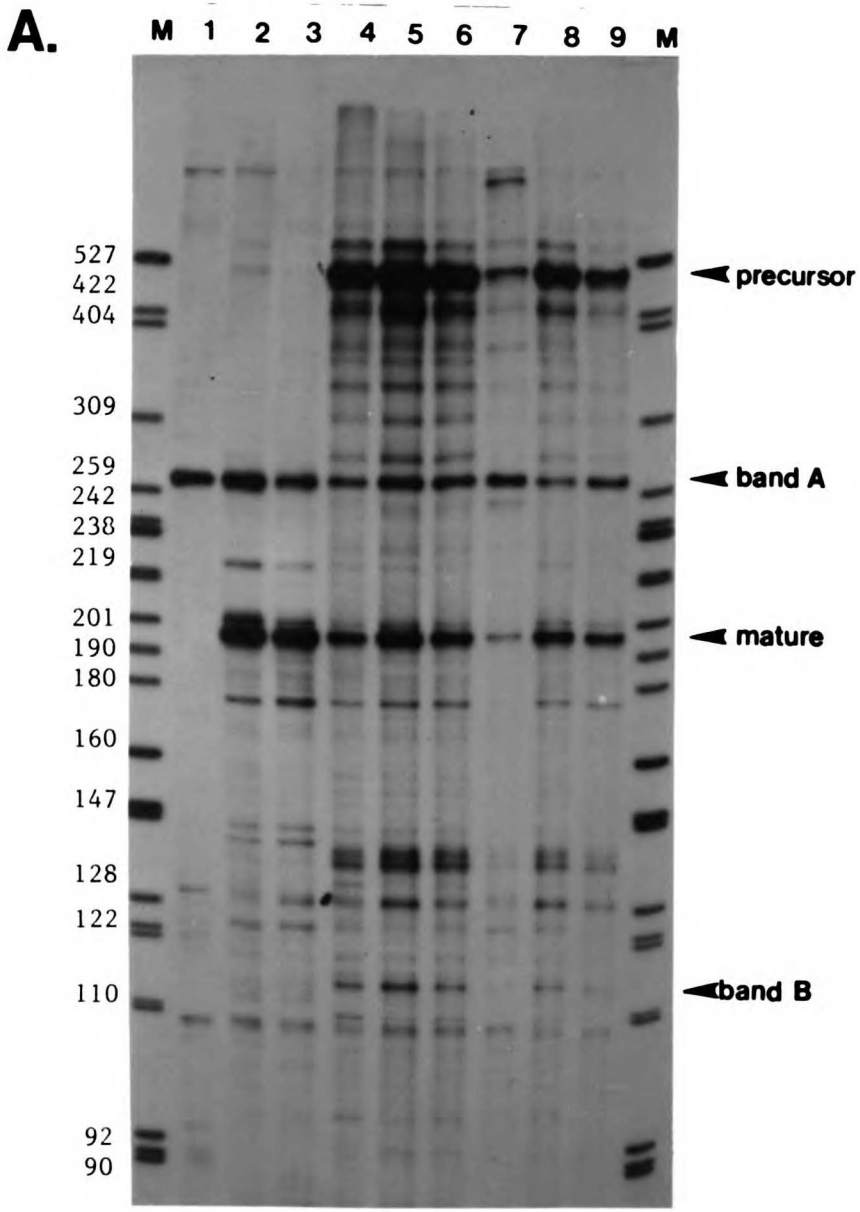
Though the suppressors are not affecting splicing directly, there are other classes of mutations which we might expect to isolate in this selection which would be of interest. For example, as discussed in Appendix 1 (Chapter 1) the fusion transcript appears to be inherently unstable. Mutations which stabilize the fusion transcripts might then be expected to give a suppressor phenotype. In that RNA stability is beginning to rear its ugly head as a possible site for regulation of gene expression (for example, see Wilson and Darnell, 1981) these mutants would not be uninteresting.

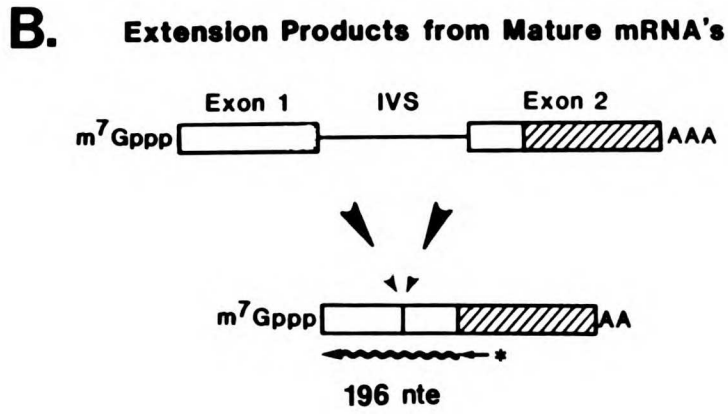
To determine if the suppressors were affecting any aspect of the expression of the fusion at the RNA level we examined the fusion transcripts in these strains. Strains were grown at 25C in YEPD liquid cultures. RNA was prepared as described previously and run over polyA+ columns. The resulting RNA was then analyzed by primer extension from a

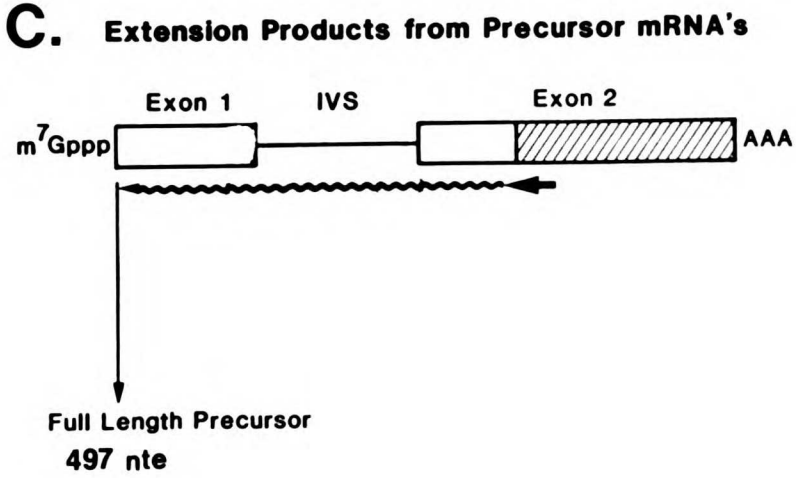
primer in the 3' exon (see Experimental Procedures, Chapter 2, for experimental details). This primer has the advantage of being specific for the fusion transcripts. In addition, because the 3' exon is found in both precursor and mature species the primer will produce cDNA products from both spliced and unspliced transcripts (see schematic to Figures 2 and 3).

The results for the E series of suppressors are shown in Figure 2. In the unmutated fusion (lane 2), the predominant cDNA products correspond to the correct length to the products from mature mRNA, 196 nucleotides. In the starting mutant, delta8, approximately 10% of the cDNA products correspond to mature mRNA (lane 9), while the remainder are the correct length (497 nucleotides) to be derived from full-length precursor (lane 9, precursor band). In all lanes, including the untransformed parent, there is an additional, fortuitous cDNA extension product (band A). This product, while not related to the fusion, provides an internal control for the amount of RNA in each lane.

The suppressors give slightly different results with the general phenotype described below. For suppressors in group I, E46 and E202, there seems to be an increase in the amount of fusion transcripts relative to the control. This is particularly clear in E202 (compare precursor and mature,lane 8 with lane 9), and appears to predominantly affect the precursor. The group II mutant, E87, also appears to accumulate additional fusion transcripts (lane 4). In this mutant the effect again predominantly affects precursor species. This would


A.
M 1 2 3 4 5 6 7 8 9 M
527
422
404
309
259
242
238
219
201
190
180
160
147
128
122
110
92
90
precursor
band A
mature
band B


B.
Extension Products from Mature mRNA's
Exon 1
IVS
Exon 2
m7Gppp
AAA
m7Gppp
AA
196 nte


C.
Extension Products from Precursor mRNA's
Exon 1
IVS
Exon 2
m7Gppp
AAA
Full Length Precursor
497 nte

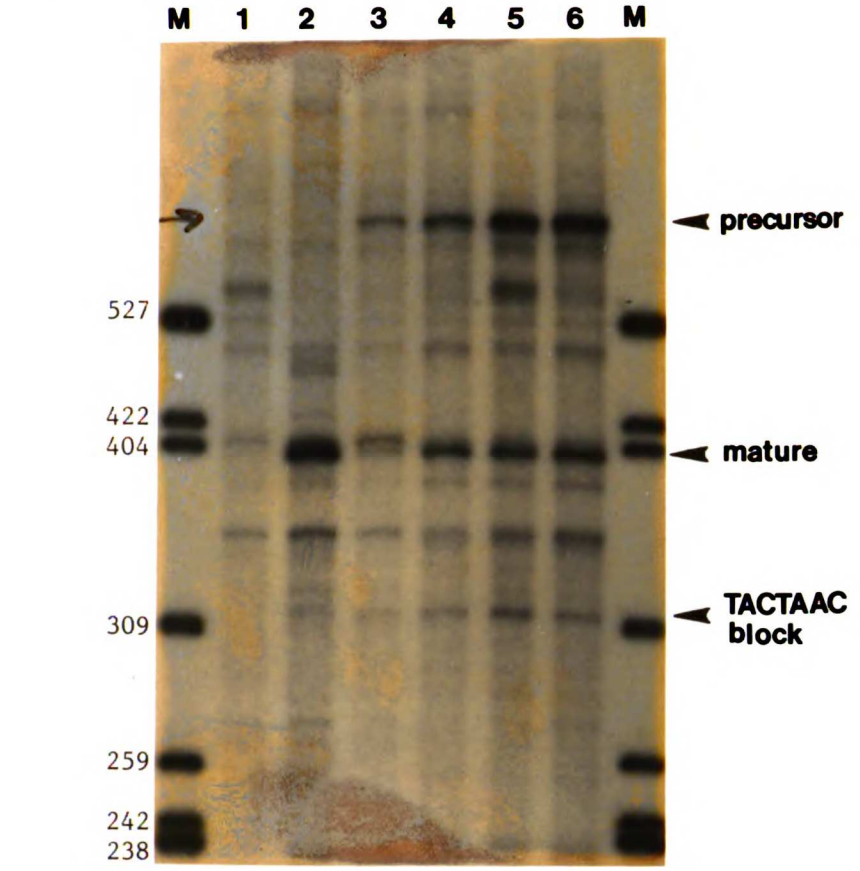

A.
M 1 2 3 4 5 6 M
527
422
404
309
259
242
238
precursor
mature
TACTAAC block
B.
Extension Products from Mature mRNA's
Exon 1
IVS
Exon 2
m7Gppp
AAA
m7Gppp
AAA
*
(398nte)
C.
Extension Products from Precursor mRNA's
Exon 1
IVS
Exon 2
m7Gppp
AAA
3' Cleavage (269nte)
Block at TACTAAC (311nte)
Aberrant 5' Cleavage (584nte)
Full Length Precursor (707nte)

FIGURE 2 Primer Extension Analysis of delta8 Suppressors

A. 5 micrograms of polyA+ RNA was extended as described in the text from the primer located at the fusion junction. After extension with reverse transcriptase the cDNA products were electrophoresed on a 6% acrylamide 8M Urea sequencing gel. Lane 1, untransformed parent FC2-12B; lane 2, wild-type fusion; lane 3, wild-type fusion with the E46 suppressor; lane 4, E87; lane 5, E124; lane 6, E170; lane 7, E183; lane 8, E202; lane 9, delta8 fusion in starting strain. Strains labeled M are kinased HpaII cut pBR328 for size markers. Band A is a band unrelated to the fusion. Band B corresponds to a primer extension block at the cryptic TACTAAC box, TACTAAG.

B. Extension products predicted from mature mRNAs.

C. Extension products predicted from precursor mRNA.

FIGURE 3 Primer Extension Analysis of A5 Suppressors

A. Results of primer extension of 5 micrograms of polyA+ RNA from a primer located in the 3' exon. Lane 1, untransfromed parent strain, FC2-12B; lane 2, wild-type fusion; lane 3, A5 mutant fusion; lane 4, T19; lane 5, T20; lane 6, T211.

B. Extension products predicted from mature mRNAs.

C. Extension products predicted from precursor mRNA.

correlate well with the observed weak suppressor behavior of this mutant (see Table 1). Again, the group III mutant, E170 (lane 6), and the group IV mutant, E124, (lane 5) appear to accumulate additional fusion transcripts. The group IV suppressor, E183, is clearly complicated. Additional analysis (data not shown) raises the possibility that in this mutant there is an addtional transcriptional start 5' of the normal gene and a secondary alteration in the splicing pattern as well. At the present this is not understood.

As shown in Figure 3, a few additional suppressors of group IV, isolated as A5 suppressors, and the lone suppressor in group V were examined. While not as dramatic, the results are similar; the suppressors appear to accumulate higher levels of fusion transcripts. These results suggest that the mechanism of suppression in these mutants is by stabilizing the unstable fusion transcripts. An alternative possibility, that the suppressors are increasing levels of transcription is still formally possible, but considered unlikely for the reasons discussed below.

DISCUSSION

In this project we set out to attempt to identify the components of the splicing machinery which interact with the critical primary sequence signals within the intron. However, due to what appears to be a low stability of the fusion transcripts, the primary class of suppressors derived appear to exert their affect by increasing the amount of the fusion transcripts.

This affect could in principle be a result of either increased transcription or a stabilization of fusion transcripts. For the reasons discussed in Appendix 1, Chapter 1, the fusion appears to produce an inherently unstable transcript. Thus, while we cannot formally rule out that the effect is on the level of transcription (experiments to test this are now in progress), the most likely explanation consistent with the genetic and biochemical data is that the suppressors are stabilizing the fusion transcripts. A piece of data which supports this argument is that the suppressors do <u>not</u> appear to increase the level of expression of the resident actin gene (as judged by Northern analysis, data not shown), a gene which has the same promoter elements.

Extragenic mutations which stabilize unstable mutant transcripts from the <u>cyc1</u> gene in yeast have been identified in the <u>sut2</u> locus (Zaret and Sherman, 1984). It would be of great interest to determine if the suppressors identified in this screen are allelic to the <u>sut2</u>

suppressors. However, the difficulty in scoring the sut2 suppressor -- slight differences in strain backgrounds can cause wide variations in the phenotype (K. Zaret, pers. comm.)-- has prevented this experiment.

If the suppressors are acting by increasing the stability of unstable transcripts, then this stabilization could be exerted at two different levels. Inactivation of a nuclear degradation pathway would lead to the stabilization of precursor species. The increased levels of precursor could then, by mass action, produce more mature mRNA. Alternatively, stabilization of cytoplasmic mRNAs could in principle lead to increased levels of mature mRNA while not affecting the levels of precursor species. The results of the primer extensions tend to suggest that the effect is most pronounced at the level of precursor (see Figure 2). This would argue that the suppressors are at least stabilizing precursor species, if not both. Thus, the molecular basis of the suppressors may be to inactivate or decrease the efficiency of a nuclear degradation pathway.

In this light, it is not immediately clear why the suppressors should increase the expression of genes without a splicing defect. For the lacz fusion the answer may lie in the observation that these constructs, because they lack a defined yeast terminator, terminate at several places within the plasmid sequences (Parker, unpublished observations). This undefined termination phenomenon may lead to the production of an array of unstable transcripts, which are then stabilized by the suppressors (see Zaret and Sherman, 1984).

Alternatively, the suppressors may affect both nuclear and cytoplasmic transcripts. Or finally, and perhaps most provocative, the fusion transcripts may be inherently unstable because they are defective in exiting the nucleus and are more subject to nuclear degradation. This may explain why the actin-HIS4 fusion, which does have a defined terminator appears to produce an unstable transcript.

The speculation that the suppressors are involved in stability of nuclear transcripts raises an interesting question, why do the suppressors confer a conditional lethal phentotype? By the simplest interpretation (though certainly not the only one), at low temperatures the suppressors are partially inactivated, leading to the suppressor phenotype, while at high temperatures, the gene product of the suppressor is now completely nonfunctional (recall the suppressors are all recessive). This would suggest that for cell growth, proper and efficient degradation of some class(es) of normal transcripts is required. Preliminary experiments examining the suppressors at high temperatures have not revealed any striking alterations in the distribution of transcripts from the fusion or the actin gene. However, the experiments were not done in a way to determine if stability of particular transcipts was affected.

The suppressors in the selection defined five complementation groups, all of which have a similar phenotype. There are two possible explanations for this; in the first, each complementation group represents a single nuclease, the inactivation of any one of which

leads to the stabilization of transcripts. This seems unlikely in that if there are at least five degradation enzymes, why would the loss of one lead to the stabilization phenotype. An alternative possibilty is that these gene products may all be part of a degradation complex, within which the loss of any one component will lead to inactivation. One prediction from this is that double mutants containing two suppressors will have the same phenotype as the single suppressors.

In summary, in this mutant hunt we have failed to isolate the class of suppressors we set out to find, those defining the genes of the splicing machinery. Instead, we have found a class of genes which appear to have an affect on the stability of the fusion transcripts. In that these suppressors confer a recessive ts lethal phenotype we speculate that proper and efficient degradation of transcripts is required for viability. In addition to opening our eyes to another feature of RNA metabolsim within the cell, this information has been valuable in designing additional selections to isolate the class of splicing suppressors originally desired.

STRAIN LIST

Strain	Genotype	Source
FC2-12B	MATalpha ura3-52 his4-401 leu2 trp1 HOL1 can1	FC1-6BxTD77
FC2-16D	MATa ura3-52 his4-401 leu2 ade5 HOL1 sac1	FC1-6BxTD77
EC1-1A	MATa ura3-52 his4-401 leu2 trp1 HOL1 E46 pYAHS-I(delta8) at HIS4	E46xFC2-16D
EC3-1C	MATa ura3-52 his4-401 leu2 trp1 ade5 HOL1 E87 pYAHS-I(delta8) at HIS4	E87xFC2-16D
EC4-6C	MATa ura3-52 his4-401 leu2 E124 can1 HOL1 pYAHS-I(delta8) at HIS4	E124xFC2-16D
EC5-3B	MATa ura3-52 his4-401 leu2 HOL1 sac1 E170	E170xFC2-16D

EC6-9C	MATa ura3-52 his4-401 trp1 ade5 HOL1 E183 pYAHS-I(delta8) at HIS4	E183xFC2-16D
TC1-2B	MATa ura3-52 his4-401 leu2 trp1 T211 HOL1 pYAH-I2(A5) integrated at HIS4	T211xFC2-16D
TC2-1B	MATa ura3-52 his4-401 leu2 trp1 ade5 can1 HOL1 T6 pYAH-I2(A5) integrated at HIS4	T6xFC2-16D
TC3-3A	MATa ura3-52 his4-401 leu2 trp1 ade5 sac1 HOL1 pYAH-I2(A5) integrated at HIS4	T19xFC2-16D
TC4-13B	MATa ura3-52 his4-401 leu2 trp1 ade5 sac1 HOL1 T20 pYAH-I2(A5) integrated at HIS4	T20xFC2-16D

TC5-9B	MATa ura3-52 his4-401 leu2 trp1 ade5 sac1 HOL1 T26 pYAH-I2(A5) integrated at HIS4	T26xFC2-16D
TC6-4D	MATa ura3-52 his4-401 leu2 trp1 ade5 can1 sac1 HOL1 T27	T27xFC2-16D
NNY1	MATalpha ura3-52 trp1 his3	J.Rossi (City of Hope)

The following strains are a collection of the RNA mutants obtained from D. J. Hurt in Anita Hoppers laboratory. I have given his name and the genotype as I know it.

Strain	Genotype
rna1-2B	MATa ura3-52 leu2
rna1-3B	MATalpha ura3-52 leu2
368alpha2-5	MATalpha rna2

125alpha2-93	MATalpha rna3
339alpha2-1	MATalpha rna4
J08alpha2-57	MATalpha rna5
202alpha2-67	MATalpha rna7
219alpha2-85	MATalpha rna8
257alpha2-82	MATalpha rna9
382alpha2-23	MATalpha rna11

REFERENCES

Abelson, J. (1979). RNA processing and the intervening sequence problem. Ann. Rev. Biochem. 48, 1035-1069.

Botstein, D., and Maurer, R. (1982). Genetic approaches to the analysis of microbial development. Annu. Rev. Genet. 16 61-83.

Gallwitz, D. (1982). Construction of a yeast actin gene intron deletion mutant that is defective in splicing and leads to the accumulation of precursor RNA in transformed cells. Proc. Nat. Acad. Sci. USA 79, 3493-3497.

Hartman, P.E., and J.R. Roth. (1973). Mechanisms of suppression. Adv. Genet. 17, 1-105.

Langford, C., and Gallwitz, D. (1983). Evidence for an intron-contained sequence required for the splicing of yeast RNA polymerase II transcripts. Cell 33, 519-527.

Pikielny, C.W., Teem, J.L., and Rosbash, M. (1983). Evidence for the biochemical role of an internal sequence in yeast nuclear mRNA introns: implications for U1 RNA and metazoan splicing. Cell 34 , 395-402.

Szostak, J.W. (1983). Replication and resolution of teleomeres in yeast. Cold Spring Harbor Symp. Quant. Biol. 47, 1187-1194.

Teem, J.L., Abovich, N., Kaufer, N.F., Schwindinger, W.F., Warner, J.R., Levy, A., Woolford, J., Leer, R.J., van Raamsdonk-Duin, M.M., Mager, W.H., Planta, R.J., Schultz, L., Friessen, J.D., and Rosbash, M. (1984). A comparison of yeast ribosomal protein gene DNA sequences. Nucl. Acids. Res. 12:22, 8295-8312.

Wilson, M.C., and Darnell, J.E. (1981) Control of messenger RNA concentration by differential ctyoplasmic half-life. J. Mol. Biol. 148, 231-251.

Zaret, K.S., and Sherman, F. (1984). Mutationally altered 3' ends of the yeast CYC1 mRNA affect transcript stability and translational efficiency. J. Mol. Biol. 177, 107-135.

APPENDIX 3

Linkage of TS Lethality and Suppressor Phenotypes

Crosses

The following is a list of the outcrosses of the revertants obtained from A5 and delta8. Crosses of A5 revertants are referred to as TCx (T series cross), where x is the exact cross number. Similarily, outcrosses of the delta8 revertants are numbered ECx.

Cross	Parent 1	Parent 2
EC1	E46 isolate	FC2-12B (MATa)
EC3	E87 isolate	FC2-16D
EC4	E124 "	FC2-16D
EC5	E170 "	FC2-16D
EC6	E183 "	FC2-16D
TC1	T211 "	FC2-16D
TC2	T6 "	FC2-16D
TC3	T19 "	FC2-16D
TC4	T20 "	FC2-16D
TC5	T26 "	FC2-16D
TC6	T27 "	FC2-16D

The pages that follow give the scoring of the segregants from these outcrosses. The segregation of the suppressor and the ts lethal phenotype are summarized for each individual revertant following the cross. In addition, when appropriate, comments follow immediately after the relevant cross. Unfortunately, in some cases, poor sporulation and spore viability limited the number of tetrads scored.

EC1

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
1a	+	-	-	-
1b	+	-	+	+
1c	+	+	+	++/-
2a	+	-	+	+
2b	+	-	-	-
2c	+	+	-	-
3a	+	-	+	+
3b	+	+	-	-
3c	+	-	-	-
3d	+	+	+	++/-
4a	+	-	-	-
4b	+	+	-	-
4c	+	-	-	-
5a	+	+	+	++/-
5b	+	-	-	-
5c	+	-	+	+
6a	+	+	+	++/-
6b	+	-	-	-
7a	+	+	+	++/-
7b	+	-	+	+
7c	+	-	-	-

Summary of Fusion Containing Segregants

	Total	Hol+	Hol+/-	Hol-
ts lethal	5	5	0	0
wild-type	5	0	5	0

Comments

Careful comparison of the ts lethal segregants and the temperature resistant segregants showed that the ts lethal spores grow faster on Hol media. However, the observation that every fusion-containing segregant can now grow on Hol to some extent indicates that there has been a reversion event linked to the fusion. Southern analysis of the E46 revertant has shown that a ClaI restriction site, deleted in the

delta8 deletion, has been restored. In this light, the cis-acting reversion event is likely to be the result of a gene conversion event between the resident actin gene and the fusion. It is important to note that though the fusion now contains and wild-type intron, and accordingly splices at high efficiency (see Figure 3, lane 3) the Hol phenotype is still increased by the ts lethal suppressor. This indicates that this suppressor's mechanism of suppression is not at the level of splicing.

A final note, in this cross I have used as the MATa parent a strain derived by transformation of FC2-12B with a centromere plasmid followed by loss of the plasmid and dissection of the resulting diploids. The result of this manipulation is a perfectly isogenic strain, ideal for outcrosses of revertants. However, poor sporulation and spore viability of this isogenic cross is a hindrance to analysis. For this reason the remainder of the outcrosses were done using a related, but not perfectly isogenic strain, FC2-16D.

EC3

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
1a	+	+	-	-
1b	+	+	-	-
1c	+	-	+	++/-
2a	+	+	-	-
2b	+	+	-	-
3a	+	-	-	-
3b	+	+	-	-
3c	+	+	+	-
3d	+	-	+	+
4a	+	+	-	-
5a	+	+	+	-
5b	+	-	+	+/-
6a	+	-	-	-
6b	+	+	-	-
7a	+	+	+	-
7b	+	+	-	-
8a	+	-	-	-
9a	+	+	-	-
9b	+	+	-	-
10a	+	-	+	+
11a	+	+	+	-
11b	+	+	-	-
12a	+	+	+	-
12b	+	+	+	-
12c	+	+	-	-
12d	+	+	-	-
13a	+	+	+	-
13b	+	+	-	-
14a	+	+	+	-
14b	+	+	+	-
15a	+	+	-	-
15b	+	+	-	-

EC3 (continued)

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
16a	+	+	+	-
16b	+	-	-	-
16c	+	+	-	-
17a	+	+	+	-
17b	+	+	-	-
18a	+	+	-	-
18b	+	+	+	-
19a	+	+	-	-
19b	-/+	-	-	-
19c	+	+	-	-
19d	-/+	-	-/+	+/-
20a	+	+	-	-

Summary of Fusion-Containing Segregants

	Total	Hol+	Hol+/-	Hol-
ts lethal	4	2	2	0
wild-type	12	0	0	12

EC4

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
1a	+	+	+	-
1b	+	-	+	-
2a	+	-	+	-
2b	+	-	+	-
3a	+	+	+	-
3b	+	-	+	+
3c	+	-	-	-
4a	+	-	-	-
4b	+	+	-	-
4c	+	+	+	-
4d	+	-	+	-
5a	+	+	-	-
5b	+	+	-	-
6a	+	+	-	-
6b	+	+	+	-
6c	+	-	+	+
6d	+	-	-	-
7a	+	-	+	+
7b	+	-	+	+
8a	+	-	+	-
9a	+	-	-	-

Summary of Fusion-Containing Segregants

	Total	Hol+	Hol+/-	Hol-
ts lethal	9	4	0	5
wild-type	4	0	0	4

Comments

The results of this outcross indicated that while the ts lethal mutation was required for growth on Hol media, only half of the expected spores grew on Hol. This suggests that there is a second unlinked allele segregating which is required for suppression. A second allele in this complementation group also has this same behavior, E183,

cross EC6. However, alleles in this locus which are obtained by selection of suppressors of the A5 mutation (T6,T19,T20,T26, and T27) do not appear to have a requirement for an unlinked modifier (see cross TC2 for example).

EC5

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
1a	+	-	-	-
1b	+	+	-	-
1c	+	+	+	-
1d	+	-	+	+
2a	+	+	+	-
2b	+	-	+	+
2c	+	+	-	-
3a	+	-	+	+
3b	+	-	-	-
3c	+	+	+	-
3d	+	+	-	-
4a	+	-	-	-
4b	+	-	+	+
4c	+	+	+	-
4d	+	+	-	-
5a	+	+	-	-
5b	+	+	+	-
5c	+	-	-	-
5d	+	-	+	+
6a	+	-	-	-
6b	+	-	+	+
6c	+	+	+	-
6d	+	+	-	-
7a	+	+	+	-
7b	+	-	-	-
7c	+	-	+	+
8a	+	+	-	-
8b	+	-	-	-
8c	+	+	+	-
8d	+	-	+	+
9a	+	-	-	-
9b	+	-	-	-
9c	+	+	-	-
9d	+	+	+	-

Summary of Fusion-Containing Segregants

	Total	Hol+	Hol+/-	Hol-
ts lethal	8	8	0	0
wild-type	9	0	0	9

EC6

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
1a	+	+	+	-
1b	+	-	+	++/-
2a	+	+	+	-
2b	+	-	-	-
3a	+	+	+	-
3b	+	-	-	-
4a	+	-	-	-
4b	+	-	+	-
5a	+	-	-	-
5b	+	+	-	-
6a	+	-	-	-
6b	+	+	+	-
6c	+	+	-	-
6d	+	-	+	+
7a	+	+	+	-
8a	+	+	-	-
8b	+	-	-	-
8c	+	-	+	++/-
9a	+	+	+	-
9b	+	+	+	-
9c	+	-	-	-
9d	+	-	-	-
10a	-/+	-	-	-
10b	+	+	+	-
10c	+	-	-	-
10d	+	+	-	-
11a	+	-	-	-
11b	+	+	-	-
11c	+	-	+	+
11d	+	+	+	-
12a	+	+	+	-
13a	+	-	-	-
13b	+	+	+	-
13c	+	-	+	+
13d	+	+	-	-

EC6 (continued)

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
14a	+	+	+	-
15a	+	-	-	-
15b	+	+	+	-
15c	+	-	-	-
16a	+	-	+	+
16b	+	+	-	-
16c	+	-	-	-
17a	+	+	+	-
17b	+	-	+	-
17c	+	-	-	-
18a	+	+	-	-
18b	+	-	-	-
18c	+	-	+	-

Summary of Fusion-Containing Segregants

	Total	Hol+	Hol+/-	Hol-
ts lethal	9	6	0	3
wild-type	14	0	0	14

Comments

As discussed for cross EC4, this allele seems to require the presence of an additional unlinked modifer for suppression of the Hol- phenoptype.

TC1

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
1a	+	+	-	-
1b	+	-	+	+
1c	+	+	-	-
2a	+	+	-	-
2b	+	-	+	+
2c	+	+	-	-
2d	+	-	+	+
3a	+	-	+	+
3b	+	-	+	+
3c	+	+	-	-
3d	+	+	-	-
4a	+	-	+	+
4b	+	-	-	-
4c	+	+	-	-
5a	+	+	+	-
5b	+	-	-	-
5c	+	+	-	-
5d	+	-	+	+
6a	+	-	+/-	-
6b	+	-	+	+
7a	+	-	-	-
7b	+	+	+	-
7c	+	+	-	-
7d	+	-	+	-
8a	+	+	-	-
8b	+	-	-	-
8c	+	-	+	+
9a	+	+	-	-
9b	+	+	+	-
10a	+	-	-	-
10b	+	+	+	-

Summary of Fusion-Containing Segregants

	Total	Hol+	Hol+/-	Hol-
ts lethal	10	9	0	1
wild-type	4	0	0	4

TC2

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
1a	+	+	-	-
1b	+	-	+	+
1c	+	-	+	+
2a	+	-	+	+
2b	+	+	+	-
2c	+	-	-	-
3a	+	-	-	-
3b	+	+	-	-
3c	+	+	+	-
3d	+	-	+	+
4a	+	+	+	-
4b	+	-	-	-
4c	+	-	+	+
5a	+	+	+	-
5b	+	-	-	-
5c	+	-	-	-
6a	+	-	-	-
6b	+	-	+	+
7a	+	-	+	+
7b	+	+	-	-
7c	+	+	+	-
7d	+	-	-	-
8a	+	-	+	+
8b	+	-	+	+
8c	+	+	-	-
8d	+	+	-	-
9a	+	-	+	+
9b	+	+	-	-
9c	+	-	+	+
9d	+	+	-	-
10a	+	+	-	-
10b	+	-	+	+

Summary of Fusion-Containing Segregants

	Total	Hol+	Hol+/-	Hol-
ts lethal	12	12	0	0
wild-type	5	0	0	5

TC3

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
1a	+	-	+	+
1b	+	+	-	-
2a	+	+	+	--/+
2b	+	-	-	-
2c	+	-	-	-
3a	+	-	+	+
3b	+	-	-	-
3c	+	+	-	-
3d	+	+	+	--/+
4a	+	+	+	-
4b	+	-	+	+
4c	+	-	-	-
4d	+	+	-	-
5a	+	-	+	+
5b	+	+	-	-
6a	+	-	+	+
6b	+	-	+	+
6c	+	+	-	-
7a	+	-	+	+
7b	+	+	-	-
7c	+	-	+	+
7d	+	+	-	-
8a	+	+	-	-
8b	+	-	+	+
8c	+	+	-	-
8d	+	+	+	+

Summary of Fusion-Containing Segregants

	Total	Hol+	Hol+/-	Hol-
ts lethal	9	9	0	0
wild-type	4	1	0	3

TC4

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
1a	+	+	+	-
1b	+	-	-	-
1c	+	-	-	-
2a	+	-	-	-
2b	+	+	+	-
2c	+	-	+	+
3a	+	-	-	-
3b	+	+	+	-
3c	+	-	-	-
4a	+	-	+	+
4b	+	-	+	+
4c	+	+	-	-
4d	+	+	-	-
5a	+	-	+	+
5b	+	+	+	+
6a	+	-	-	-
6b	+	+	+	-
6c	+	-	+	+
7a	+	-	-	-
7b	+	-	+	+
7c	+	+	+	-
8a	+	+	-	-
8b	+	+	+	-
8c	+	-	+	+
8d	+	-	-	-
9a	+	-	+	+
9b	+	-	+	+
9c	+	+	-	-
9d	+	+	-	-
10a	+	-	+	+
10b	+	-	-	-
10c	+	+	-	-
11a	+	+	-	-
11b	+	+	+	-
12a	+	-	+	+
12b	+	+	+	-
12c	+	+	-	-
12d	+	-	-	-

TC4 (continued)

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
13a	+	+	+	-
13b	+	-	+	+
13c	+	+	-	-
13d	+	-	-	-
14a	+	-	-	-
14b	+	-	+	+
14c	+	+	-	-
14d	+	+	+	-
15a	+	+	+	-
15b	+	-	-	-
16a	+	-	+	+
16b	+	+	-	-
16c	+	-	+	+
16d	+	+	-	-
17a	+	+	-	-
17b	+	-	-	-
18a	+	+	-	-
18b	+	-	-	-
18c	+	-	+	+
18d	+	+	+	-
19a	+	+	+	-
19b	+	+	+	-
19c	+	-	-	-
20a	+	-	+	+
20b	+	+	+	-
20c	+	-	-	-

Summary of Fusion-Containing Segregants

	Total	Hol+	Hol+/-	Hol-
ts lethal	17	17	0	0
wild-type	15	0	0	15

TC5

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
1a	+	+	-	-
1b	+	-	+	+
1c	+	-	+	+
2a	+	+	-	-
2b	+	-	+	+
2c	+	-	+	+
3a	+	+	-	-
3b	+	-	+	+
3c	+	+	-	-
4a	+	+	+	++/-
4b	+	+	+	++/-
4c	+	-	-	-
4d	+	-	-	-
5a	+	+	+	++/-
5b	+	+	-	-
6a	+	+	+	++/-
6b	+	+	+	++/-
6c	+	+	-	-
6d	+	-	-	-
7a	+	-	-	-
7b	+	-	+	+
8a	+	-	-	-
8b	+	-	+	+
9a	+	-	-	-
9b	+	-	+	+
9c	+	+	+	++/-
9d	+	+	-	-
10a	+	-	-	-
10b	+	+	-	-
10c	+	+	+	++/-
11a	+	+	-	-
11b	+	+	+	+/-
12a	+	-	+	+
12b	+	+	-	-
12c	+	+	+	+/-

TC5 (continued)

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
13a	+	-	-	-
13b	+	+	+	+
13c	+	-	-	-
13d	+	+	+	+/-
14a	+	+	+	++/-
14b	+	-	-	-
14c	+	-	+	+

Summary of Fusion-Containing Segregants

	Total	Hol+	Hol+/-	Hol-
ts lethal	10	10	0	0
wild-type	12	1	11	0

Comments

The ability of every fusion-containing segregant to grow on Hol at some level indicates the presence of a reversion event linked to the fusion. Careful examination of the ts lethal and wild-type spores demonstrates that the cis-acting suppressor does not confer a full Hol+ phenotype. This indicates that the reversion event linked to the fusion is probably not a gene conversion event.

TC6

Segregant	YEPD (25)	YEPD (36)	-URA	HOL
1a	+	-	+	+
2a	+	-	+/-	+/-
3a	+	+	+	-
3b	+	+	+	-
3c	+	-	+	-
3d	+	-	+	+
4a	+	-	+	+
4b	+	+	+	-
4c	+	+	+	-
4d	+	-	-	-
5a	+	-	+	+
5b	+	-	+	+
5c	+	+	-	-
6a	+	-	-	-
6b	+	-	-	-
7a	+	+	+	-
7b	+	-	+	+
7c	+	-	+	+
7d	+	+	-	-
8a	+	-	-	-
8b	+	+	-	-
9a	+	+	+	-
9b	+	+	-	-
9c	+	-	+	+
9d	+	-	-	-
10a	+	-	+	+
10b	+	-	-	-

Summary of Fusion-Containing Segregants

	Total	Hol+	Hol+/-	Hol-
ts lethal	11	10	0	1
wild-type	6	0	0	6

EPILOGUE

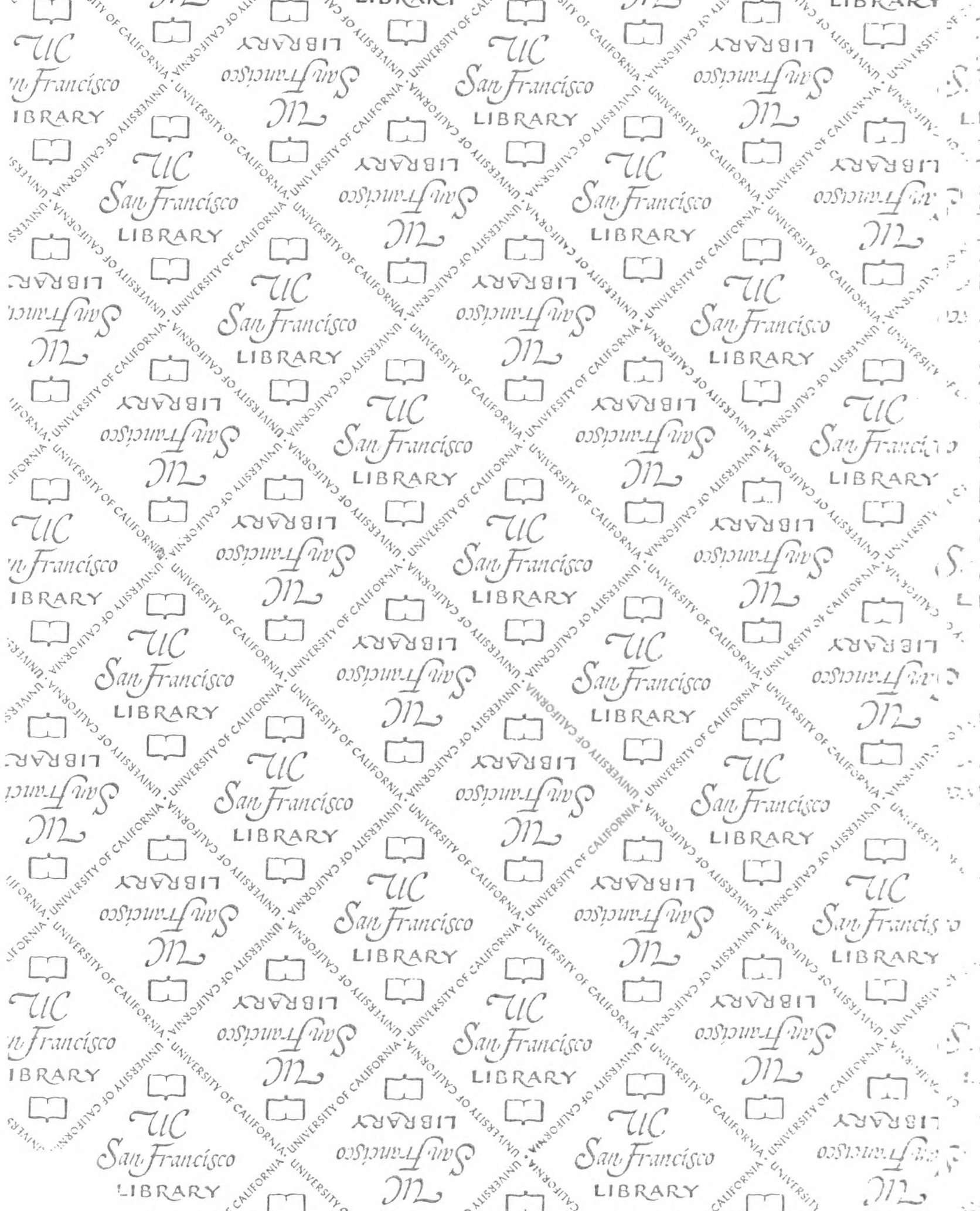
The results of this dissertation have demonstrated the potential of a genetic analysis of mRNA splicing in yeast. We have succeeded in constructing a spliced gene with an easily scored biological phenotype which we used in an in vivo mutant hunt for splicing mutations. In that this approach was sucessful (resulting in the isolation of the A5 mutation), we have great confidence in applying this genetic system to the analysis of the mRNA splicing machinery itself.

The analysis of the A5 mutation proved to be particularly informative because the yeast splicing machinery failed to activate a cryptic junction. This suggests that further analysis of point mutations in yeast may be more informative than in mammalian cells by allowing the study of partial reactions. In principle this will allow the nucleotides required for each step in the process to be identified. With this goal we have entered into a collaboration with the laboratories of John Rossi (City of Hope) and John Abelson (Cal. Tech.), to construct and characterize a number of mutations within the intron which by specific models are expected to separate the various steps of the pathway.

While the detailed analysis of the TACTAAC box and the 5' consensus sequence are proceeding, we are still struggling to understand how intron structure and sequence can influence the use of these splicing signals. The observation that the TACTAAC-like sequence

in the actin intron can be recognized by the splicing machinery has suggested a possible mechanism by which the use of a TACTAAC box could be enhanced. To further this observation it is necessary to determine if the TACTAAC-like sequence plays any role in the splicing of the wild-type actin intron. To this end, we have precisely removed the TACTAAC-like sequence from the actin intron by a small deletion with the goal of determining if this has any effect on the splicing efficiency of the intron.

We still do not understand the fundamental nature of the splicing machinery. Our first experience with the selection of extragenic suppressors has not identified the particular genes we most desire, those defining the splicing machinery itself. However, we plan to take advantage of our results to design more powerful selection schemes, now focusing on dominant suppressors which should suggest excellent candidates for the splicing enzymes themselves.


UC
San Francisco
LIBRARY
UNIVERSITY OF CALIFORNIA · UNIVERSITY OF CALIFORNIA