This is to certify that the

dissertation entitled

COMPONENTS INVOLVED IN THE
DEGRADATION OF RNA IN PLANTS.

presented by

Crispin Barwick Taylor

has been accepted towards fulfillment
of the requirements for

PH.D. degree in GENETICS

Pamela J. Green
Major professor

Date 9/24/93

O-12771

COMPONENTS INVOLVED IN THE DEGRADATION OF RNA IN PLANTS

By

Crispin Barwick Taylor

A DISSERTATION

Submitted to
Michigan State University
in partial fulfilment of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

Graduate Program in Genetics

1993

ABSTRACT

COMPONENTS INVOLVED IN THE DEGRADATION OF RNA IN PLANTS

By

Crispin B. Taylor

As a first step towards the elucidation of fundamental principles governing the control of gene expression at the level of mRNA stability in plants, two types of molecular components that may be involved in the process have been identified: ribonuclease genes in *Arabidopsis*, and sequence elements that target transcripts for rapid degradation in tobacco. Three *Arabidopsis* RNase genes, *RNS1*, *RNS2* and *RNS3* were identified by virtue of their homology to a class of fungal RNases. The RNS RNases are also related to the S-RNases, components of a plant self-recognition response termed self-incompatibility. The expression of the *RNS2* gene was found to be induced during senescence, and in response to phosphate starvation in *Arabidopsis*, implying a role for RNS2 in the remobilization of phosphate from RNA. Sequence comparisons demonstrated that RNS2 falls into a subclass within the T_2 family that is distinct from that formed by the S-RNases. Further, using a system developed to measure mRNA half-lives in stably transformed tobacco cell lines, two sequence elements have been found to confer instability on reporter transcripts.

When the first, termed DST, was present in two copies in the 3′ untranslated regions of two different reporter transcripts, it caused a pronounced decrease in the half-lives of those transcripts. The second element, termed the AUUUA repeat, also caused a marked destabilization of reporter transcripts. In both cases, the effect of the elements on transcript stability in the cell lines was mirrored by a marked decrease in the accumulation of the corresponding transcripts in transgenic plants. Finally, a novel differential screening strategy has been employed to identify eight genes with unstable transcripts (*GUT*s) in tobacco, solely on the basis of the instability of their mRNAs. Future identification of endogenous instability determinants within the *GUT*s should lead to a better understanding of the different mechanisms by which unstable transcripts may be recognized and targeted for rapid degradation in plants. Taken together, this work should provide the foundation for future efforts aimed at elucidating mRNA decay pathways in plants.

To **Cyn** and to **Emma** who have, quite simply, made it all worthwhile.

ACKNOWLEDGEMENTS

I would like to take this opportunity to thank the many individuals who have contributed directly and indirectly to the work described in this Dissertation. I thank Pam Green for her unfailing and enthusiastic support of my endeavors, and Tom Friedman, Natasha Raikhel, Mike Thomashow and Steve Triezenberg, for their patience, and for generously dispensing advice and encouragement whenever it has been solicited. Thanks are also due to each of the rotation students, undergraduates, post-docs and graduate students who have enriched my "PhD experience". It has truly been a pleasure to work in Pam's lab, not least because of the atmosphere created by Pam and all of these individuals. In particular, I want to thank André Dandridge, who has diligently provided expert technical assistance and good company almost from the start, and my co-authors Pauline Bariola, Tom Newman, Masaru Ohme-Takagi, whose individual contributions to this work are indicated throughout.

I would also like to acknowledge the contribution of all the people who work behind the scenes to keep the PRL and its constituent labs functioning. Without past and present office staff, and individuals such as Marlene Cameron, Barb Hoeffler, Bob Keever, Elliot Light and Kurt Stepnitz, completing my PhD would have been much harder.

Finally, I must recognize the contribution of my family. Emma has accepted the ephemeral presence of her father at home with grace, and Cynthia has supported, guided, encouraged and loved. Without their unfailing friendship I could never have done this. Thank you.

TABLE OF CONTENTS

LIST OF TABLES..... ix

LIST OF FIGURES..... x

CHAPTER 1: Introduction..... 1

CHAPTER 2: Genes with homology to fungal and S-gene ribonucleases are expressed in *Arabidopsis thaliana*..... 14

ABSTRACT..... 15

INTRODUCTION..... 16

MATERIALS AND METHODS..... 17

RESULTS AND DISCUSSION..... 20

LITERATURE CITED..... 30

CHAPTER 3: RNS2: a senescence-associated RNase of *Arabidopsis* that diverged from the S-RNases prior to speciation..... 32

ABSTRACT..... 33

INTRODUCTION..... 34

MATERIALS AND METHODS..... 36

RESULTS AND DISCUSSION..... 39

LITERATURE CITED..... 57

CHAPTER 4: DST sequences, highly conserved among plant *SAUR* genes, target reporter transcripts for rapid degradation in tobacco..... 60

ABSTRACT..... 61

INTRODUCTION..... 62

MATERIALS AND METHODS..... 67

RESULTS..... 76

DISCUSSION..... 99

LITERATURE CITED..... 107

CHPATER 5: The effect of sequences with high AU content on transcript stability in tobacco..... 113

ABSTRACT..... 114

INTRODUCTION..... 115

MATERIALS AND METHODS..... 118

RESULTS..... 119

DISCUSSION..... 131

LITERATURE CITED..... 139

CHAPTER 6: The identification of genes with unstable transcripts (*GUT*s) in tobacco..... 142

ABSTRACT..... 143

INTRODUCTION..... 144

MATERIALS AND METHODS..... 147

RESULTS AND DISCUSSION..... 150

CONCLUSIONS AND FUTURE PROSPECTS..... 169

CHAPTER 6 (cont.)
LITERATURE CITED..... 171

CHAPTER 7: Conclusions..... 173

LIST OF TABLES

Table 2-1: Amino acid homologies among the *RNS* gene products and related RNases..... 24

Table 4-1: Inhibition of mRNA synthesis in NT-1 cells..... 77

Table 4-2: mRNA half-lives in NT-1 cells stably transformed with pMON505..... 82

Table 4-3: Effect of DST sequences on mRNA stability in NT-1 cells..... 92

Table 5-1: Effect of the AUUUA repeat on *GUS* mRNA half-lives in NT-1 cells..... 125

Table 6-1: *GUT* and *NOTGUT* mRNA levels and half-lives..... 155

Table 6-2: Structural characteristics of *GUT*s and *NOTGUT*s..... 162

LIST OF FIGURES

Figure 2-1: Sequence alignment of the *RNS* gene products..... 22

Figure 2-2: Southern analysis of the *A. thaliana RNS* genes..... 25

Figure 2-3: Northern analysis of the *A. thaliana RNS* transcripts..... 27

Figure 3-1: Alignment of S- and S-like RNase amino acid sequences..... 43

Figure 3-2: Expression of *RNS2* in yeast..... 45

Figure 3-3: Gene genealogy of the RNase superfamily..... 47

Figure 3-4: Expression of *RNS2* in different organs of *Arabidopsis*..... 50

Figure 3-5: Induction of the *RNS2* transcript during senescence in *Arabidopsis*..... 51

Figure 3-6: Effect of phosphate starvation on *RNS2* expression..... 54

Figure 4-1: Location and structure of DST elements..... 66

Figure 4-2: Representation of the *CAT* and *GUS* genes following integration of pMON505 into the tobacco genome..... 79

Figure 4-3: Decay of transcripts in tobacco (NT-1) cells stably transformed with pMON505..... 81

Figure 4-4: Disappearance of transcripts encoded by *CAT* and *GUS* genes expressed under the control of copies of the *Cab-1* promoter in transgenic tobacco..... 86

Figure 4-5: Representation of the test and reference genes following integration into the tobacco genome..... 88

Figure 4-6: Destabilization of the *GUS* transcript by DST sequences in stably transformed NT-1 cells..... 91

Figure 4-7: Effects of DST and control inserts on the stability of *GUS* or globin test transcripts..... 94

Figure 4-8: Mapping of the 3′ ends of the *GUS* transcripts..... 96

FIgure 4-9: Reduced globin mRNA accumulation caused by DST sequences in transgenic plants..... 98

Figure 5-1: Constructs used to test the effect of AU-containing sequences..... 122

Figure 5-2: Rapid mRNA decay caused by the AUUUA repeat in tobacco cells..... 124

Figure 5-3: Histograms comparing the effect of the AU-containing inserts on *GUS* and globin mRNA stability..... 127

Figure 5-4: S1 nuclease protection analysis of the 3′ ends of the test transcripts..... 130

Figure 5-5: Effect of the AUUUA repeat in transgenic plants..... 132

Figure 6-1: Preliminary characterization of potential *GUT*s..... 153

Figure 6-2: Time-course experiments to determine *GUT* transcript half-lives..... 159

Figure 6-3: Sequence similarities of *GUT* and *NOTGUT* cDNAs..... 166

CHAPTER 1

INTRODUCTION

In eukaryotes, each of the steps between transcription of a gene and degradation of the corresponding protein are potential targets for regulation of expression. Of these, transcriptional control mechanisms are the best understood, both from the perspective of basic components of the transcriptional apparatus, and with respect to the mechanisms that regulate specific genes or sets of genes. However, it has become apparent that many eukaryotic genes are controlled to some extent by post-transcriptional events (Peltz *et al.*, 1991). The contribution of mRNA processing, transport, and stability, and of translational and post-translational mechanisms to the regulation of gene expression has therefore received increasing attention recently (Green, 1993; Gallie, 1993; Sachs, 1993; Belasco and Brawerman, 1993). Of these, regulation at the level of mRNA stability is among the most widely evoked, but the mechanisms by which such control is exerted remain, on the whole, poorly understood.

Unstable transcripts, with half-lives on the order of minutes, are the exception rather than the norm in higher eukaryotes. Average transcript half-lives are on the order of several hours (Siflow and Key, 1979; Brock and Shapiro, 1983), and it has therefore been suggested that the destabilization of specific transcripts is an active process (Hunt, 1988). The most extensively studied unstable transcripts in mammalian systems often encode proteins involved in the regulation of cellular growth and differentiation, such as c-*myc* (Jones and Cole, 1987), granulocyte/macrophage - colony stimulating factor

(GM-CSF) (Shaw and Kamen, 1986), and others. The proteins encoded by these genes are required only transiently in the cell, and their aberrant expression can have catastrophic consequences, such as cellular transformation and oncogenesis (Eick *et al.*, 1986). Rapid changes in the levels of these proteins is thought to be facilitated by the instability of the corresponding transcripts, as unstable transcripts will more quickly attain new steady-state levels (both higher and lower) following changes in the transcription rate of their genes (Peltz, *et al.*, 1991).

The rate of synthesis of an individual transcript is regulated via the interactions between components of the transcriptional machinery with specific sequences in the promoter of the corresponding gene. Following this paradigm, it seems likely that the degradation rate of that transcript will largely be defined via the interactions between specific sequence elements within that transcript and cellular factors that mediate it's degradation. Indeed, a number of sequence determinants of mRNA instability have been identified in mammalian systems. Among the first instability sequences to be described was an AU-rich element (ARE) in the 3' untranslated region (UTR) of GM-CSF. When this element was placed in the 3'UTR of the otherwise stable β-globin transcript, it caused the transcript to be rapidly degraded (Shaw and Kamen, 1986). Other characterized instability sequences include, for example, an element in the 3'UTR of the transferrin receptor mRNA, which facilitates the destabilization of this transcript in response to increases in cellular iron concentration (Klausner

et al., 1993), and a coding-region portion of the c-*fos* transcript which only appears to destabilize this transcript in serum-stimulated cells (Shyu *et al.*, 1989). In all of these cases, one or more factors have been identified that specifically interact with each instability sequence. Presumably some or all of these factors are involved in targeting the transcripts for rapid degradation.

Rapid changes in gene expression are likely to be particularly important in plants (eg. Theologis, 1986). As sessile organisms plants cannot move away from adverse stimuli but must respond by altering their metabolism, often quite quickly, if they are to survive (Green, 1993). Plants are therefore likely to have a large number of genes whose expression must change quickly in response to environmental stimuli and, accordingly, mRNA stability may figure prominently in the regulation of many plant genes (Green, 1993). As yet, however, there are only a few reports of unstable mRNAs in plants, and none in which the mechanisms that mediate a transcripts' instability are defined.

In addition to sequence elements that confer instability on transcripts, RNases are also likely to figure prominently in the degradation pathways of all mRNAs, in animals as well as plants. A large number of RNases have been described from various plant species, and have been classified into four major groups on the basis of characteristics such as molecular mass, pH optima, homopolymer specificity and cation requirements (Wilson, 1982). The levels of some of these RNases, or their activities, are known to respond to a variety of environmental stimuli and developmental cues, such as senescence (Matile

and Winkenbach, 1971; Blank and McKeon, 1991). However, the RNases that play a role in mRNA degradation have not been differentiated from those with other functions in any eukaryote. Without knowing which enzymes are involved, it is difficult to elucidate the mechanisms by which mRNAs are recognized and degraded.

This Dissertation addresses efforts to identify sequence elements that confer instability on reporter transcripts in tobacco, and also focusses on the initial isolation and characterization of RNase genes in *Arabidopsis*. The bases for each of the major research projects described in the Dissertation are outlined below.

Arabidopsis RNases and RNase Genes

One approach toward the identification of plant RNases that participate in the degradation of messenger RNA is to biochemically characterize as many RNases as possible from a model plant species. Using this work as a basis, it should then be possible to screen for mutants deficient in any of the RNases identified. These mutants can then be used to address the function of the corresponding RNase. To this end, initial studies have led to the identification of some 16 bands of RNase activity in extracts of *Arabidopsis* tissue using a gel-based RNase assay (Yen and Green, 1991). Among these are

representatives of at least three of the four major classes of RNase previously described in plants (Wilson, 1982; Yen and Green, 1991). More recently a number of *Arabidopsis* mutants have been identified that exhibit differences from the wild type profile of RNase activities on these gels (Abler, M.A., Danhof, L. and Green, P.J. unpublished). By examining these mutants for defects in the degradation of specific mRNAs, it may be eventually be possible to identify those mutants affected in the activity of a RNase involved in mRNA degradation.

A complementary approach toward elucidating the function of RNases in *Arabidopsis* is to identify RNase genes in this species. Until recently, the only plant RNases that had been cloned were the S-RNases (McClure *et al.*, 1989b). The S-RNases, which are associated with self-incompatibility (SI) in the Solanaceae, were identified as RNases through their homology to a class of fungal RNases typified by RNase T_2 of *Aspergillus oryzae* (Kawata *et al.*, 1988; McClure *et al.*, 1989b). The S-RNases exhibit many of the properties expected for cellular recognition molecules, but their precise roles in mediating the SI response are still not clear (Haring *et al.*, 1990). Nonetheless, the sequence conservation between the S-RNases and the fungal RNases lent itself to a PCR approach toward detecting related RNase genes in *Arabidopsis*. The application of this strategy, and the identification of three *Arabidopsis* RNase genes, *RNS1*, *RNS2*, and *RNS3*, are described in Chapter 2 of this Dissertation. Chapter 3 includes a discussion of the evolutionary relationships among the

T_2/S class of RNases. The further characterization of *RNS2* expression in *Arabidopsis*, and the potential functions of the RNS2 protein are also discussed in this Chapter.

Identification of *cis*-acting sequences that mediate rapid mRNA decay

As discussed above, unstable transcripts are likely to be targeted for rapid degradation via interactions between specific sequences within that transcript and components of the cellular RNA degradation machinery, such as RNases. The identification of sequences that confer instability on transcripts is thus a prerequisite to understanding how mRNA decay rates are controlled. One approach that has been widely employed in mammalian systems is to examine the effect of putative instability determinants on the half-lives of otherwise stable reporter transcripts (reviewed in Peltz, *et al.*,) In many of these experiments, reporter transcript half-lives are measured by blocking transcription with inhibitors, such as Actinomycin D (ActD). The subsequent disappearance of the reporter transcript is followed by, for example, northern blotting. Chapter 4 of this Dissertation describes the development of an analogous system for measuring transcript half-lives in stably transformed plant cells. The system makes use of a tobacco cell line, NT-1, that can be stably transformed (An, 1985) with reporter constructs containing putative instability

determinants, and that is also amenable to ActD treatment.

The NT-1 cell system has been used to identify two very different sequence elements that function to destabilize transcripts in tobacco. The first, termed DST, is derived from the Small Auxin Up RNAs (*SAUR*s), which were initially characterized in soybean (McClure *et al.*, 1989a). This element was chosen to be tested for its effect on transcript stability because the *SAUR*s appear to be among the most inherently unstable plant transcripts yet characterized, with half lives estimated to be on the order of 10 - 50 minutes (McClure and Guilfoyle, 1989; Franco *et al.*, 1990). Also, the DST element is a short (~40 bp) sequence that is conserved in the 3'UTRs of all the *SAUR* genes whose sequences have been reported to date (McClure *et al.*, 1989a; Yamamoto *et al.*, 1992; Gil, P. Liu, Y., Orbović, V., Verkamp, E., Poff, K.L. and Green, P.J., submitted). The conservation of this element in a non-coding region supports its functional significance; its location is similar to that of several mammalian instability determinants, which suggested that it may mediate the instability of the *SAUR* transcripts (McClure *et al.*, 1989a). When present in two copies in the 3'UTRs of reporter transcripts, the DST element caused a pronounced decrease in the half-lives of those transcripts in stably transformed NT-1 cells. The same transcripts also accumulated to markedly reduced levels in transgenic tobacco plants, implying that the DST element can target transcripts for rapid degradation in plants as well as NT-1 cells. The experiments to analyze the effect of the DST element on transcript stability in

NT-1 cells, and on transcript accumulation in transformed tobacco are also described in Chapter 4.

The DST element differs from a prominent class of mRNA instability sequences that have been shown to function in mammalian cells. Sequences in the 3'UTRs of a number of proto-oncogene, lymphokine and cytokine transcripts are AU-rich and contain multiple copies of an AUUUA motif. The AREs from GM-CSF (Shaw and Kamen, 1986) and c-*fos* (Shyu *et al.*, 1989) are known to target reporter transcripts for rapid degradation in mammalian cells. To test whether similar AU-rich sequences may also destabilize transcripts in plants, two sequence elements with identical A + U content, but different primary structures, were inserted into the 3'UTRs of reporter transcripts. The effect of these elements on reporter transcript half-life in NT-1 cells was analyzed as outlined above for the DST element. Of the sequences tested, only an element containing 11 overlapping copies of the AUUUA sequence (termed the AUUUA repeat) conferred instability on the reporter transcripts. This element also caused a marked decrease in the accumulation of the reporter transcript in transformed tobacco plants. These experiments are described in Chapter 5 of this Dissertation.

The identification of the DST and AUUUA elements as instability determinants in plants is an important step toward the goal of understanding the mechanisms controlling mRNA stability in plants. However the examination of the role of the DST and AUUUA elements was prompted by previous

evidence suggesting that they may affect transcript stability. They are therefore unlikely to be representative of the repertoire of instability sequences that may function in endogenous plant transcripts. An initial step toward the identification of additional plant instability sequences is to first identify unstable transcripts, before determining which portions of those transcripts mediate their instability. We have recently used a novel differential screening strategy to search for genes with unstable transcripts (*GUT*s) in tobacco. One major advantage of this approach is that it facilitates the identification of *GUT*s solely on the basis of the instability of their transcripts. The screen is based on the rationale that levels of unstable mRNAs will rapidly fall in cells treated with ActD. Poly(A)$^+$ RNA isolated from ActD-treated cells will therefore be depleted of unstable transcripts, relative to the levels of those transcripts in control, untreated cells. The differential screen, and the preliminary characterization of the eight *GUT*s we have isolated using this strategy, are the subjects of Chapter 6 of this Dissertation.

In Chpater 7, some of the recent work that has had as its impetus the studies performed in this Dissertation is summarized, and suggestions for future research efforts are outlined.

LITERATURE CITED

An, G. (1985) High efficiency transformation of cultured tobacco cells. Plant Physiol. 79: 568-570

Belasco, J., and Brawerman, G. (1993) eds. Control of messenger RNA stability. (San Diego, California: Academic Press)

Blank, A., and McKeon, T.A. (1991) Expression of three RNase activities during natural and dark-induced senescence of wheat leaves. Plant Physiol. 97: 1409-1413

Brock, M.L., and Shapiro, D.J. (1983) Estrogen stabilizes vitellogenin mRNA against cytoplasmic degradation. Cell 34: 207-214

Eick, D., Piechaczyk, M., Henglein, B., Blanchard, J.M., Traub, B., Kofler, E., Wiest, S., Lenoir, G.M., and Bornkamm, G.W. (1986) Aberrant c-myc RNAs of Burkitt's lymphoma cells have longer half-lives. EMBO J. 4: 3717-3724

Franco, A.R., Gee, M.A., and Guilfoyle, T.J. (1990) Induction and super-induction of auxin-responsive mRNAs with auxin and protein synthesis inhibitors. J. Biol. Chem. 265: 15845-15849

Gallie, D.R. (1993) Post-transcriptional regulation of gene expression in plants. Annu. Rev. Plant Physiol. Mol. Biol. 44: 77-106

Green, P.J. (1993) Control of mRNA stability in plants. Plant Physiol. 102: 1065-1070

Haring, V., Gray, J.E., McClure, B.A., Anderson, M.A., and Clarke, A.E. (1990) Self-incompatibility: a self-recognition system in plants. Science 250: 937-941

Hunt, T. (1988) Controlling mRNA lifespan. Nature 38: 567-568

Jones, T.R., and Cole, M.D. (1987) Rapid cytoplasmic turnover of c-myc mRNA: requirement of the 3′ untranslated sequences. Mol. Cell. Biol. 7: 4513-4521

Kawata, Y., Sakiyama, F., and Tamaoki, H. (1988) Amino acid sequence of ribonuclease T_2 from *Aspergillus oryzae*. Eur. J. Biochem. 176: 683-697

Klausner, R.D., Rouault, T.A., and Harford, J.B. (1993) Regulating the fate of mRNA: the control of cellular iron metabolism. Cell 72, 19-28

Matile, P., and Winkenbach, F. (1971) Function of lysosomes and lysosomal enzymes in the senescing corolla of the Morning Glory (*Ipomoea purpurea*). J. Exp. Bot. 22: 759-771

McClure, B.A., and Guilfoyle, T. (1989) Rapid redistribution of auxin-regulated RNAs during gravitropism. Science 243: 91-93

McClure, B.A., Hagen, G., Brown, C.S., Gee, M.A., and Guilfoyle, T.J. (1989a) Transcription, organization, and sequence of an auxin-regulated gene cluster in soybean. Plant Cell 1: 229-239

McClure, B.A., Haring, V., Ebert, P., Anderson, M.A., Simpson, R.J., Sakiyama, F., and Clarke, A.E. (1989b) Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. Nature 342: 955-957

Peltz, S.W., Brewer, G., Bernstein, P., Hart, P.A., and Ross, J. (1991) Regulation of mRNA turnover in eukaryotic cells. In Critical Reviews in Eukaryotic Gene Expression, G.S. Stein, J.L. Stein, and J.B. Lian, eds (Boca Raton: CRC Press), pp. 99-126

Sachs, A.B. (1993) Messenger RNA degradation in eukaryotes. Cell 74, 413-421

Shaw, G., and Kamen, R. (1986) A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell 46: 659-667

Shyu, A.-B., Greenberg, M.E., and Belasco, J.G. (1989) The c-*fos* transcript is targeted for rapid decay by two distinct mRNA degradation pathways. Genes and Devel. 3: 60-72

Siflow, C.D. and Key, J.L. (1979) Stability of polysome-associated polyadenylated RNA from soybean suspension culture cells. Biochemistry 18: 1013-1018

Theologis, A. (1986) Rapid gene regulation by auxin. Annu. Rev. Plant Physiol. 37: 407-438

Wilson, C.M. (1982) Plant nucleases: biochemistry and development of multiple molecular forms. *in* M.C. Rattazzi, S.C. Scandalios, G.S. Whitt, eds, Isozymes. Current Topics in Biological and Medical Research, Vol. 6.

Alan R. Liss, Inc, New York, pp 33-54

Yamamoto, K.T., Mori, H., and Imaseki H. (1992) cDNA cloning of indole-3-acetic acid-regulated genes: Aux22 and SAUR from mung bean (*Vigna radiata*) hypocotyl tissue. Plant Cell Physiol. 33: 93-97

Yen, Y., and Green, P.J. (1991) Identification and properties of the major ribonucleases of *Arabidopsis thaliana*. Plant Physiol. 97: 1487-1493

CHAPTER 2

GENES WITH HOMOLOGY TO FUNGAL AND S-GENE RNASES ARE EXPRESSED IN *Arabidopsis thaliana*

This Chapter was originally published in *Plant Physiology*.

Reference: Taylor, C.B. and Green, P.J. (1991). Genes with homology to fungal and S-gene RNases are expressed in *Arabidopsis thaliana*. Plant Physiol. 96: 980-984.

ABSTRACT

The only RNase genes that have been cloned from higher plants are those expressed in species that are known to exhibit self-incompatibility, such as the S-genes of *Nicotiana alata*. In this report, we investigated whether the expression of this particular type of RNase gene is restricted to self-incompatible species, or if similar genes are expressed in other plants. Using a polymerase chain reaction approach we have identified a set of three putative RNase genes in the self-compatible plant *Arabidopsis thaliana (L.) Heynh.* ecotype Columbia. These *A. thaliana* genes, designated RNS1, RNS2, and RNS3, do not cross-hybridize to each other, and their products are homologous to both the S-gene RNases of *N. alata* and a set of related fungal enzymes. The *A. thaliana* RNS1, RNS2 and RNS3 genes encode transcripts of 1.1, 1.3 and 1.1 kb respectively, and of the three genes, RNS2 is the most highly expressed in whole plants. The identification of the RNS genes in a self-compatible species suggests that this class of RNases is of general significance in RNA catabolism in higher plants.

INTRODUCTION

RNA catabolism is an important component of many cellular processes including, for example, the post-transcriptional regulation of gene expression through mRNA decay. However, little is known about the control of RNA degradation in any eukaryote. This is especially true in higher plants where, despite the biochemical characterization of numerous RNase activities [18, 6], the only RNase genes that have been cloned are those coding for the style-specific products of the S-genes, components of the self-incompatible response in species such as *Nicotiana alata* [2].

The S-genes had been cloned for some time before comparison of deduced amino acid sequences [12] showed that their products were related to a group of fungal RNases consisting of RNase T_2 of *Aspergillus oryzae* [10], RNase Rh of *Rhizopus niveus* [8] and RNase M of *Aspergillus saitoi* [9]. Style extracts and purified S-gene products of *N. alata* were shown to have RNase activity *in vitro* [12], but the precise roles of the S-gene products, termed S-RNases [12], in arresting the growth of incompatible pollen tubes remains unclear. One model is that the S-RNases produced in the style enter incompatible pollen tubes and attack their rRNA, hence inhibiting their growth [13, 7].

We are interested in identifying RNases that have fundamental roles in RNA catabolism in higher plants. This interest prompted us to ask whether the expression of RNases similar to the S-RNases is only associated with the self-

incompatible response, or if related enzymes are also expressed in self-compatible species. The latter would suggest that these RNases have more widespread functions in higher plants. Our approach was to use DNA primers, corresponding to amino acids conserved among fungal and S-RNases [12], for PCR[2] amplification of an *Arabidopsis thaliana (L.) Heynh.* ecotype Columbia cDNA library.

Following PCR amplification of the *A. thaliana* cDNA library, we have identified three putative RNase genes, termed RNS1, RNS2 and RNS3. These genes are all expressed in the self-compatible species *A. thaliana*, and their products are strikingly similar to the *N. alata* S-RNases and the fungal RNases typified by RNase T_2. This suggests that RNases of this type may be expressed in many plant species, and therefore are likely to be involved in other aspects of RNA turnover beyond those associated with self-incompatibility.

MATERIALS AND METHODS

cDNA Library

An *Arabidopsis thaliana (L.) Heynh.* ecotype Columbia cDNA library constructed from total green tissue using the Lambda Zap vector (Stratagene) was kindly provided by N.-H. Chua and A. Gasch. The library was 'zapped' into plasmid form [4], and DNA representing 180,000 recombinants was amplified using

PCR, as described below.

PCR Cloning

The sequences of the primers used for PCR were:

```
                                                          E
                                                          N
                                                          I   Y
          H   G   L   W   P   D             T   G   H   K   V   W   E
5' GAATTCAT GGN TTN TGG CCN GA 3' 5' CTCGAGT NCC ATG TTT TTT ATA TTC 3'
   EcoRI  C     C                   XhoI        G   C   AAC GC  C
                                                      G   C
```

Nucleotide sequences were derived from previously published deduced amino acid sequences [2, 3, 8, 9, 10] indicated in bold type, where N represents all four possible nucleotides. The second primer was the complement of the sequence coding for the amino acids indicated. Amplification of sequences in the library flanked by these primers was carried out using two sequential PCR experiments. The initial amplification consisted of 25 cycles of 94°C for 1 min, 42°C for 2 min, and 72°C for 3 min. The reaction, in a volume of 25μL, contained 0.3μg of plasmid library DNA linearized with EcoRI, 42 pmol of each primer, 50mM KCl, 10mM Tris-HCl, pH 8.3, 1.5mM $MgCl_2$, 0.01% (w/v) gelatin, 200μM of each dNTP and 0.6 units of Taq DNA polymerase (Perkin Elmer). The products of the reaction were separated on a low melting temperature agarose gel and the major product (175-225bp) was excised and used for a second round of PCR under the same conditions as above, except that the 42°C step was carried out at 50°C. The major PCR product was excised, blunt-end ligated into a plasmid vector, and

sequenced using the dideoxy method of Sanger *et al* [16].

Southern and Northern Analyses

Genomic DNA was isolated from all above-ground tissues of mature, five-week old soil-grown *A. thaliana*, using the method of Dellaporta *et al.* [5]. After digestion with restriction endonucleases, DNA was separated on an agarose gel and blotted onto BioTrace RP (Gelman Scientific). Prehybridization was for five hours at 52°C in a buffer containing 5XSSC, 10X Denhardt's solution [15], 0.1% SDS, 0.1M potassium phosphate, pH6.8 and 100μg/mL denatured salmon sperm DNA. Hybridization was for 16 hours at 52°C in the same buffer, except that 10% (w/v) dextran sulphate and 30% (v/v) formamide were added, and the SDS was omitted.

Total RNA was isolated from similar tissues essentially as described by Puisant and Houdebine [14]. PolyA$^+$ RNA was purified by oligo dT-cellulose affinity chromatography as recommended by the supplier (Pharmacia). PolyA$^+$ and polyA$^-$(the flow-through from the oligo dT-cellulose column) RNAs were denatured and separated on formaldehyde/agarose gels. Following transfer to BioTrace RP, the blots were treated as described above.

RESULTS AND DISCUSSION

Amplification of Putative RNase Sequences

The primers for PCR amplification described above were designed to code for two of the longest stretches of amino acid homology among fungal and S-RNases [12]. These blocks of amino acid homology have subsequently been found in the S-gene products of a number of other plant species that exhibit self-incompatibility, including *Petunia hybrida* [3], *Petunia inflata* [1] and *Solanum tuberosum* [7], as well as in additional *N. alata* S-alleles [11]. In between these blocks of homology, McClure *et al* [12] also noted a number of individual amino acid residues that were absolutely conserved among three S-RNases (S_2, S_3 and S_6) and two fungal RNases (T_2 and Rh). The presence of these residues in the PCR products from *A. thaliana* would provide additional confirmation that the corresponding polypeptides were related to the aforementioned RNases.

When the primers described above were used to amplify hybridizing sequences from an *A. thaliana* cDNA library, the major PCR product was approximately 200bp. This could encode some 67 amino acids, which was within the range of sizes expected for the region between the primers from the deduced amino acid sequences of the RNases discussed above.

Relationship of the RNS Gene Products to Other RNases

Full or partial (single nucleotide) sequencing of 21 individual recombinants obtained following cloning of the major PCR product demonstrated that there are at least three putative RNases in *A. thaliana* related in deduced amino acid sequence to the fungal and S-RNases shown in Figure 2-1. These have been designated RNS1, RNS2 and RNS3. Of the 21 clones that were sequenced, and two additional clones analyzed by dot blot hybridization (not shown), twelve were identical to RNS1, five to RNS2 and three to RNS3. The three remaining clones showed no similarity to the RNases shown in Figure 2-1.

In Figure 2-1 the amino acid residues that are absolutely conserved among the RNS gene products of *A. thaliana* and the S_2, S_3, S_6, T_2 and Rh RNases are boxed. These identical residues include the blocks of amino acid homology used for design of the PCR primers. Moreover, all of the residues between these blocks that were shown to be completely conserved among the fungal and S-RNases in Figure 2-1 [12], are also found in each of the three RNS clones. This strongly suggests that the RNS gene products are indeed RNases.

This argument is strengthened by the fact that two of the identical amino acids that fall within the regions covered by the PCR primers, His 1 and His 59 in Figure 2-1, are known to be part of the active site of RNase T_2 [10] and RNase M [9]. Two other amino acids that are identical among all the RNases shown in Figure 2-1, Trp 38 and Pro 39, have been suggested to contribute to the base non-specific character of RNase T_2 [10]. Another residue that is

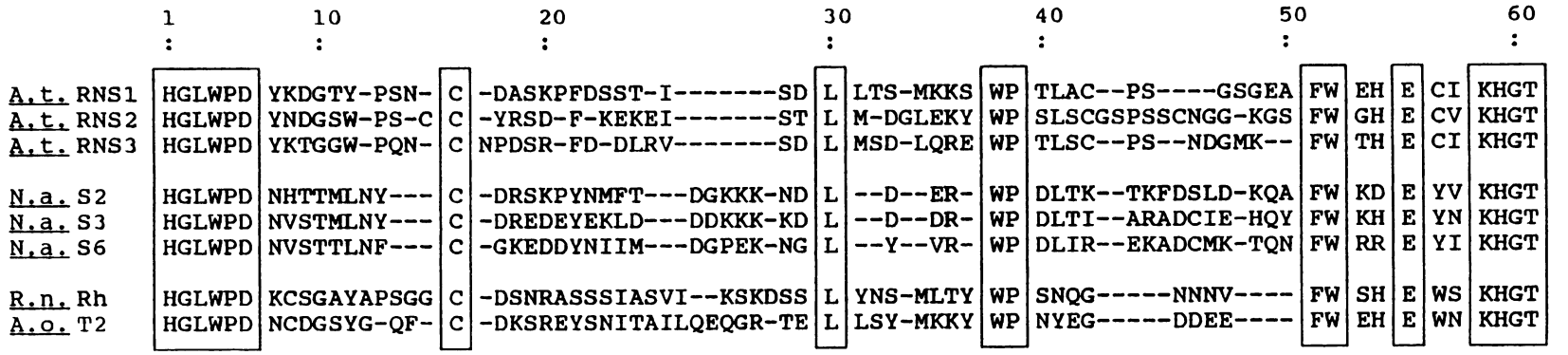

Figure 2-1. Sequence alignment of the RNS gene products with fungal and S-RNases.

Alignment of deduced amino acid sequences of the *A. thaliana* RNS genes with the S-RNases of *N. alata* [12], RNase T_2 of *A. oryzae* [10] and RNase Rh of *R. niveus* [8]. Sequence alignment was based on that of McClure *et al.* [12]. Residues that are identical in all of the RNases are boxed. The sequences are numbered beginning with 1 for the first amino acid in the cloned fragment of RNS 1. Dashes indicate gaps introduced to optimize the alignment. The abbreviations to the left of the figure are as follows: A.t., *A. thaliana*; N.a., *N. alata*; R.n., *R. niveus*; A.o., *A. oryzae*.

identical in all the sequences in Figure 2-1, Cys 16, has been found (in RNase T_2) to form a disulfide bond with a Cys located just outside the region shown in Figure 2-1 [10].

In addition to the identical amino acids, there is significant homology between the RNS gene products and the fungal and S-RNases throughout the cloned region. When the percentages of functionally identical amino acids are compared (Table 2-1), the homology between individual RNS sequences and the S-RNases is between 34 and 48%, while that between the RNS sequences and the fungal RNases ranges from 46 to 57%. The highest degree of homology (between 65 and 72%) is found when the individual RNS sequences are compared to each other. These extensive homologies are also reflected in hydrophobicity plots of this region of the RNases (data not shown).

Structure and Expression of the RNS Genes

Analysis of DNA sequence data showed that the percentages of identical nucleotides for the RNS genes of *A. thaliana* were 59% for RNS1 and RNS2, 57% for RNS1 and RNS3, and 61% for RNS2 and RNS3 (data not shown). This suggested that the RNS sequences would not cross-hybridize, which was confirmed by Southern analysis (Figure 2-2). As shown in Figure 2-2, each RNS clone hybridizes to a unique set of restriction fragments in *A. thaliana* genomic DNA, indicating that RNS1, RNS2 and RNS3 correspond to different genes. It is likely that the RNS genes are not members of large gene families

Table 2-1. Amino Acid Homologies Among RNS Gene Products and Related RNases

	RNS2	RNS3	N.a.S2	N.a.S3	N.a.S6	R.n.Rh	A.oT2
RNS1	67	72	48	40	34	52	57
RNS2		65	44	45	38	46	48
RNS3			47	37	35	46	51

Deduced amino acid sequences were aligned as shown in Figure 2-1, and the number of functionally identical amino acids for each pair of gene products was totalled. Percent homologies were calculated by dividing the number of functionally identical amino acids by the average total number of amino acids for each pair of sequences. Functionally identical amino acids are grouped as: A,S,T; N,Q; D,E; I,L,M,V; H,K,R; and F,W,Y. Abbreviations are as described in the legend to Figure 2-1.

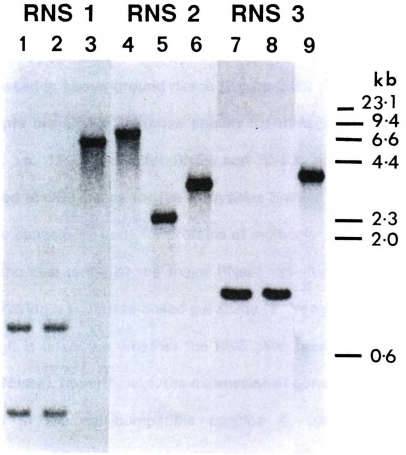

Figure 2-2. Southern analysis of the *A. thaliana* RNS genes.

10μg of genomic DNA was cut with EcoRV (lanes 1,4,7), EcoRV and HindIII (lanes 2,5,8) or HindIII (lanes 3,6,9) and separated on a 1% agarose gel. Blots were hybridized with the [^{32}P]-labelled RNS probes indicated above the lanes.

because, with one exception, each clone hybridizes to a single band in genomic Southerns. The two bands in lanes 1 and 2 of Figure 2-1 are expected because of the presence of a site for the restriction endonuclease EcoRV within the cloned fragment of RNS1.

Analysis of polyA$^+$ RNA shows that each of the RNS genes of *A. thaliana* is expressed in above-ground tissue (Figure 2-3). The sizes of the RNS gene transcripts are similar to those shown for transcripts of the S-genes of *P. hybrida*, i.e. 1100 bases for RNS1 and RNS3, and 1300 bases for RNS2, compared to 900 bases for the *P. hybrida* S-allele 3A [3]. These transcripts have the capacity to code for proteins of molecular mass up to 40kD, which is within the size range of the major RNase activities recently identified in *A. thaliana* using a substrate-based gel assay [Y. Yen and P.J.G., submitted]. At present it is unknown whether the RNS gene products correspond to any of these RNases. Nevertheless, the expression of genes highly homologous to the S-genes in the self-compatible species *A. thaliana* suggests that the corresponding RNases may be involved in more general aspects of RNA catabolism in higher plants.

Interestingly, expression levels for each of the three *A. thaliana* genes differ markedly in this tissue sample. The northern blots with RNS1 and RNS3 probes had to be exposed several times longer than that with an RNS2 probe to produce signals of similar intensities (Figure 2-3). These data obtained using Poly A$^+$ RNA, and other experiments performed with total RNA (not shown)

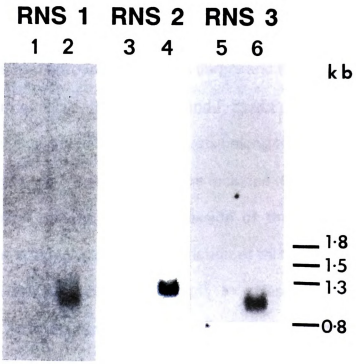

Figure 2-3. Northern analysis of the *A. thaliana* RNS transcripts.

6μg of poly A^+ (lanes 2,4,6) and 20μg of poly A^- (lanes 1,3,5) RNA isolated from above-ground tissue were denatured and separated on a 1% agarose/2% formaldehyde gel. Blots were hybridized with the [^{32}P]-labelled RNS probes indicated above the lanes. Lanes 3 and 4 were exposed for 1 day; lanes 1,2,5 and 6 for 7 days.

clearly indicate that RNS2 is the most highly expressed of the RNS genes. However, the relative transcript levels in the tissues analyzed in the northern blots do not correlate with the frequency with which the individual RNS clones were isolated by PCR amplification of the cDNA library (see above). This discrepancy may reflect differences in primer homology to each of the RNS cDNAs. Alternatively, it may be due to a difference between the tissue samples used for preparation of the cDNA library and those used to prepare the RNA for northern analysis. Implications and Future Prospects

Our data indicate that the RNS gene products belong to a class of enzymes that have a homologous active site. This class includes the S-RNases of *N. alata*, and several fungal RNases. The expression of the RNS genes in a self-compatible plant strongly argues for a function of this class of RNases beyond their role in self-incompatibility, perhaps in stress related responses or senescence [6, 18]. It remains to be determined whether the expression patterns of the RNS genes are the same as those of the S-genes and whether the corresponding enzymes share any common functions.

The cloning of the RNS genes from *A. thaliana* should facilitate the characterization of their functions via antisense RNA techniques, or the use of insertional mutagenesis in the future [17]. Further studies on the temporal and spatial regulation of RNS gene expression in *A. thaliana* are in progress. These approaches, in conjunction with a detailed biochemical study of the RNases of *A. thaliana*, will help to define the roles of the RNS gene products in RNA

catabolism, and may also provide insight into the roles of related proteins in the self-incompatible response.

ACKNOWLEDGEMENTS

We are greatful to Drs. Nam-Hai Chua and Alex Gasch for providing the *A. thaliana* cDNA library, and to Jim Dombrowski and Robin Steinman for zapping the library into plasmid form. We thank Dr. Natasha Raikhel and members of the Green lab for helpful comments on the manuscript.

LITERATURE CITED

[1] Ai Y, Singh A, Coleman CE, Ioeger TR, Kheyr-Pour A, Kao TH (1990) Self-incompatibility in *Petunia inflata*. Isolation and characterization of complementary DNA encoding three S-allele-associated proteins. Sex. Plant Reprod. **3**: 130-138

[2] Anderson MA, Cornish EC, Mau S-L, Williams EG, Hoggart R, Atkinson A, Bonig I, Grego B, Simpson R, Roche PJ, Haley JD, Penshow JD, Niall HD, Tregear G, Coghlan JP, Crawford RJ, Clarke AE (1986) Cloning of cDNA for a stylar glycoprotein associated with expression of self-incompatibility in *Nicotiana alata*. Nature **321**: 38-44

[3] Clark KR, Okuley JJ, Collins PD, Sims TL (1990) Sequence variability and developmental expression of S-alleles in self-incompatible and pseudo self-compatible Petunia. Plant Cell **2**: 815-826

[4] Delauny AJ, Verma DPS (1990) Isolation of plant genes by heterologous complementation in *Escherichia coli*. *In* SB Gelvin, RA Schilperoort, DPS Verma eds, Plant Molecular Biology Manual. Kulwer Academic, Dordrecht, pp. A14/1-A14/23

[5] Dellaporta SL, Wood J, Hicks JB (1983) A plant DNA minipreparation: Version II. Plant Mol. Biol. Rep. **1**: 19-21

[6] Farkas GL (1982) Ribonucleases and ribonucleic acid breakdown. *In* D Boulter, B Parthier, eds, Nucleic Acids and Proteins in Plants, Encyclopedia of Plant Physiology, Vol 16. Springer-Verlag, Berlin. pp 224-262

[7] Haring V, Gray JE, McClure BA, Anderson MA, Clarke AE (1990) Self-incompatibility: a self-recognition system in plants. Science **250**: 937-941

[8] Horiuchi H, Yanai K, Takagi M, Yano K, Wakabayashi E, Sanda A, Mine S, Ohgi K, Irie M. (1988) Primary structure of a base non-specific ribonuclease from *Rhizopus niveus*. J. Biochem (Tokyo) **103**: 408-418

[9] Irie M, Watanabe H, Ohgi K, Harada M (1986) Site of alkylation of the major ribonuclease from *Aspergillus saitoi* with iodoacetate. J. Biochem (Tokyo) **99**: 627-633

[10] Kawata Y, Sakiyama F, Tamaoki H (1988) Amino-acid sequence of ribonuclease T_2 from *Aspergillus oryzae*. Eur. J. Biochem. **176**: 683-697

[11] Kheyr-Pour A, Bintrin SB, Ioerger TR, Remy R, Hammond SA, Kao TH (1990) Sequence diversity of pistil S-proteins associated with gametophytic self-incompatibility in *Nicotiana alata*. Sex. Plant Reprod. **3**: 88-97

[12] McClure BA, Haring V, Ebert PR, Anderson MA, Simpson RJ, Sakiyama F, Clarke AE (1989) Style self-incompatibility gene products of *Nicotiana alata* are ribonucleases. Nature **342**: 955-957

[13] McClure BA, Gray JE, Anderson MA, Clarke AE (1990) Self-incompatibility in *Nicotiana alata* involves degradation of pollen rRNA. Nature **347**: 757-760

[14] Puissant C, Houdebine L-M (1990) An improvement of the single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. BioTechniques **8**: 148-149

[15] Sambrook J, Fritsch EF, Maniatis T (1989) Molecular Cloning: A Laboratory Manual 2nd Edition. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY

[16] Sanger F, Nicklen S, Coulsen AR (1977) DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**: 5463-5467

[17] Somerville C (1989) *Arabidopsis* blooms. Plant Cell **1**: 1131-1135

[18] Wilson CM (1982) Plant nucleases. Biochemistry and development of multiple molecular forms. *In* MC Ratizzi, JG Scandalios and GS Whitt, eds, Isozymes: Current Topics in Biological and Molecular Research, Vol 6. Alan R. Liss, New York, pp 33-54

CHAPTER 3

RNS2: A SENESCENCE-ASSOCIATED RNASE OF *ARABIDOPSIS* THAT DIVERGED FROM THE S-RNASES BEFORE SPECIATION

This Chapter was originally published in *Proc. Natl. Acad. Sci.*

Reference: Taylor, C.B., Bariola, P.A., delCardayré, S.B., Raines, R.T. and Green, P.J. (1993) RNS2: A senescence-associated RNase of *Arabidopsis* that diverged from the S-RNases prior to speciation. Proc. Natl. Acad. Sci. 90: 5118-5122

ABSTRACT

Several self-compatible species of higher plants, such as *Arabidopsis thaliana*, have recently been found to contain S-like ribonucleases (RNases). These S-like RNases are homologous to the S-RNases that have been hypothesized to control self-incompatibility in Solanaceous species. However, the relationship of the S-like RNases to the S-RNases is unknown, and their roles in self-compatible plants are not understood. To address these questions, we have investigated the *RNS2* gene, which encodes an S-like RNase (RNS2) of *Arabidopsis*. Amino acid sequence comparisons indicate that RNS2 and other S-like RNases comprise a subclass within an RNase superfamily, which is distinct from the subclass formed by the S-RNases. RNS2 is most similar to RNase LE [Jost, W., Bak, H., Glund, K., Terpstra, P., Beintema, J.J. (1991) *Eur. J. Biochem.* **198**, 1-6.], an S-like RNase from *Lycopersicon esculentum*, a Solanaceous species. The fact that RNase LE is more similar to RNS2 than to the S-RNases from other Solanaceous plants indicates that the S-like RNases diverged from the S-RNases prior to speciation. Like the S-RNase genes, *RNS2* is most highly expressed in flowers, but unlike the S-RNase genes, *RNS2* is also expressed in roots, stems and leaves of *Arabidopsis*. Moreover, the expression of *RNS2* is increased in both leaves and petals of *Arabidopsis* during senescence. Phosphate starvation can also induce the expression of *RNS2*. Based on these observations we suggest that one role of RNS2 in *Arabidopsis* may be to remobilize phosphate, particularly when cells senesce, or when phosphate becomes limiting.

INTRODUCTION

Fundamental insights into the relationship between protein structure and function and gene evolution have been gained from the study of members of the pancreatic ribonuclease (RNase) superfamily typified by RNase A [1]. Another RNase family has recently been identified in the plant kingdom [2]. RNases in this family are not homologous to the pancreatic RNase superfamily, but rather share homology with a class of fungal RNases that includes RNase T_2 of *Aspergillus oryzae* [3]. The plant members of this family include the S-RNases, proteins associated with a self-recognition response known as self-incompatibility (SI) in certain species of higher plants. In *Nicotiana alata* and other members of the Solanaceae, pollen carrying a particular allele at the S-locus, which controls SI, are unable to fertilize plants carrying the same S-allele. The mechanism by which S-RNases may participate in the rejection of incompatible pollen is unknown; however, their genes co-segregate with the S-locus [4].

Initially it seemed plausible that the S-RNases were highly specialized polymorphic enzymes, without homologs in self-compatible species. However, several recent studies have shown that this is not the case. Among these are data obtained from PCR experiments performed with cDNA from the self-compatible crucifer, *Arabidopsis thaliana*. Using primers directed against the regions most highly conserved between the fungal and S-RNases, amplification products corresponding to three *Arabidopsis* RNase genes, *RNS1*, *RNS2*, and

RNS3, were identified [5]. Further evidence for S-RNase homologs in self-compatible plants is indicated by the protein sequences of RNase LE, isolated from cultured cells of *Lycopersicon esculentum* (tomato) [6], and RNase MC, isolated from seeds of *Momordica charantia* (bitter gourd) [7]; both share homology with fungal RNases and S-RNases. It has also been shown that two RNases isolated from tomato fruit, Tf1 and Tf2, have a number of biochemical properties in common with the S-RNases [8].

The finding of S-RNase homologs ("S-like" RNases) in self-compatible species indicates that enzymes related to the S-RNases play a fundamental role in RNA catabolism in higher plants. Plants contain a large number of RNase activities that are regulated in response to a variety of stimuli [9, 10]. Several plant RNases have been purified and characterized biochemically, but nothing is known about the corresponding genes. It is therefore difficult to assess whether previously described RNase activities, such as those induced during senescence [11, 12], are encoded by S-like RNase genes. In an effort to gain a better understanding of the role of S-like RNases in plants and their relationships with the S-RNases, we have investigated *RNS2*, the most highly expressed of the *Arabidopsis RNS* genes [5].

MATERIALS AND METHODS

Isolation and Sequencing of *RNS2* cDNAs

The *Arabidopsis* cDNA library in Lambda Zap [13] that was initially used to detect the *RNS* genes by PCR amplification [5] was used to screen for *RNS2* cDNA clones by plaque hybridization [14] using the *RNS2* PCR product as a probe. Positive plaques were purified, the cDNA clones converted into plasmid form [15] and sequenced with Sequenase (United States Biochemical), using the chain-termination method [16].

Expression of RNS2 in Yeast

A yeast expression system developed previously for the expression of RNase A (delCardayré, S.B. Quirk, D.J., Rutter, W.J., and Raines, R.T., submitted) was used in conjunction with a substrate-based gel assay to confirm the RNase activity of RNS2. Briefly, an *RNS2* cDNA covering the entire coding region including the signal sequence was inserted between the yeast *PHO5* promoter and *GADPH* terminator [17], and the resulting *PHO5-RNS2-GADPH* gene was then cloned into the yeast shuttle vector pWL. The detailed structure of pWL will be published elsewhere (delCardayré, S.B. Quirk, D.J., Rutter, W.J., and Raines, R.T., submitted). Transformants of *Saccharomyces cerevisiae* strain BJ2168 (*MAT* **a** *prc1-407 prb1-1122 pep4-3 leu2 trp1 ura3-52*) harboring these plasmids were grown in minimal dextrose liquid lacking Trp and containing 0.2 mM KH_2PO_4 (which induces expression from the *PHO5*

promoter) [18], or 11 mM KH_2PO_4 (which does not induce the *PHO5* promoter), for 2 days at 30°C. Following clarification of the cultures by two centrifugation steps (both for two minutes in a microfuge), 5 μL of the resulting supernatant was separated on RNase activity gels [19]. The gels were run, and subsequently treated to visualize RNase activities, essentially as described previously [19].

Multiple Sequence Alignment

The amino acid sequence of RNS2 deduced from the cDNA clones was aligned with the amino acid sequences of fifteen Solanaceous S-RNases [20-25], four fungal RNases [3, 26-28], and two S-like RNases for which protein sequence data are available [6, 7], using the University of Wisconsin Genetics Computer Group program PILEUP with a gap weight of 3.0 and a gap length weight of 0.1 [29]. The RNase sequences were aligned from the positions corresponding to the presumptive mature N-termini of the *N. alata* S-RNases [4]. The gene genealogy was also generated using PILEUP, which produces a similarity score for each possible pair of sequences. The genealogy is a representation of these similarity scores, which are used to order the alignment. The horizontal branch distances in the genealogy are proportional to the similarities between the sequences.

Expression Analyses

Arabidopsis thaliana (L.) Heynh. ecotype RLD was grown under conditions of 12 hours light - 12 hours dark with a relative humidity of 50% at 20°C . Total RNA was isolated from roots, stems, leaves and flowers of 4-5 week old *Arabidopsis* plants using the method of Puissant and Houdebine [30]. For the senescence experiments, flowers were staged based on the morphological characteristics as defined in Smyth *et al.* [31]. Senescing leaves were those showing visible signs of senescence, including chlorosis at the leaf margins and wilting. To starve *Arabidopsis* for phosphate, 1-week-old etiolated seedlings were removed from solid AGM medium (4.3g/L MS salts (Sigma), 30g/L sucrose, 2mg/L glycine, 100mg/L myoinositol, 0.5mg/L pyridoxine, 0.5mg/L nicotinic acid, 0.1mg/L thiamine-HCl, buffered with 2.5mM 2(N-morpholino ethanesulfonic acid), pH5.7) and shaken in liquid media with or without 1.25 mM KH_2PO_4 as described in [32], for 12 hours in the dark. RNA was separated on formaldehyde-agarose gels and transferred to Biotrace HP membrane (Gelman). The northern blots were hybridized as described previously [5] to a 0.7kb gene-specific probe for *RNS2*, corresponding to the EcoRI - XbaI fragment of the longest *RNS2* cDNA clone. This probe, which includes 15 bp of the 5′ untranslated region and the coding region up to nucleotide position 696, does not hybridize to *RNS1* and *RNS3*, and produces an identical pattern of hybridization on genomic Southern blots to that reported for the PCR fragment of *RNS2* ([5]; C.B.T., P.J.G., results not shown). For use as an internal standard in the senescence and phosphate starvation experiments, a

probe for the ubiquitous, highly expressed [33] translation initiation factor eIF4A was generated using PCR. The 380bp *EIF4A* probe was isolated following PCR amplification of the same *Arabidopsis* cDNA library, using oligonucleotide primers corresponding to amino acids 197 - 203 (5′ GGN TT(C/T) AA(A/G) GAN CA(A/G) AT 3′) and to amino acids 317 - 323 (reverse complement: 5′ TC N(C/T)(T/G) CAT (A/G/T)AT N(A/G)(C/T) (A/G)TC 3′) of *Nicotiana tabacum* eIF4A2, which are highly conserved in *N. tabacum* eIF4A3 and related gene products [33]. The PCR reaction was performed using the conditions previously described for the amplification of the *RNS* sequences [5]. The major product of the reaction was blunt-end ligated into a plasmid vector and sequenced [16]. Between the primers used for PCR amplification, the deduced amino acid sequence of the major PCR product shows 96.5% identity with the same region of *N. tabacum* eIF4A2, confirming its identification as an *Arabidopsis* eIF4A sequence (C.B.T., P.J.G., unpublished observations).

RESULTS AND DISCUSSION

RNS2 was initially identified as a PCR product amplified from an *Arabidopsis* cDNA library using primers corresponding to the regions most conserved between the S-RNases and a class of fungal RNases [5]. This PCR product was used as a hybridization probe to isolate *RNS2* cDNA clones from the same library. The nucleotide and deduced amino acid sequences of the longest clone

containing a full-length coding region have been deposited in GENBANK under Accession Number M98336. Analysis of several independent cDNA clones showed that transcripts from the *RNS2* gene can be polyadenylated at three different sites, a feature common to many plant genes [34]. The 19 amino acids at the N-terminus of the RNS2 protein (see legend to Figure 3-1) are typical of a eukaryotic signal sequence, with Ala and Leu making up over half of the residues in this region. This indicates that RNS2 is targeted to the secretory pathway in *Arabidopsis*, as are the S-RNases of the Solanaceae, which are secreted enzymes.

To confirm that RNS2 is indeed an RNase, the coding region of the *RNS2* cDNA, including the putative signal sequence was expressed in *S. cerevisiae* under the control of the *PHO5* promoter which is induced (derepressed) in medium low in phosphate [18]. RNase activity secreted into the culture medium by yeast transformed with the vector control (CON) or the RNS2 expression construct (RNS2), was then detected following electrophoresis on RNase activity gels [19], (see Figure 3-2). Under inducing conditions, two bands of RNS2 activity were observed that have apparent molecular weights of 28 to 33 kDa, slightly higher than the predicted molecular weight of 27 kDa (if the RNS2 signal sequence is cleaved as indicated in the legend to Figure 3-1). This slight difference in molecular weight and the presence of two RNS2 bands may result from processing of the RNS2 signal sequence at multiple sites [35] or differences in glycosylation [36] that are known to affect the mobility of heterologous proteins produced in yeast. It should also be noted that the

RNase activity gels are run under non-reducing conditions [19] which may also contribute to the differences. However, both RNase bands are specific to the RNS2 clone and correlate with the induction of the *PHO5* promoter as expected. This demonstrates that the *RNS2* gene encodes an active RNase.

Comparison of RNS2 to Related RNases

To compare the similarity of the deduced amino acid sequence of RNS2 with those of related plant RNases, the alignment shown in Figure 3-1 was generated as described in *Materials and Methods*. The alignment demonstrates that the similarity of RNS2 to the S-RNases is dispersed throughout the coding region (light shading in Figure 3-1). Moreover, each of the five regions most conserved among the S-RNases (numbered C1 - C5 by Kao and co-workers [21, 37], and boxed in Figure 3-1), is also evident in RNS2. In contrast to the extensive similarity of RNS2 with the S-RNases, its similarity to the related fungal enzymes is highest in the central region of the protein (positions 37 - 110), but is markedly less pronounced towards the N- and C-termini (see asterisks above sequence in Figure 3-1). Twenty-five residues are absolutely conserved among all the RNases of this superfamily. The most prominent of these are the His residues at positions 41 and 106 in Figure 3-2 that have been shown to be important for catalysis in RNase T2 [3]. Others include the pairs of Cys residues at positions 56 and 109, and 174 and 214 that have been shown to form disulfide bonds that are critical for maintaining the structure of RNase M of *Aspergillus saitoi* [28]. Trp 80, which has been proposed to

Legend to Figure 3-1. Alignment of S- and S-like RNase Amino Acid Sequences

The alignment was performed as described in *Materials and Methods* using the U.W.G.C.G. package [29]. Sequences are of RNS2 of *Arabidopsis*, LE from *L. esculentum* [6], MC from *M. charantia* [7], *N. alata* S-alleles 2, 3 and 6 (sequences labelled 2Nic, 3Nic and 6Nic [20]) and 1, F11 and Z (sequences labelled 1Nic, F11 and Z [21]), *P. inflata* S-alleles 1, 2 and 3 (sequences labelled 1Pet, 2Pet, 3Pet [22], *S. chacoense* S-alleles 2 and 3 (sequences labelled 2Sol and 3Sol [23]), *P. hybrida* S-alleles 1, 2 and 3 (sequences labelled Ps1B, Ps2A, Ps3A [24]), *S. tuberosum* S1 allele (sequence labelled 1Stb [25]). The RNases are aligned from the presumptive mature N-termini of the *N. alata* S-alleles, as in [4]. Light shading indicates residues in any sequence that are identical or functionally identical to a residue at the same position in the RNS2 sequence, ellipses denote residues that are identical or functionally identical in the three S-like RNases, RNS2, LE and MC, and heavy shading denotes residues that are identical or functionally identical in at least twelve of the fifteen S-RNases. Conserved regions C1 - C5 of Ioerger *et al.* [37] are boxed. Asterisks above the RNS2 sequence denote residues in the fungal RNases T2 [3], Rh [26], M [27] and Trv [28] that are identical or functionally identical to the RNS2 residues. Functionally identical residues are grouped as: A,S,T; I,L,M,V; H,K,R; F,W,Y; D,E; Q,N. The deduced amino acid sequence of RNS2 includes the following residues that precede the sequence shown in the alignment: MASRLCLLLLVACIAGAFA↓GDVIELNRSQR. The arrow indicates the most likely site for cleavage of the signal sequence, based on a statistical comparison of amino acid sequences in the vicinity of known cleavage sites [38].

The sequence of the full length RNS2 cDNA clones were obtained by Pauline Bariola

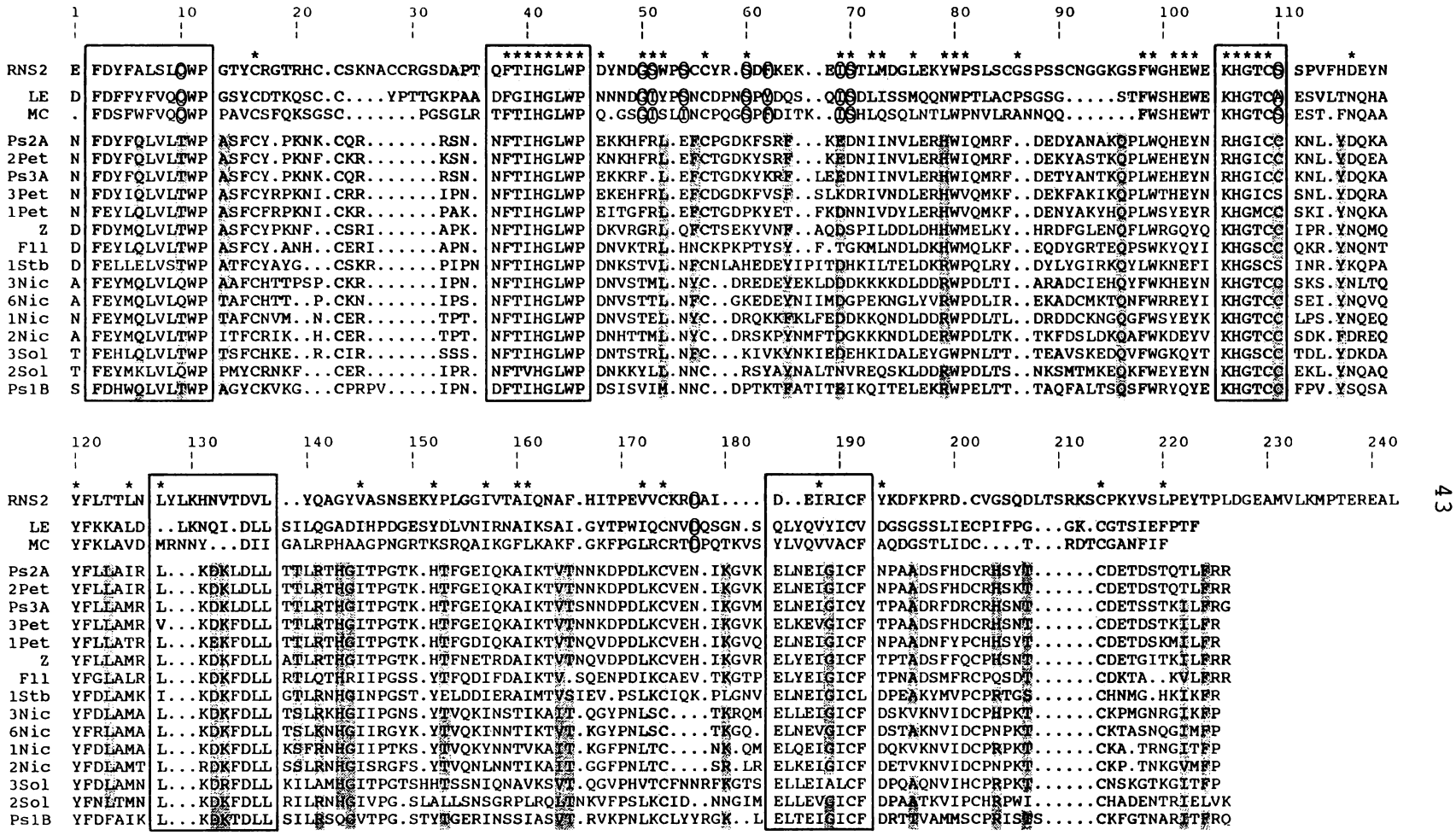

Figure 3-1. Alignment of S- and S-like RNase Amino Acid Sequences

contribute to the broad substrate specificity of RNase T2 [3], is also invariant.

Residues at a number of positions are conserved in mutually exclusive sets, either only among the S-RNases or only among the S-like RNases. Many of the residues that distinguish these subclasses are clustered and therefore correspond to regions within the S-RNases and the S-like RNases that are potentially related to their disparate functions. The residues that are conserved among the S-like RNases (RNS2, LE and MC), but absent from the S-RNases, are highlighted by ellipses in Figure 3-1. Many of these fall between the His residues of the putative active site and most are also conserved in all of the fungal enzymes, indicating that the S-like RNases may share important catalytic properties with the fungal enzymes. Although all members of this class of RNase that have been tested are able to cleave RNA nonspecifically, they do exhibit some distinct substrate preferences. RNases LE and all the fungal enzymes cleave RNA preferentially 3′ of purine nucleotides. However, RNase LE has a somewhat higher preference for guanosine [39], whereas the fungal RNases appear to prefer adenosine [3, 26-28]. The sequence similarities between the fungal and S-like RNases in this region could therefore contribute to this substrate preference. Interestingly, the S-RNases that have been tested tend to cleave preferentially after cytosines or uridylate residues [2, 40]. Some of the residues that are highly conserved among the S-RNases but absent from the S-like RNases are clustered between positions 132 and 165 (see dark shading in Figure 3-1). This region may constitutea novel domain required for

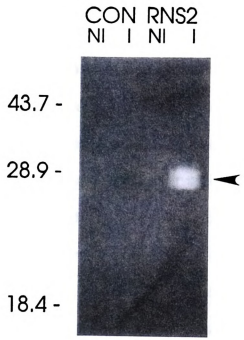

Figure 3-2. Expression of *RNS2* in Yeast

Yeast cultures transformed with the control pWL (CON) or RNS2 (RNS2) constructs were grown under conditions that do not induce (NI) or do induce (I) transcription from the *PHO5* promoter. Supernatants from these cultures were run on RNase activity gels as described in *Materials and Methods*. Positions of molecular weight markers (kDa) are shown to the left of the gel. The arrowhead indicates the RNS2 activity bands.

The system for the heterologous expression of RNS2 in yeast was developed by Pauline Bariola, with vectors constructed by Stephen delCardayré and Ron Raines.

the specialized functions of the S-RNases in SI. At present it is not possible to test this hypothesis directly, as a functional test for S-RNase action in SI has not yet been developed.

The most striking difference between the amino acid sequence of RNS2 and those of all the other enzymes is the carboxy-terminal extension of 20 amino acids. This sequence has features in common with C-terminal vacuolar-targeting signals from other plant proteins, such as lectins and seed storage proteins [41]. Although a minimal consensus sequence for vacuolar targeting in plants has not yet been defined, proven C-terminal vacuolar targeting signals contain a preponderance of hydrophobic amino acids, especially in stretches of three to four [41]. From the Pro at position 225 to the C-terminus at position 242, the sequence of RNS2 includes 8 hydrophobic out of 18 residues, with the sequence Ala-Met-Val-Leu between positions 230 and 233. If RNS2 is targeted to the vacuole, as these sequences imply, this would constitute a significant distinction between RNS2 and the 18 related RNases in Figure 3-1, all of which are thought to be extracellular [4, 6, 7].

A broader illustration of the relationship of RNS2 to the other related RNases was obtained by constructing a gene genealogy (see Figure 3-3) based on the deduced amino acid sequences in Figure 3-1. This genealogy initially divides the RNase superfamily into two lineages, the fungal RNases and the plant RNases. The most significant feature, however, is the placing of the S-like and the S-RNases into two discrete lineages, an observation that indicates

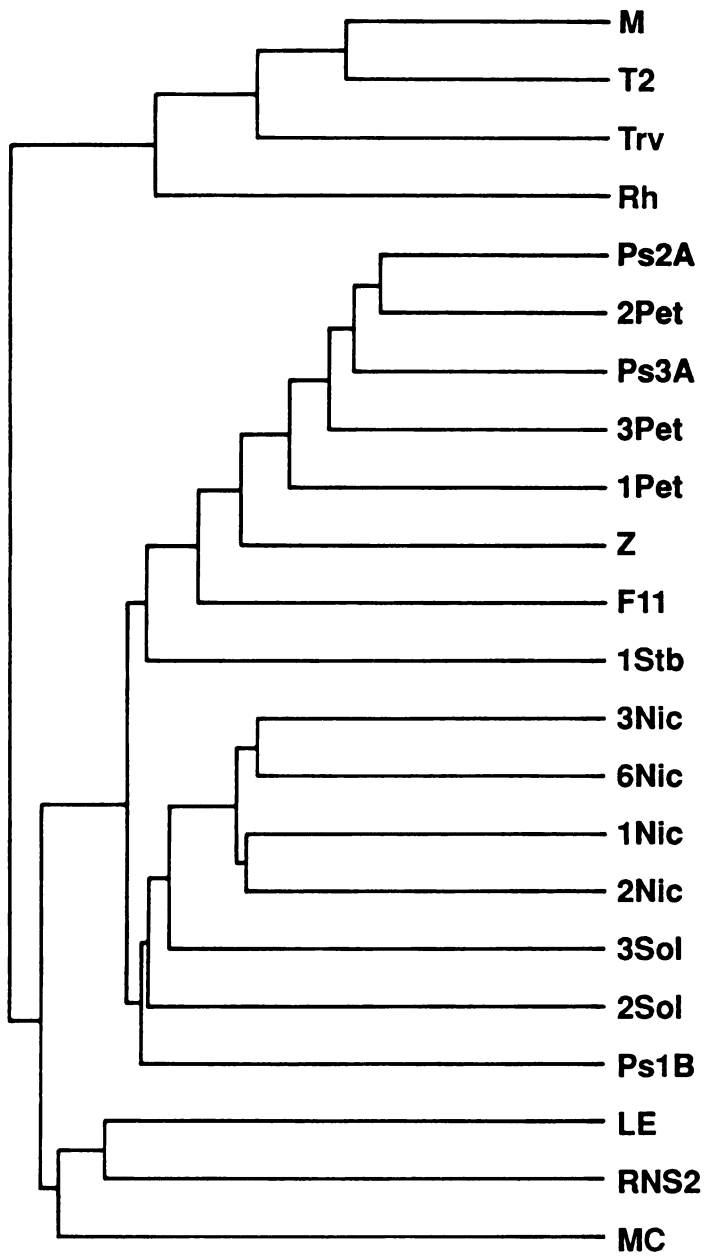

Figure 3-3. Gene Genealogy of the RNase Superfamily

The genealogy was generated from the alignment depicted in Figure 3-1 as described in *Materials and Methods*. Horizontal branch lengths represent overall similarity between the RNases. Designations for each RNase are as in Figure 3-1. T2, Rh, M and Trv are the fungal RNases (refs [3], [26], [27], [28] respectively).

that they form distinct categories of RNases in plants. This arrangement is consistent with current models proposing a specialized function for the S-RNases in SI [4]. Had the S-RNases and S-like RNases been intermingled in the genealogy, a specialized role for the S-RNases would have been more difficult to reconcile. The grouping of the three S-like RNases is also of note because it implies that their sequences, and thus presumably their function, may be evolutionarily conserved across a broad range of plant species. This is most clearly illustrated by RNase LE from *L. esculentum*, a Solanaceous species. RNase LE is placed on the same branch as RNases RNS2 and MC, and is therefore more closely related to these RNases than it is to any of the S-RNases (Figure 3-3) which were isolated from other Solanaceous species. Moreover, preliminary sequence data (P.A.B., C.B.T. and P.J.G., unpublished) indicate that RNS3, one of the other RNases of *Arabidopsis*, is more similar to RNase LE than it is to any of the other RNases, including RNS2. The placing of the S-like RNases on a separate branch of the genealogy from the S-RNases, and the close relationship of LE to other S-like RNases from non-Solanaceous species, strongly indicate that these two groups of RNases diverged prior to speciation. These conclusions greatly extend the information gleaned from the gene genealogy reported previously by Ioerger *et al.* [42] for the S-RNases. This earlier genealogy did not include any S-like RNases, and thus did not address the relationships between S- and S-like RNases. However the genealogy of Ioerger *et al.* [42] included most of the S-RNases that were analyzed in the

present study, and the genealogy in Figure 3-3 is consistent with the evolutionary relationships proposed previously [42] for those enzymes.

Control of *RNS2* Expression

As a first step toward elucidating the function of RNS2, the expression of *RNS2* in various organs of *Arabidopsis* was investigated. RNA was isolated from roots, leaves, stems and flowers of *Arabidopsis* plants, and *RNS2* expression was analyzed by northern blotting (Figure 3-4A). Similar to the S-RNases of the Solanaceae [43], *RNS2* is most highly expressed in flowers of *Arabidopsis*. However, *RNS2* is also expressed in other organs, notably leaf and stem, albeit at a much lower level than in flowers. The expression of *RNS2* in all organs that were examined implies that RNS2 is a fundamental component of the RNA degradation machinery in *Arabidopsis*. To localize further *RNS2* expression in flowers of *Arabidopsis*, RNA was isolated from *Arabidopsis* pistils (stigma and style) and petals harvested at anthesis. The data shown in Figure 3-4B demonstrate that *RNS2* is expressed in both of these flower organs, with slightly higher expression apparent in petals. This is in contrast to the expression of the S-RNases, which is restricted to the gynoecium, and most prominent in the transmitting tissue of the style [43]. Thus, in analogy with the sequence data described above, there are some distinct similarities between the expression properties of *RNS2* and the S-RNases, but also some significant differences.

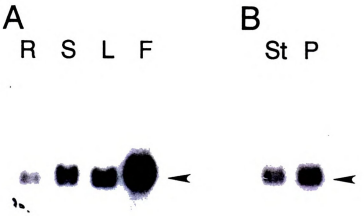

Figure 3-4. Expression of *RNS2* in Different Organs of *Arabidopsis*

(A) Samples of 12μg of total RNA isolated from roots (lane R), stems (lane S), leaves (L) and flowers (F) were hybridized to the RNS2 probe following northern blotting. (B) Samples of 5μg of total RNA isolated from style and stigma (St) and petals (P) dissected from *Arabidopsis* flowers were subjected to northern blotting and hybridization to the *RNS2* probe. The *RNS2* transcript is indicated by the arrowhead.

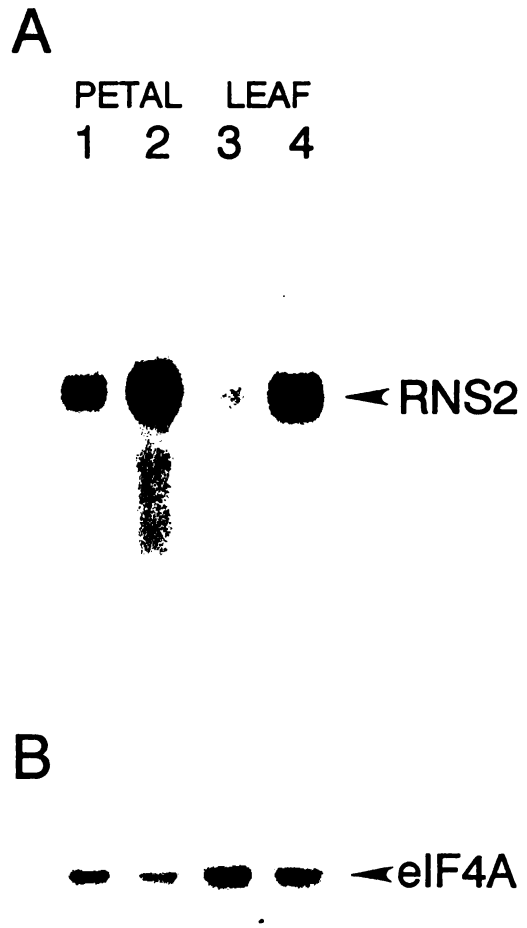

Figure 3-5. Induction of the *RNS2* Transcript During Senescence in *Arabidopsis*

(A) Samples of 5μg of total RNA isolated from nonsenescing (lanes 1 and 3) and senescing (lanes 2 and 4) petals and leaves, as indicated above the lanes, were hybridized to the *RNS2* probe following northern blotting. The *RNS2* transcript is indicated. (B) The northern blot shown in Figure 3-5A was stripped of the *RNS2* probe following the manufacturer's protocol, and hybridized with an *EIF4A* probe (see *Materials and Methods*). The arrowhead shows the *EIF4A* transcript.

The presence of *RNS2* transcripts in petals indicates that RNS2 may contribute in part to the increase in RNase activity that is known to occur in petals during senescence in plants [11]. RNA was therefore isolated from *Arabidopsis* petals harvested at anthesis (stages 13 and 14 in [31]) and during senescence (stage 16 in [31]), and probed for the *RNS2* transcript. A clear increase in *RNS2* expression in senescing petals was observed in these experiments, as shown in Figure 3-5A (lanes 1 and 2). The increase in *RNS2* expression during senescence is even more evident in leaves, where the basal level of the *RNS2* transcript is lower than in petals (Figure 3-5A, lanes 3 and 4; see also Figure 3-4A, lane L). As a control, the blot shown in Figure 3-5A was stripped of the *RNS2* probe and hybridized with a probe for the translation factor eIF4A from *Arabidopsis* (see *Materials and Methods*). The levels of the *EIF4A* transcript are approximately equal in each pair of samples (Figure 3-5B). Moreover, the level of the *Arabidopsis CAB-1* transcript, which encodes the chlorophyll a/b binding protein [44], decreased in leaves during senescence (C.B.T., P.J.G., results not shown). These results confirm that the senescence-induced accumulation of *RNS2* mRNA is a specific effect. RNS2 is therefore likely to be a component of the major change in gene expression that is associated with the onset of senescence [45]. This includes the induction of a large number of hydrolytic enzymes [46], which are thought to be involved in the recycling of nutrients from the vegetative to the reproductive organs [47].

A role for RNase LE in phosphate rescue has been suggested because it is secreted from tomato cells following phosphate starvation [39; 48]. To test whether phosphate limitation could induce the expression of the *RNS2* gene, RNA was isolated from *Arabidopsis* seedlings that had been placed in phosphate-free medium, or in medium containing 1.25 mM phosphate, for 12 hours. As shown in the northern blot in Figure 3-6A, accumulation of the *RNS2* transcript increases markedly following phosphate limitation, while the level of the *EIF4A* transcript remains constant (Figure 3- 6B). Examples of other enzymatic activities induced by phosphate starvation have been described in plants [49; 50], but in these cases it is unknown whether regulation is exerted at the mRNA or protein levels.

Conclusions and Future Prospects

To our knowledge *RNS2* is the first senescence-associated RNase gene identified in higher plants. Senescence-associated RNase activities have been demonstrated in a number of plant species [11, 12, 51], but it is not known whether they are encoded by S-like RNase genes. Since plants often grow under phosphate-limiting conditions [52], it is not surprising that a ribonuclease gene might be induced to facilitate the recovery of phosphate from dying cells. A vacuolar localization for RNS2 would not be inconsistent with this hypothesis, as small fragments of RNA have been found in this organelle [53]. It is also possible that RNS2 may participate in the remobilization of phosphate

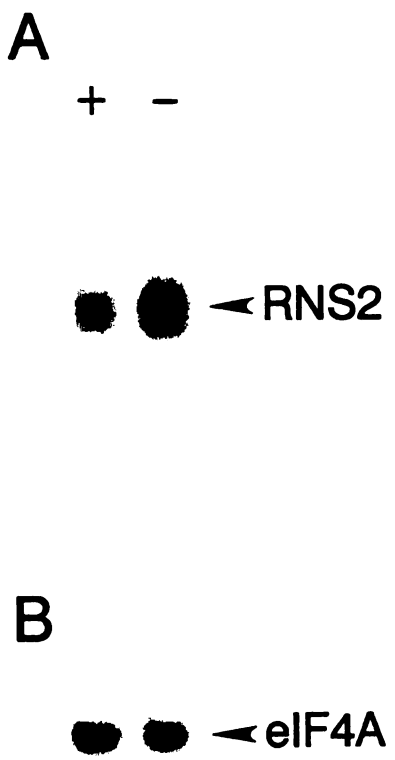

Figure 3-6. Effect of Phosphate Starvation on *RNS2* Expression

(A) Etiolated *Arabidopsis* seedlings were incubated in the presence (+) or absence (-) of phosphate as described in *Materials and Methods*. Samples of 10μg of total RNA isolated from these seedlings were hybridized to the *RNS2* probe following northern blotting. (B) The northern blot in Figure 3-6A was stripped of the *RNS2* probe, and hybridized to the *EIF4A* probe. The *EIF4A* transcript is indicated.

in non-senescing cells, for example during phosphate starvation.

In this study we have demonstrated that RNS2 is an active RNase using a yeast expression system. *In vitro* mutagenesis of RNS2, followed by expression of mutant proteins in this system, should permit direct assessment of the importance of specific residues for the RNase activity of RNS2. This information should be particularly important in elucidating whether the corresponding residues, and the RNase activity, are also required for the action of the S-RNases in SI. At present it is unknown whether RNS2 corresponds to any of the major RNase activities identified previously from *Arabidopsis* [19]. Our initial PCR experiments indicate that *RNS2* is one of at least three closely related RNase genes in *Arabidopsis* [5]. In the future it may be possible to compare the RNS2 sequence to all the related *Arabidopsis* RNases in order to identify regions unique to each enzyme. This will facilitate the production of monospecific antisera that could be used to differentiate among the *RNS* gene products.

ACKNOWLEDGEMENTS

We are most grateful to André Dandridge for expert technical assistance, Brett McLarney for contributing to the *RNS2* sequencing, Dr. Joe White for help with the sequence comparisons and Dr. Steve Kay for the *CAB-1* probe. We thank Linda Danhof for running the RNase activity gel in Figure 3-2, and Dr. Hans Kende and members of the Green Laboratory for comments on the manuscript. This work was supported by grants from the National Science Foundation, the McKnight Foundation and the Department of Energy to P.J.G. P.A.B. was supported in part by a Graduate Fellowship from the National Science Foundation.

LITERATURE CITED

1. Beintema, J.J., Schüller, C., Irie, M., Carsana, A. (1988) *Prog. Biophys. Molec. Biol.* **51,** 165-192.

2. McClure, B.A., Haring, V., Ebert, P.R., Anderson, M.A., Simpson, R.J., Sakiyama, F., Clarke, A.E. (1989) *Nature* **342,** 955-957.

3. Kawata, Y., Sakiyama, F., Tamaoki, H. (1988) *Eur. J. Biochem.* **176,** 683-697.

4. Haring, V., Gray, J.E., McClure, B.A., Anderson, M.A., Clarke, A.E. (1990) *Science* **250,** 937-941.

5. Taylor, C.B., Green, P.J. (1991) *Plant Physiol.* **96,** 980-984.

6. Jost, W., Bak, H., Glund, K., Terpstra, P., Beintema, J.J. (1991) *Eur. J. Biochem.* **198,** 1-6.

7. Ide, H., Kimura, M., Arai, M., Funatsu, G. (1991) *F.E.B.S. Lett.* **284,** 161-164.

8. McKeon, T.A., Lyman, M.L., Prestamo, G. (1991) *Arch. Biochem. Biophys.* **290,** 303-311.

9. Farkas, G.L. (1982) in *Encyclopedia of Plant Physiology N.S.*, eds. Boulter, D., Parthier, B. (Springer-Verlag, Berlin), Vol. 16, pp. 224-262.

10. Wilson, C.M. (1982) in *Isozymes: Current Topics in Biological and Molecular Research*, eds Ratizzi, M.C., Scandalios, J.G., Whitt, G.S. (Alan R. Liss, N.Y.), Vol. 6, pp. 33-54.

11. Matile, P., Winkenbach, F. (1971) *J. Exp. Bot.* **22,** 759-771.

12. Blank, A., McKeon, T.A. (1991) *Plant Physiol.* **97,** 1409-1413.

13. Short, J.M., Fernandez, J.M., Sorge, J.A., Huse, W.D. (1988) *Nucl. Acid Res.* **16,** 7583-7600

14. Sambrook, J., Fritsch, E.F., Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual 2nd Edition* (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY).

15. Delauny, A.J., Verma, D.P.S. (1990) in *Plant Molecular Biology Manual*, eds. Gelvin, S.B., Schilperoort, R.A., Verma, D.P.S. (Kluwer Academic, Dordrecht), pp. A14/1-A14/23.

16. Sanger, F., Nicklen, S., Coulsen, A.R. (1977) *Proc. Natl. Acad. Sci. USA* **74,** 5463-5467.

17. Rosenberg, S., Barr, P.J., Najarian, R.C., and Hallewell, R.A. (1984) *Nature* **312,** 77-80.

18. Thill, G.P., Kramer, R.A., Turner, K.J., Bostian, K.A. (1983) *Mol. Cell Biol.* **3,** 570-579.

19. Yen, Y., Green, P.J. (1991) *Plant Physiol.* **97,** 1487-1493.

20. Anderson, M.A., McFadden, G.I., Bernatzky, R., Atkinson, A., Orpin, T., Dedman, H., Tregear, G., Fernley, R., Clarke, A.E. (1989) *Plant Cell* **1,** 483-491.

21. Kheyr-Pour, A., Bintrin, S.B., Ioerger, T.R., Remy, R., Hammond S,A., Kao, T.-h. (1990) *Sex. Plant Reprod.* **3,** 88-97.

22. Ai, Y., Singh, A., Coleman, C.E., Ioeger, T.R., Kheyr-Pour, A., Kao, T.-h. (1990). *Sex. Plant Reprod.* **3,** 130-138.

23. Xu, B., Mu, J., Nevins, D., Grun, P., Kao, T.-h. (1990) *Mol. Gen. Genet.* **224,** 341-346.

24. Clark, K.R., Okuley, J.J., Collins, P.D., Sims, T.L. (1990) *Plant Cell* **2,** 815-826.

25. Kaufmann, H., Salamini, F., Thompson, R.D. (1991) *Mol. Gen. Genet.* **236,** 457-466.

26. Horiuchi, H., Yanai, K., Takagi, M., Yano, K., Wakabayashi, E., Sanda, A., Mine, S., Ohgi, K., Irie, M. (1988) *J. Biochem. (Tokyo)* **103,** 408-418.

27. Inada, Y., Watanabe, H., Ohgi, K., Irie, M. (1991) *J. Biochem. (Tokyo)* **110,** 896-904.

28. Irie, M., Watanabe, H., Ohgi, K., Harada, M. (1986) *J. Biochem. (Tokyo)* **99,** 627-633.

29. Devereux, J., Haeberli, P., Smithies, O. (1984) *Nucl. Acid. Res.* **12,** 387-395.

30. Puissant, C., Houdebine, L.-M. (1990) *BioTechniques* **8,** 148-149.

31. Smyth, D.R., Bowman, J.L., Meyerowitz, E.M. (1990) *Plant Cell* **2,** 755-767.

32. Tewes, A., Glund, K., Walther, R., Reinbothe, H. (1984) *Z. Pflanzenphysiol* **113,** 141-150.

33. Owttrim, G.W., Hofmann, S., Kuhlemeier, C. (1991) *Nucl. Acid. Res.* **19,** 5491-5496.

34. Dean, C., Tamaki, S., Dunsmuir, P., Favreau, M., Katayama, C., Dooner, H., Bedbrook, J. (1986) *Nucl. Acid. Res.* **14,** 2229-2240.

35. Ohgi, K., Horiuchi, H., Watanabe, H., Takagi, M., Yano, K., Irie, M. (1991) *J. Biochem. (Tokyo)* **109,** 776-785.

36. Tague, B.W., Chrispeels, M.J. (1987) *J. Cell Biol.* **105,** 1971-1979.

37. Ioerger, T.R., Gohlke, J.R., Xu, B., Kao, T.-h. (1991) *Sex. Plant. Reprod* **4,** 81-87.

38. von Heijne, G. (1986) *Nucl. Acid. Res.* **14,** 4683-4690.

39. Nürnberger, T., Abel, S., Jost, W., Glund, K. (1990) *Plant Physiol.* **92,** 970-976.

40. Singh, A., Ai, Y., Kao, T.-h. (1991) *Plant Physiol.* **96,** 61-6834.

41. Chrispeels, M.J., Raikhel, N.V. (1992) *Cell* **68,** 613-616.

42. Ioerger, T.R., Clark, A.G., Kao, T.-h. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 9732-9735.

43. Cornish, E.C., Pettitt, J.M., Bonig, I., Clarke, A.E. (1987) *Nature* **326,** 99-102.

44. Millar, A.J., Kay, S.A. (1991) *Plant Cell* **3,** 541-550.

45. Woodson, W.R. (1987) *Physiol. Plantarum* **71,** 495-502.

46. Borochov, A., Woodson, W.R. (1989) *Hortic. Rev.* **11,** 15-43.

47. Kelly, M.O., Davies, P.J. (1988) *C.R.C. Crit. Rev. Plant Sci.* **7,** 139-173.

48. Löffler, A., Abel, S., Jost, W., Beintema, J.J., Glund, K. (1992) *Plant Physiol.* **98,** 1472-1478

49. Duff, S.M.G., Plaxton, W.C., Lefebvre, D.D. (1991) *Proc. Natl. Acad. Sci. USA* **88,** 9538-9542

50. Usada, H., Shimogawara, K. (1992) *Plant Physiol.* **99,** 1680-1685

51. Woolhouse, H.W. (1982) in *The Molecular Biology of Plant Development*, eds. Smith, H., Grierson, D. (Blackwell, Oxford), pp. 256 - 281.

52. Fried, M., Brosehart, H. (1967) *The Soil-Plant System in Relation to Organic Mineral Nutrition*, (Academic Press, N.Y.).

53. Abel, S., Blume, B., Glund, K. (1990) *Plant Physiol.* **94,** 1163-1171.

CHAPTER 4

DST SEQUENCES, HIGHLY CONSERVED AMONG PLANT *SAUR* GENES, TARGET REPORTER TRANSCRIPTS FOR RAPID DECAY IN TOBACCO

This Chapter was originally published in *Plant Cell*.

Reference: Newman, T.C., Ohme-Takagi, M., Taylor, C.B. and Green, P.J. (1993). DST sequences, highly conserved among plant *SAUR* genes, target reporter transcripts for rapid decay in tobacco. Plant Cell 5: 701-714

ABSTRACT

DST elements are highly conserved sequences located in the 3' untranslated regions (UTRs) of a set of unstable soybean transcripts known as the Small - Auxin - Up - RNAs (*SAUR*s; McClure et al., 1989). To test whether DST sequences could function as mRNA instability determinants in plants, a model system was developed to facilitate the direct measurement of mRNA decay rates in stably transformed cells of tobacco. Initial experiments established that the chloramphenicol acetyltransferase (*CAT*) and β-glucuronidase (*GUS*) transcripts degraded with similar half-lives in this system. In addition, their decay kinetics mirrored the apparent decay kinetics of the corresponding transcripts produced in transgenic plants under the control of a regulated promoter *(Cab-1)*. The model system was then used to measure the decay rates of *GUS* reporter transcripts containing copies of the DST sequence inserted into the 3' UTR. An unmodified *CAT* gene introduced on the same vector served as the internal reference. These experiments and a parallel set utilizing a β-globin reporter gene demonstrated that a synthetic dimer of the DST sequence was sufficient to destabilize both reporter transcripts in stably transformed tobacco cells. The decrease in transcript stability caused by the DST sequences in cultured cells was paralleled by a coordinate decrease in transcript abundance in transgenic tobacco plants. The implications of these results for the potential function of DST sequences within the *SAUR* transcripts are discussed.

INTRODUCTION

Most studies focusing on the regulation of nuclear genes in plants at the RNA level have concentrated on transcriptional control mechanisms. However, the steady - state level of every mRNA is dictated both by its rate of synthesis and by its rate of degradation. Although virtually nothing is known about the mechanisms of mRNA degradation in plants, considerable evidence has been accumulating that indicates that mRNA stability contributes to the control of many plant genes. Walling et al. (1986) were among the first to emphasize the potential importance of post-transcriptional control of transcript accumulation when they showed that differences in transcription rates measured in isolated nuclei could not fully explain observed differences in mRNA level for a number of genes expressed in soybean embryos. Subsequently, many groups have documented similar discrepancies between nuclear run-on transcription and mRNA levels for a variety of plant genes. Examples include the *E17* gene of tomato (Lincoln and Fischer, 1988), the ribulose bisphosphate carboxylase small subunit (*rbcS*) genes in soybean (Shirley and Meagher, 1990) and the *ADH1* gene of maize (Rowland and Strommer, 1986). Because differences in mRNA stability could clearly account for these discrepancies, this explanation is the most common, albeit not the only interpretation of such data (see Thompson and White, 1991).

Additional evidence for post-transcriptional control that may occur at the

level of mRNA stability comes from the finding that several sequences located downstream of transcriptional start sites can affect gene expression. Downstream sequences that function in this manner include the 5′ untranslated region (UTR) of the *Em* gene of wheat (Marcotte et al., 1989), a 230- bp sequence encoding the 5′ UTR and the beginning of the coding region of the *Fed-1* gene of pea (Dickey et al., 1992), and the 3′ ends of several plant genes (Inglelbrecht et al., 1989). It is highly likely that at least some of these sequences act as determinants of mRNA stability.

Direct information about mRNA decay rates in plants is quite limited. The average half-life of an mRNA in animal cells is on the order of several hours (Brock and Shapiro, 1983) and estimates based on experiments done in soybean cells indicate that the half-lives of most plant mRNAs may fall into a similar range (Siflow and Key, 1979). Both the 5′ cap and the 3′ poly(A) tail have been found to stabilize in vitro synthesized transcripts electroporated into plant protoplasts (Gallie et al., 1991). Therefore, these elements may act as general determinants of mRNA stability in plants. However, it is clear that some plant mRNAs decay more rapidly than average despite the presence of a cap and poly(A) tail. The disappearance kinetics of transcripts such as the phytochrome mRNA following red light treatment in oat (Colbert et al., 1985) and the proteinase inhibitor II mRNA following wounding in potato (Logemann et al., 1988) indicate that plants contain unstable mRNAs that decay with half-lives on the order of minutes. Decay rates of unstable transcripts with half-

lives in this range are most commonly measured in mammalian cells following the inhibition of RNA synthesis with a transcriptional inhibitor. Recently, this approach was used to demonstrate that the half-life of the oat phytochrome A (*phyA*) mRNA is about 1 hr in etiolated oat seedlings (Seeley et al., 1992). Degradation of the *phyA* mRNA was not accelerated by treatment with red light indicating that the instability is an inherent property of the mRNA that is independent of light (Seeley et al., 1992).

Because most eukaryotic transcripts appear to be relatively stable, the selective degradation of highly unstable transcripts is generally considered to be an active process (Hunt, 1988). In mammalian cells, unstable transcripts with half-lives of less than 30 min often encode proteins involved in regulating cell growth and differentiation such as lymphokines, cytokines, or proto-oncogene products. The instability of these transcripts is one means by which the cell can facilitate the transient production of the corresponding protein (Peltz et al. 1991). Many unstable transcripts in mammalian cells contain sequences within their 3′ UTRs that are AU-rich and contain one or more copies of the motif AUUUA (reviewed in Cleveland and Yen, 1989). The AU-rich sequences in the GM-CSF (Shaw and Kamen, 1986) and c-*fos* (Shyu et al., 1989) transcripts have been shown to cause mRNA instability when inserted into an otherwise stable transcript (β-globin).

None of the sequences that target transcripts for rapid decay in plants have been reported. However, a good candidate is common to a prominent set

of unstable transcripts encoded by the auxin-inducible *SAUR* (Small - Auxin - Up - RNA) genes of soybean (McClure et al., 1989). Although the exact function of the SAUR proteins is unknown, the temporal and spatial expression characteristics of the *SAUR* genes are generally correlated with auxin-induced cell elongation (McClure and Guilfoyle, 1989; Gee et al., 1991). The half-life of the soybean *SAUR* transcripts has been estimated to be between 10-50 min (McClure and Guilfoyle, 1989; Franco et al., 1990). Unlike unstable mammalian transcripts such as GM-CSF, c-*myc*, and c-*fos*, the 3′ UTRs of *SAUR* transcripts are not particularly AU-rich, and they do not contain multiple AUUUA sequences (McClure et al., 1989). Instead, all known *SAUR* genes contain a novel conserved sequence centered about 30 to 100 bp downstream from the coding region. This downstream sequence, designated the DST element (McClure et al., 1989), is about 40 bp long and consists of three highly conserved sequences separated by two variable sequences (see Figure 4-1). In addition to the DST sequences identified in *SAUR* genes from soybean (McClure et al., 1989), mung bean (Yamamoto et al., 1992), and Arabidopsis (Y. Liu, P. Gil, and P.J. Green, unpublished results), DST-like sequences have also been identified in the 3′ UTRs of two auxin-inducible cDNAs from tobacco (van der Zaal et al., 1991) that are unrelated to the *SAURs* as shown in Figure 4-1. The role of the DST sequence is yet to be resolved, but it has been suggested to control the stability of the soybean *SAUR* transcripts (McClure et al., 1989; Franco et al., 1990).

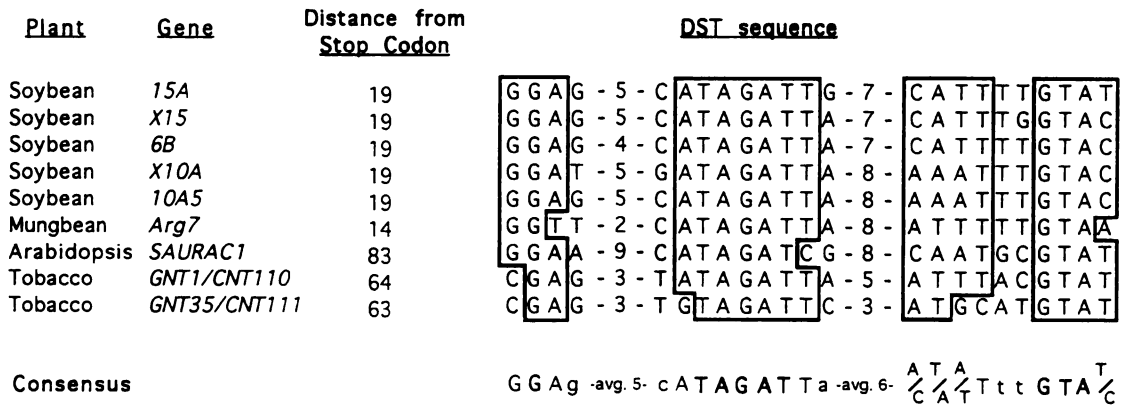

Figure 4-1. Location and structure of DST elements.

DST sequences identified in seven *SAUR* genes and two unrelated auxin-inducible transcripts are aligned. The elements consist of three highly conserved regions (boxed) separated by two variable regions. Within the consensus, invariant bases are indicated in bold and bases conserved in at least seven of nine sequences are capitalized. All elements are located in 3' UTRs at the indicated distances downstream of the stop codon.

To investigate directly the function of DST sequences in plants, we established a model system where rates of mRNA decay could be measured in stably transformed cells of tobacco. Using this system, we found that an insert containing two copies of the DST sequence could confer mRNA instability on both β-glucuronidase (*GUS*) and β-globin reporter transcripts. It is likely that DST sequences are also capable of targeting reporter transcripts for rapid degradation in intact plants because we observed a marked decrease in mRNA accumulation due to the presence of DST sequences in transgenic tobacco.

MATERIALS AND METHODS

Plant Materials, Growth and Transformation

NT-1 Cells

Untransformed *Nicotiana tabacum* cv Bright Yellow 2 (NT-1) cells (An, 1985) were grown at 28°C on a rotary shaker (150 rpm) in NT medium (MS salts (Sigma; St. Louis, MO), 30 g/L sucrose, 3 μM thiamine, 0.56 mM myo-inositol, 1.3 mM KH_2PO_4, 1 μM 2,4-Dichlorophenoxyacetic acid (2,4-D), buffered with 2.5 mM 2[N-morpholino]ethanesulfonic acid (MES), pH 5.7), and were subcultured using 5% inoculum every 7 days. To facilitate transformation, constructs were introduced into Agrobacterium strains GV3111SE or LBA4404 by triparental mating (Fraley et al., 1985), or by electroporation under

conditions recommended by the manufacturer (BioRad; Melville, NY: 25 μF, 2.5 kV, 200 Ω). Agrobacterium strains containing the construct of interest were used to transform NT-1 cells as described previously (An, 1985) with minor modifications. Briefly, three days after subculture, acetosyringone was added to the NT-1 culture to a final concentration of 20 μM and the cells were abraded by repeated pipeting (about 30 times in a 10 mL pipet). Four mLs of cells were co-cultivated with 75-100 μl Agrobacterium for 3 days. The cells were washed 3-4 times in 50 mL NTC (NT medium containing 500 μg/mL carbenicillin), followed by sedimentation at 200 x g for 4 min to remove the Agrobacterium. Following resuspension in 5-7 mL of NTC, cells were plated on 2-3 NTKC (NT medium containing 100 μg/mL kanamycin and 500 μg/mL carbenicillin) plates solidified with 0.7% phytagar (Gibco/BRL; Gaithersberg, MD). After 3-4 weeks, individual transformed calli were transferred to fresh plates or to liquid NTKC and cultured at 28°C. Transformed lines were screened for expression of the reporter genes by ELISA (for chloramphenicol acetyltransferase (CAT) expression; 5 Prime - 3 Prime; Boulder, CO.) or enzymatic assay (β-glucuronidase (GUS); Jefferson et al., 1986) and subsequently by RNA gel blotting.

Transgenic tobacco plants

N.tabacum SR-1 plants used for transformation were grown axenically on MS media (MS salts (Sigma), 30 g/L sucrose, 0.56 mM myo-inositol, buffered to

pH 5.7 with 2.5 mM MES, plus 0.8% phytagar). Leaves from these plants were transferred to No. 3 medium (MS salts, 30 g/L sucrose, 1.2 μM thiamine, 0.56 mM myo-inositol 1 μM indole-3-acetic acid (IAA), 10 μM benzylaminopurine (BAP), buffered to pH 5.6 with 2.5 mM MES plus 0.65% phytagar) and infected with desired Agrobacterium strains using syringe needles. After 3-4 days, the leaves were transferred to No. 3 medium, containing 200 μg/mL kanamycin and 500 μg/mL carbenicillin. Emerging shoots were transferred to MS medium containing 100 μg/mL kanamycin and 500 μg/mL carbenicillin to induce rooting. Plantlets testing positive for expression of *CAT* or *GUS* reference genes were transferred to soil, and grown at under conditions of 16 hr light/ 8 hr dark, 27°C, 75% humidity. For analysis of steady state RNA levels (Figure 4-9), plants were grown to the 10-14 leaf stage, and a fully expanded leaf from the third pair from the apex was harvested and frozen in liquid nitrogen prior to extraction of RNA. For the chlorophyll a/b binding protein (*Cab-1*) promoter time-course experiments (Figure 4-4), transgenic T1 seed were surface sterilized and plated on MS medium containing 100 μg/mL kanamycin and 500 μg/mL carbenicillin at a density of 50-70 seeds per plate. Seedlings were grown under conditions of 16 hr Light (26°C) and 8 hr dark (23°C) and the time-course was carried out 17-18 days after plating. At each time-point, seedlings from one plate were harvested and frozen in liquid nitrogen.

Gene Constructions

pMON505-70, containing *CAT-E9* and *GUS-3C* genes under the control of the 35S promoter (-940 - +9), has been described previously (Fang et al., 1989). A pMON505 (Rogers et al., 1987) derivative containing *CAT-E9* and *GUS-3C* genes under control of the full *Cab-1* promoter (-4000 - +31; Nagy et al., 1987) was constructed by insertion of *Cab-1-CAT-E9* and *Cab-1-GUS-3C* genes into the SmaI site and into the polylinker (between the SacI and ClaI sites), respectively. This construct was designated p844. All other constructs for expression in NT-1 cells and transgenic plants were made from derivatives of pMON505 containing either a 35S-*CAT-3C* or 35S-*GUS-3C* reference gene, (designated p847 and p851, respectively), inserted into the HpaI site of pMON505 oriented towards the right border. Test genes were constructed in pBluescript SKII$^+$ (Stratagene; La Jolla, CA) or pUC (New England Biolabs; Beverly, MA) intermediate vectors that had the following cassette structure: SacI-35S-BglII-XbaI-*GUS*(or β-globin)-BamHI-*E9*-ClaI. All putative instability determinants and control sequences were inserted into the unique BamHI site in the 3'UTR approximately 250 bp upstream of the *E9* poly(A) sites (Fang et al., 1989). Each test gene was moved into the appropriate pMON505 derivative (p847 for *GUS* and p851 for globin test genes, respectively) as a SacI-ClaI fragment that was inserted between the corresponding sites of the polylinker.

The human β-globin coding region was derived from pSP6βc (Lang and Spritz, 1987) and comprised sequences from the NcoI site at the ATG to the

MseI site 30 bp downstream of the stop codon. Before construction of the globin test gene, the BamHI site within the β-globin coding region was filled in and an SphI linker (GGCATGCC) was inserted to restore the proper reading frame. The DST sequence used in this study was identical to that of the soybean *SAUR-15A* gene (McClure et al., 1989) and was synthesized as follows: AGATCTAGGAGACTGACATAGATTGGAGGAGACATTTTGTATAAT AGGATCC. The first and last six bp consisted of BglII and BamHI sites, respectively, to facilitate cloning in the orientation shown, into the BamHI site in the 3'UTRs of the test genes. The following polylinker sequence was used as the spacer control: GGATCCGATCTAGAGTCGACCTGCAGGCATGCAAG CTATCTCTAGAAGATCT.

RNA Manipulations

RNA Isolation

In most cases, RNA was isolated from 2-5 g of NT-1 cells, or transgenic plant tissue ground in liquid nitrogen by a modification of the method of Puissant and Houdebine (1990) as follows: Ground frozen tissue was resuspended in 5 mL of GTC buffer (4M guanidinium thiocyanate, 25 mM sodium citrate, 0.5% sarkosyl and 0.1 M β-mercaptoethanol). 0.5 mL of 2 M sodium acetate, pH 4, 5 mL of phenol, and 1 mL of chloroform were added sequentially followed by vortexing. Samples were centrifuged at 10,000 x g for 10 min and RNA was precipitated from the aqueous phase with 5 mL isopropanol. Following

sedimentation for 10 min at 10,000 x g, pellets were resuspended in 1.5 mL of 4 M lithium chloride and RNA was sedimented for 10 min in a microfuge. After a second lithium chloride precipitation, pellets were resuspended in 0.75 mL of 0.5% SDS, 10 mM Tris-HCl, pH 7.5, 1mM EDTA, extracted with an equal volume of chloroform/isoamyl alcohol (24:1) and precipitated with 0.2 M sodium acetate, pH 5, and 0.6 mL isopropanol. For some of the initial half-life measurements, and for the *Cab-1* promoter time-course, RNA was isolated according to the method of Nagy et. al., (1988b). Both RNA isolation methods yielded equivalent results.

RNA gel blot analysis

RNA samples were separated by electrophoresis in 2% formaldehyde/ 1% agarose gels run in 1X MOPS buffer (20mM 3-[N-Morpholino]propanesulfonic acid (MOPS), 5mM sodium acetate, 1mM EDTA, pH 7.0) and transferred to Biotrace HP membrane (Gelman; Ann Arbor, MI). RNA gel blots were prehybridized and hybridized as described previously (Taylor and Green, 1991). Probes spanning the coding regions of the *CAT*, *GUS* and globin transcripts, were isolated from the intermediate cassette vectors described above as XbaI (or BglII) - BamHI fragments and were labeled with ^{32}P by random priming. Because of the small size of the globin transcripts, blots from globin constructs were probed with globin and quantitated as described below, before probing with *GUS* to minimize background. The heat shock protein 70 probe was the

4.8 kb EcoRI fragment of pHSP4.8 (Wu et al., 1988). The β-ATPase probe corresponded to a 1.5 kb NcoI - SalI fragment covering 93% of the cDNA of the *Atp2-1* gene of *Nicotiana plumbaginifolia* (Boutry and Chua, 1985).

Half-life Determination

Half-life determinations were performed on stably transformed NT-1 cell lines 3-4 days after subculture. ActD was added to the cultures to a final concentration of 100 μg/mL and 10 mL samples were removed from the culture every 30 min for 2-2.5 h. Each sample was immediately sedimented at 700 x g for 1 min, and frozen in liquid nitrogen. Following analysis of the RNA on RNA gel blots as described above, signals were quantitated using a Betascope (Betagen) or PhosphorImager (Molecular Dynamics; Sunnyvale, CA). Data analyzed on both instruments yielded identical results. The values for each time course were subjected to linear regression analysis using SigmaPlot (Jandel Scientific; Corte Madera, CA) to calculate mRNA half-lives.

S1 nuclease protection

To map the 3′ ends of test transcripts, S1 nuclease protection analyses were performed using a 690 base single-stranded end-labeled probe from the *rbcS-E9* gene, essentially as described by Fang et al., (1989) except that the annealing reaction contained 70% formamide. These experiments were designed to be qualitative rather than quantitative. For this reason, the amount of RNA for

each sample was adjusted to allow signals of approximately equal intensity to be compared. The amounts of RNA used are as follows: no insert, 10 μg; DST x1, 20 μg; spacer x1, 10 μg; spacer x2, 40 μg; DST x2, 40 μg. To bring the total amount of RNA to 40 μg for all S1 reactions, 30 μg, 20 μg and 30 μg of RNA from untransformed NT-1 cells were added to the no insert, DST x1 and spacer x1 samples respectively. The *rbcS-E9* probe used in these experiments also detects a signal from the reference gene due to its homology to the *rbcS-3C* 3′ end (Fang et al., 1989). As expected based on Figure 4-6 and Table 4-3, the ratio of the test and reference signals is lower for the DST x2 sample than for the other samples in Figure 4-8.

For monitoring the *Cab-1* promoter time-course, single-stranded confluently labeled DNA probes were prepared that covered the 5′end of the *CAT* or the *GUS* coding regions. Specifically the *CAT* probe was 258 bases long and protected 181 bases of the *CAT* transcript (corresponding nucleotide positions 413 to 575 of the *CAT* gene (Hadfield et al., 1986) plus 16 bases of the 5′UTR located between this sequence and the XbaI site). The *GUS* probe was 190 bases long and protected 117 bases of the *GUS* transcript (corresponding to nucleotide positions 292 to 389 of the *E. coli uidA* gene (Jefferson et al., 1986) plus 16 bases of the 5′UTR located between this sequence and the XbaI site). Non-protected sequences in both the *CAT* and *GUS* probes corresponded to pBlueScript SKII+ sequences between the -20 M13 priming site and the HindII site. S1 nuclease protection experiments were

carried out as described above using 40 μg of RNA from the *Cab-1* time course samples and 50,000 cpm of each probe (labeled as described by Nagy et al., 1987).

^{3}H-Uridine incorporation

Transcriptional inhibitors were added, at the concentrations indicated in Table 4-1, to 3.5 mL of NT-1 cells (3 days after subculture) and incubated with shaking at 28°C for 30 min. 100 μCi of ^{3}H-Uridine (51 Ci/mmol) was then added and incubation was continued for 2 hr. Cells were sedimented for 3 min at 3000 x g, and resuspended in 4 packed cell-volumes of 2mM aurin tricarboxylic acid. Cells were broken by 3 sec of sonication (Heat Systems Ultrasonics; Farmingdale, NY), an equal volume of 2X extraction buffer (600 mM NaCl, 100 mM Tris-HCl, pH 8, 10 mM EDTA, 4% SDS, 20 mM β-mercaptoethanol) was immediately added, and RNA was prepared as described by Nagy et al., (1988b). Poly(A)$^+$ RNA was selected from 200 μg of total RNA on columns containing 100 mg of oligo-dT cellulose, TCA precipitated onto glass fiber filters and quantitated using a scintillation counter. The percent inhibition of RNA synthesis was calculated relative to untreated controls. ActD was added from a 5mg/mL stock solution prepared in 80% ethanol. Control experiments performed with an equivalent volume of ethanol showed no inhibition of incorporation into poly(A)$^+$ RNA (data not shown).

RESULTS

A model system for measurement of mRNA half-lives

To define *cis*-acting sequences that control the stability of mRNAs in vivo, we sought to develop a reliable method for measuring mRNA half-lives in transformed tobacco cells. An established line of *Nicotiana tabacum* cv Bright Yellow 2 (NT-1) cells (An, 1985) was chosen for this purpose because it can be grown as a homogenous suspension in liquid medium (or on plates) and is easy to manipulate. Moreover, NT-1 cells can be stably transformed with foreign genes via Agrobacterium (An, 1985). Our experiments were modeled after those in mammalian cells where mRNA decay rates are most routinely assayed in cultured cells following treatment with a transcription inhibitor. As shown in Table 4-1, several transcription inhibitors were tested for their ability to inhibit the incorporation of ^{3}H-uridine into poly(A)$^{+}$mRNA in NT-1 cells. Neither α-amanitin (which is known to be an effective inhibitor of RNA polymerase II transcription in isolated nuclei; Cox and Goldberg, 1988) nor cordycepin (3′ deoxyadenosine, the precursor of a chain - terminating ATP analog) had a detectable effect on incorporation in intact cells in these experiments and were not pursued further. Actinomycin D (ActD), which is the most commonly used inhibitor for mRNA stability studies in mammalian cells, was found to effectively inhibit mRNA synthesis in NT-1 cells as well. As was found for Dictyostelium (Chung et al., 1981), nearly complete inhibition in NT-1

Table 4-1. Inhibition of mRNA Synthesis in NT-1 Cells

Chemical treatment[a]				% Inhibition[b]
ActD	α-Ama	Chl	Cor	
-	5	-	-	0
-	-	-	25	0
-	-	10	-	25 ± 3
25	-	10	-	60 ± 1
50	-	10	-	84 ± 2
100	-	10	-	94 ± 3
100	-	-	-	94 ± 1

[a]μg/mL of media; ActD, actinomycin D; α-Ama, α-amanitin; Chl, chloroquine; Cor, cordycepin.
[b]Data for Chl and ActD treatments were derived from at least two experiments, and are expressed ± the Standard Deviation (SD).

These experiments were performed by Tom Newman

cells requires 10 - 20 fold more ActD than in mammalian cells, perhaps due to differences in uptake. Raising the concentration of ActD to 200 μg/mL did not increase the percent inhibition beyond that obtained with 100 μg/mL (data not shown). Chloroquine was also somewhat effective in NT-1 cells; however, the effects of chloroquine and ActD were not additive and maximal inhibition could be achieved with ActD alone (see Table 4-1).

For our initial half-life measurements, NT-1 cells were transformed with a construct designated pMON505-70 (Fang et al., 1989), that contains reporter genes encoding β-glucuronidase (GUS) and chloramphenicol acetyltransferase (CAT), as shown in Figure 4-2. *CAT* and *GUS* were each expressed under the control of the 35S promoter of cauliflower mosaic virus (35S), and polyadenylation signals were provided by the 3′ ends of the pea *rbcS-E9* (*E9*) and *rbcS-3C* (*3C*) genes, respectively (Fang et al., 1989; Fluhr et al., 1986). Stable transformants were isolated and transferred into liquid medium as described in Methods. For mRNA decay studies, ActD was added to exponentially growing cell lines and RNA was isolated from samples withdrawn every 30 min for 2.5 hours. As shown in Figure 4-3, when gel blots of these RNAs were probed for the *CAT-E9* and *GUS-3C* transcripts, both transcripts decayed with apparent first order kinetics as described previously for transcripts in other organisms (e.g. Herrick et al., 1990; Savant-Bhonsale and Cleveland, 1992), and have similar half-lives of about 70-80 min. In contrast, when the same RNA samples were probed for endogenous transcripts encoding β-ATPase

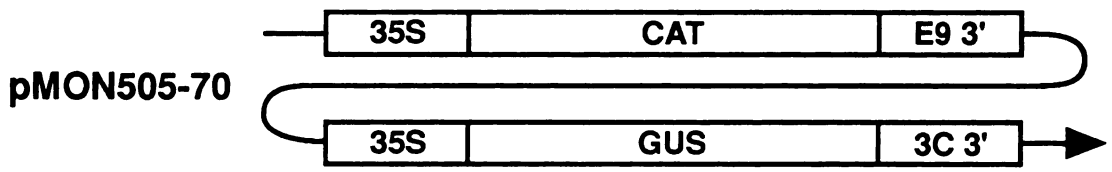

Figure 4-2. Representation of the *CAT* and *GUS* genes following integration of pMON505 into the tobacco genome.

The plant transformation vector pMON505-70 (Fang et al., 1989) contains *CAT* and *GUS* genes controlled by copies of the full-size (-940 to +9) 35S promoter (35S). E9 3′ and 3C 3′ represent the 3′ ends of the pea *rbcS-E9* and *rbcS-3C* genes, respectively. Five kilobases of vector DNA separates the two genes. The solid arrow represents the right T-DNA border. It is located 2 kilobases from the *3C* 3′ end.

Legend to Figure 4-3: Decay of transcripts in tobacco (NT-1) cells stably transformed with pMON505-70.

(A) RNA was isolated from samples harvested every 30 min after transcription was inhibited with 100 μg/mL ActD. RNA gel blots containing 10 μg of total RNA per lane were probed for the transcripts indicated on the left.
(B) Signals from **(A)** were quantitated using a β-scanning device (Betagen), normalized to the zero time point, and subjected to linear regression analysis. Data is from a single experiment but is representative of the experiments listed in Table 4-2.

These experiments were performed by Tom Newman

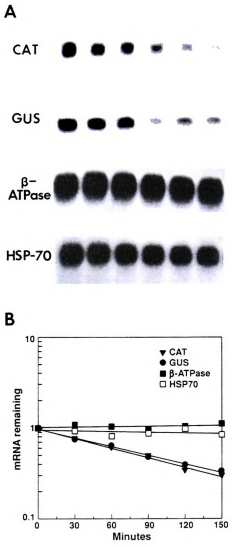

Figure 4-3. Decay of transcripts in tobacco (NT-1) cells stably transformed with pMON505-70.

Table 4-2. Messenger RNA Half-Lives in NT-1 Cells Stably Transformed with pMON505-70

Cell Line	Expt.	Blot	Absolute half-life (mins) GUS -3C	CAT -E9	GUS/ CAT
A	1	1	95	85	1.12
		2	95	100	0.95
	2	1	62	60	1.03
	3	1	90	70	1.29
	4	1	80	57	1.40
		2	73	57	1.28
B	1	1	85	80	1.06
	2	1	83	61	1.36
		2	80	65	1.23
	3	1	80	65	1.23
		2	65	54	1.20
		Mean[a]	80.4	68.9	1.18
		± SD	10.3	13.1	0.14

[a]Values obtained from duplicate blots from the same experiment were averaged prior to calculation of the mean

These experiments were performed by Tom Newman.

(Boutry and Chua, 1985) or heat shock protein 70 (Wu et al., 1988), neither transcript decayed appreciably over the 2.5 hr time course and are likely to have half-lives of greater than 4 hr. To assess the reproducibility of the data, a large number of independent half-life experiments were performed on each of two stably transformed lines as shown in Table 4-2. These data showed that the half-lives of *GUS-3C* and *CAT-E9* were 80.4 ± 10.3 and 68.9 ± 13.1 min, respectively. Although the standard deviation in these determinations was less than 20%, the minimum variation between experiments was observed when the decay rate of *GUS* was expressed relative to that of *CAT* (1.18 ± 0.14). These data suggested that it should be possible to test putative instability sequences for function by inserting them into copies of one of these reporter genes, while using the other as an internal reference.

It was also of interest to address whether the decay characteristics of the *CAT-E9* and *GUS-3C* transcripts in NT-1 cells are reflective of what occurs in transgenic plants. For these experiments, a transformation vector containing *CAT-E9* and *GUS-3C* genes under the control of the wheat chlorophyll a/b binding protein (*Cab-1*) promoter (Nagy et al., 1987) was introduced into tobacco and transgenic plants were regenerated. *Cab-1* is regulated by an endogenous circadian clock in transgenic tobacco, such that the promoter is activated just before dawn and then shuts off about midday (Nagy et al., 1988a). If the *CAT* and *GUS* transcripts decay with similar half-lives in transgenic plants, as the NT-1 cell experiments predict, then they would be

expected to disappear with approximately the same kinetics during the afternoon and evening when produced in transgenic seedlings under *Cab-1* control. This is exactly what we observed as shown in Figure 4-4. Because the *Cab-1* promoter is relatively weak compared to the 35S promoter used in NT-1 cells, we monitored the transcript levels in transgenic plants using a quantitative S1 nuclease protection assay (Kuhlemeier et al., 1987; Fang et al., 1989) to increase the sensitivity of detection (Figure 4-4A). When the transcript levels from 12:00 to 17:00 hr were subjected to linear regression analysis (Figures 4-4B and 4-4C), both *CAT* and *GUS* transcripts disappeared with apparent half-lives of approximately 100 min. These values were similar to, but slightly longer than, the absolute half-lives of the *CAT-E9* and *GUS-3C* transcripts measured in NT-1 cells (70-80 min). The slight discrepancy may be due to minor differences in overall metabolic rates between the two systems and/or to the *Cab-1* promoter not being completely shut-off during the period when the disappearance rates were measured (Giuliano et al., 1988; Millar and Kay, 1991). Nevertheless, the apparent correlation between the NT-1 cell and the transgenic plant data indicates that cultured NT-1 cells should provide an effective and expedient model system to investigate the effect of putative instability determinants such as DST sequences, on mRNA stability in plants.

Legend to Figure 4-4

Disappearance of transcripts encoded by *CAT* and *GUS* genes expressed under the control of copies of the *Cab-1* promoter in transgenic tobacco.

(A) RNA was isolated from 2 week - old transgenic tobacco seedlings (T1) grown on kanamycin-containing medium at the indicated times. Forty micrograms of total RNA was hybridized to single - stranded DNA probes corresponding to the 5′ regions of *GUS* and *CAT* and subjected to S1 nuclease analysis. Protected products corresponding to the *GUS* and *CAT* transcripts are indicated to the left of the gel. Positions of the undigested probe (CP, *CAT* probe; GP, *GUS* probe) are indicated to the right. WT-SR1 represents RNA isolated from untransformed wild-type SR1 seedlings.
(B) The *CAT* protection products from **(A)** were quantitated and normalized to the 9:00 time point. Points from 12:00 to 17:00 were subjected to linear regression analysis.
(C) Same as for **(B)** but the *GUS* protection products were quantitated.

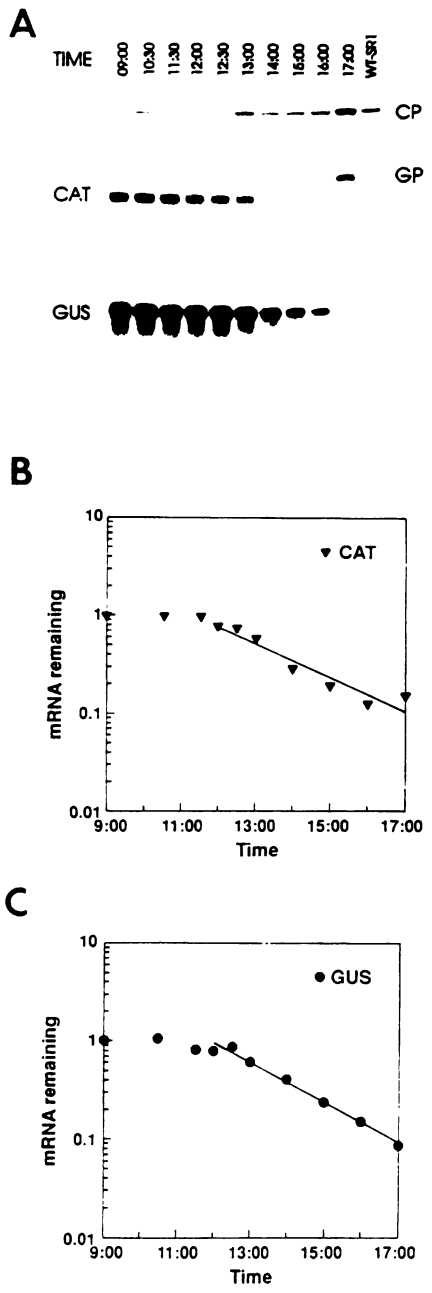

Figure 4-4. Disappearance of transcripts encoded by *CAT* and *GUS* genes expressed under the control of copies of the *Cab-1* promoter in transgenic tobacco.

Effect of DST sequences on mRNA stability in tobacco

To test directly the hypothesis that DST sequences can target transcripts for rapid decay in NT-1 cells, we inserted copies of a synthetic DST sequence into the 3'UTR of a *GUS* "test" gene. Each modified *GUS* gene was inserted into a pMON505 derivative containing an unmodified *CAT* gene (designated p847 as described in Methods). The unmodified *CAT* gene on each vector was designed to serve as an internal reference. The basic structure of the final constructs shown in Figure 4-5 was nearly identical to pMON505-70 (Figure 4-2) used in our pilot studies except that the *3C* 3' end was attached to *CAT* and the *E9* 3' end was attached to *GUS* to facilitate cloning. Because many isolated sequence motifs function better when they are multimerized (Kuhlemeier et al., 1987; Green et al., 1988; Herr and Clarke, 1986), *GUS* genes containing one (DST x1) and two copies (DST x2) of the DST sequence were made. As discussed below, a second set of test genes was made with the human β-globin gene, using an unmodified *GUS* reference gene as the internal standard, as shown in Figure 4-5. Each construct was introduced into NT-1 cells and stably transformed cell lines were isolated. To examine the effects of the inserts on mRNA stability, suspension cultures were established from the transformants, and the degradation of the *GUS* and *CAT* transcripts were monitored on RNA blots following the inhibition of transcription with ActD as outlined above.

The most marked effects on *GUS* mRNA stability resulted from the DSTx2

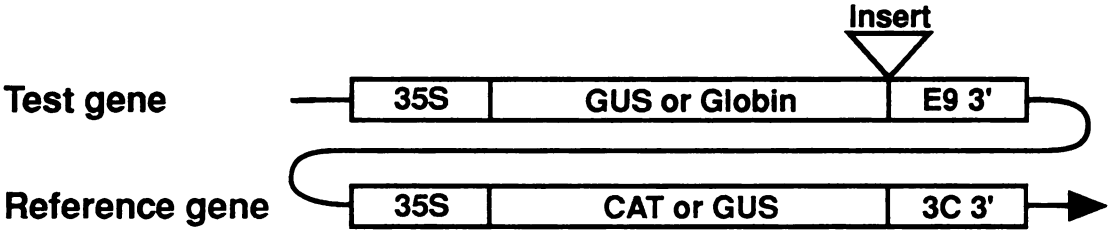

Figure 4-5. Representation of the test and reference genes following integration into the tobacco genome.

GUS and globin test genes used to investigate the effects of DST inserts, and *CAT* and *GUS* reference genes used as internal standards are illustrated. To facilitate their transfer into the tobacco genome in the relative orientations shown, the test and reference genes were cloned into the polylinker and HpaI sites, respectively, of the plant transformation vector pMON505 (Rogers et al., 1987) as described in Methods.

insert that caused a rapid degradation of the *GUS* transcript relative to the no insert control. A representative set of experiments is shown in Figure 4-6. In the transformed cell line expressing the no insert control, the *CAT* and the *GUS* transcripts decayed at approximately the same rate over a two-hour time course. In contrast, the *GUS* transcript containing two copies of the synthetic DST sequence (DST x2) was degraded much more rapidly than the *CAT* reference and the no insert control. The blot in Figure 4-6B was overexposed relative to that in Figure 4-6A because the abundance of the DST x2 transcript is also diminished. From the graphs in Figures 4-6C and 4-6D, the half-life values for the *GUS* transcripts containing no insert and DST x2 were calculated to be 120 and 37 min respectively. mRNA stability was also assayed in additional cell lines transformed with these and other constructs described below. The results of these measurements are shown in Table 4-3 and Figure 4-7. The decreased stability caused by the DSTx2 insert can be seen by comparing the absolute half-life (4.4 - fold decrease), or the relative half-life (3.6 - fold decrease) of *GUS*-DSTx2 to that of the *GUS*-no insert control. However, as expected based on our pilot studies with pMON505-70, the relative half-life values in Table 4-3 showed the least experimental variation because they were normalized to the decay of the internal reference transcripts. That is, experimental variations generally affect the test and reference mRNA half-lives coordinately. Therefore, in most cases, the SD values are lower and more consistent for the relative half-life values, than for the the absolute half-

Legend to Figure 4-6

Destabilization of the *GUS* transcript by DST sequences in stably transformed NT-1 cells.

RNA was isolated from samples harvested from stably transformed cell lines every 30 min for a period of 2 hr after the addition of ActD as described for Figure 4-3. Gel blots containing 10 μg of RNA per lane were probed with *GUS* (test) and *CAT* (reference) probes, and signals were quantitated using a β-scanning device (PhosphorImager).
(A) Gel blot of RNA isolated from a cell line transformed with a construct containing a *GUS*-no insert test gene and the *CAT* reference gene.
(B) Gel blot of RNA isolated from a cell line transformed with a construct containing a *GUS*-DSTx2 test gene and the *CAT* reference gene.
(C) and **(D)** Quantitative representation of the *GUS* test and *CAT* reference transcripts from **(A)** and **(B)** are shown in **(C)** and **(D)**, respectively. The data are derived from single experiments with individual cell lines transformed with each construct, but are representative of the data summarizing multiple half-life measurements for independent cell lines shown in Table 4-3 and Figure 4-7. The blot in **(B)** was overexposed to allow the *GUS*-DST x2 transcript, which was present in reduced amounts, to be visualized.

These experiments were performed by Masaru Takagi and Tom Newman

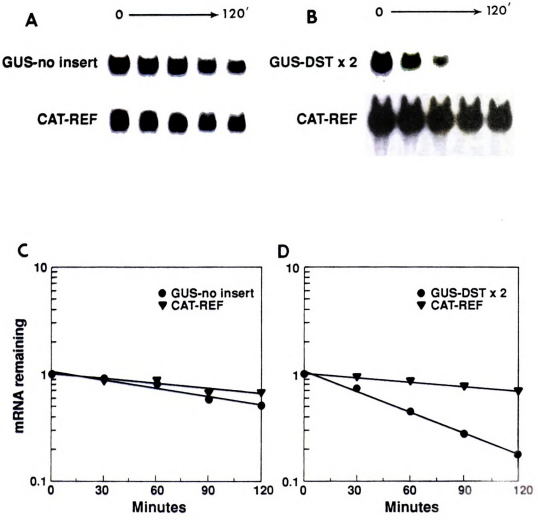

Figure 4-6. Destabilization of the *GUS* transcript by DST sequences in stably transformed NT-1 cells.

Table 4-3. Effect of DST Sequences on mRNA Stability in NT-1 Cells

Construct		Mean absolute half-life (mins)[a]			
Test gene	Ref. gene	Test mRNA	Ref. mRNA	Mean Test/Ref.[b]	Relative half-life[c]
GUS-no insert	CAT-3C	140±57	175±70	0.83±0.23	1.00±0.28
GUS-DST x1	CAT-3C	76±16	142±32	0.55±0.06	0.66±0.07
GUS-spacer x1	CAT-3C	110±15	182±6	0.59±0.04	0.71±0.05
GUS-spacer x2	CAT-3C	84±42	141±68	0.62±0.03	0.75±0.04
GUS-DST x2	CAT-3C	32±9	152±61	0.23±0.09	0.28±0.11
GLOBIN-no insert	GUS-3C	149±34	80±8	1.84±0.24	1.00±0.13
GLOBIN-DST x1	GUS-3C	102±21	82±10	1.26±0.13	0.68±0.07
GLOBIN-spacer x1	GUS-3C	114±33	80±14	1.41±0.15	0.77±0.08
GLOBIN-spacer x2	GUS-3C	113±6	62±5	1.75±0.03	0.95±0.02
GLOBIN-DST x2	GUS-3C	33±2	97±35	0.39±0.14	0.21±0.08

[a]Data were derived from at least two northern blot experiments from each of two to five independent stably transformed cell lines for each construct. Values are Means ± SD, and were calculated as described for the Mean absolute half-life values in Table 4-2.
[b]For each northern blot, the half-life of the Test transcript was divided by that of the Reference (Ref.) transcript. Values obtained from replicate blots from the same experiment were averaged prior to calculation of the Mean Test/Ref.
[c]Relative half-lives were calculated by normalizing the Mean Test/Ref. for each construct to that of the corresponding no-insert control, which was arbitrarily assigned a value of 1.

These experiments were performed by Masaru Ohme-Takagi and Tom Newman

life values in Table 4-3, as they were in Table 4-2. For this reason we have relied more on the relative half-life values which are summarized in Figure 4-7A.

The 3.6 - fold decrease in mRNA stability caused by the DST x2 insert can not simply be explained by the increased size of the UTR in the *GUS*-DST x2 transcript. As shown in Figure 4-7A, insertion of two copies (spacer x2) of a polylinker sequence, the same size as the DST sequence, into the *GUS* 3′ UTR had little effect (1.3 - fold) on the stability of the mRNA. Compared to the marked decrease in stability caused by the DST x2 insert, the effect of one copy of the DST element was modest and not significantly different from the spacer x1 and spacer x2 controls. To examine whether any of the inserts altered the sites of polyadenylation, the sizes of the *GUS* test transcripts from each of the constructs in Figure 4-7A were compared using RNA gel blots. The results showed that all test transcripts were of the expected size (data not shown). In addition, S1 nuclease protection analysis, as shown in Figure 4-8, indicated that the 3′ ends of the test transcripts were identical and corresponded to the well characterized polyadenylation sites of the *rbcS-E9* gene (Fluhr et al, 1986; Mogen et al., 1992). These data indicate that the decreases in mRNA stability caused by DST sequences in tobacco can be attributed to direct effects of the elements, not indirect effects caused by aberrant polyadenylation.

The effects of the DST and spacer inserts on mRNA stability in tobacco cells were also investigated within the context of a β-globin test transcript

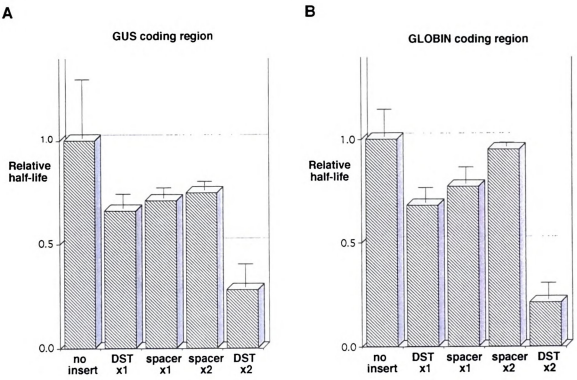

Figure 4-7. Effects of DST and control inserts on the stability of *GUS* and globin test transcripts.

Histograms summarizing the effects of the indicated inserts, on the stability of the **(A)** *GUS* or **(B)** globin test transcripts. Bars represent relative half-lives as calculated in Table 4-3 for each construct.

These experiments were performed by Masaru Ohme-Takagi and Tom Newman

Legend to Figure 4-8: Mapping of the 3′ ends of the *GUS* test transcripts.

A 690 base single stranded probe from the 3′ end of the *rbcS-E9* gene (Fang et al., 1989) was annealed to RNA samples isolated from cell lines expressing each of the *GUS* test constructs and digested with S1 nuclease as described in Methods. The protection products corresponding to the expected 3′ ends of *GUS* transcripts (230 to 249 nt) are indicated by the bracket (GUS 3′). The expected protection products (89 nt) from the *CAT* 3′ end is also indicated (CAT 3′). P indicates the position of the undigested probe. WT NT-1 represents RNA from untransformed wild - type NT-1 cells.

These experiments were performed by Masaru Ohme-Takagi

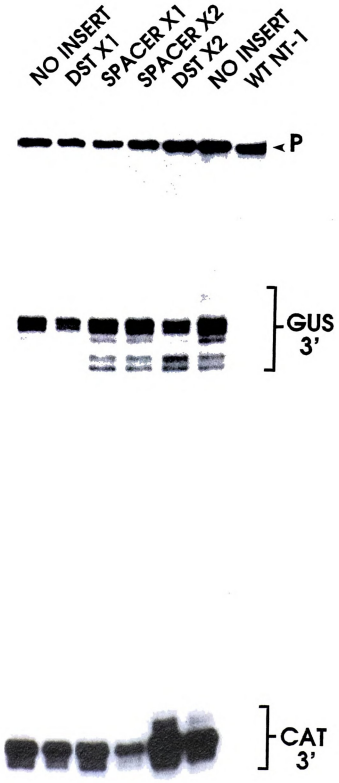

Figure 4-8. Mapping of the 3′ ends of the *GUS* test transcripts.

using an unmodified *GUS* gene as the reference (Figure 4-5). Again multiple half-life measurements were made for independent stably transformed cell lines and the results are shown in Table 4-3 and Figure 4-7B. For each of the inserts, the effects on globin mRNA stability paralleled those obtained with the *GUS* test genes discussed above. Insertion of two copies of the DST sequence destabilized the globin transcript by 4.8 - fold relative to the no insert control. The decrease in stability caused by the DST x1 insert was relatively modest (1.5 - fold) and only slightly greater than the spacer x1 (1.3 - fold) and spacer x2 (1.1 - fold) inserts.

The marked destabilizing effect of the DST x2 insert in transformed tobacco cell lines prompted us to test the function of this element in transgenic tobacco plants. To this end, we compared the transcript levels in transgenic plants transformed with the globin-no ,insert, DST x2, and spacer x2 constructs. In each case, nine to 14 plants were analyzed and data from two representative plants per construct is shown in Figure 4-9. As expected, the globin-DST x2 transcript was the same size as the globin-spacer x2 transcript, and both were slightly larger than the globin transcript containing no insert. A clear effect of the DST x2 insert on the abundance of the globin test transcript was also visible. When the transcripts were quantitated and normalized for the *GUS* reference transcript present in each plant, there was little difference between the accumulation of globin-no insert and globin-spacer x2. In contrast, the level of the globin-DST x2 transcript was about 7.5 - fold lower.

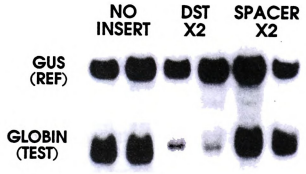

Figure 4-9. Reduced globin mRNA accumulation caused by DST sequences in transgenic plants.

Constructs containing the *GUS* reference gene and either the globin-no insert, globin-DST x2, or the globin-spacer x2 test gene were introduced into tobacco leaf discs via Agrobacterium and transgenic plants were regenerated. RNA samples from the leaves of 9 to 14 independent transgenic plants per construct were analyzed and results from two representative plants for each construct are shown. Gel blots of 10 μg RNA per lane, were hybridized with *GUS* and globin probes and the signals were quantitated using a PhosphorImager.

These data are quite consistent with the effects of the DST x2 (4.8 - fold) and spacer x2 (1.1 - fold) inserts on globin mRNA decay rates in transformed tobacco cell lines (Figure 4-7B). Thus it is highly likely that DST sequences are recognized as mRNA instability determinants in normally growing plants, as they are in cultured cells.

DISCUSSION

Measurement of mRNA decay rates

In this study we described a model system for the direct measurement of mRNA decay rates in stably transformed cells of tobacco. This system allowed us to demonstrate that a dimer of the DST sequence could act as a potent determinant of mRNA instability. For these analyses, we employed a test gene / reference gene approach much like those used to normalize for position effects when transcript levels are compared in transgenic plants (Kuhlemeier et al., 1987; Fang et al., 1989). Position effects do not appear to be a major consideration for half-life measurements because our data and that from other systems indicate that mRNA decay is first order and therefore is independent of initial mRNA concentration. Nevertheless, the use of an reference transcript as a internal standard proved to be the most effective strategy to normalize for cell line - to - cell line variations that may occur due to minor differences in

growth rate, etc. Nearly identical results were obtained with the *GUS* and the globin constructs even though different reference genes were used for each series. This emphasizes the reliability of the reference gene system and indicates that the behavior of the sequences that we tested was independent of the adjacent coding region.

The model system utilized in this study should allow the decay rates of a variety of endogenous and engineered transcripts to be assayed in NT-1 cells. In particular, this system is well suited for measuring the decay rates of short-lived mRNAs because the strong 35S promoter can be used to enhance mRNA synthesis, there is no tissue limitation, and it is convenient to isolate multiple samples over a relatively short period of time. Conversely, the system is less optimal for measuring the decay rates of long-lived mRNAs because this would require prolonged incubation with ActD which can slow or stop mRNA decay at later time points (Shyu et al., 1989; T.C. Newman and P.J. Green, unpublished results). Pulse-chase methods are a good alternative for measuring long mRNA half-lives (Brock and Shapiro, 1983) but are problematic for short half-lives because precursor pools make it difficult to achieve a rapid chase (Siflow and Key, 1979). Approaches that facilitate the analysis of short-lived mRNAs may be particularly useful in plant biology because plants are known to exhibit many rapid changes in gene expression in response to various stimuli (e.g. Theologis, 1986; Logemann et al., 1988; Thompson and White, 1991). If a given mRNA had a half-life of several hours, it would be difficult to rapidly

shut-down synthesis of the corresponding protein by simply repressing transcription. However if the mRNA were very short-lived, decreases (and increases) in transcript levels could be achieved rapidly.

It was reassuring to see that the decay kinetics of *CAT* and *GUS* mRNAs in NT-1 cells were consistent with the disappearance rate of these transcripts under control of the *Cab-1* promoter in plants. In theory, the *Cab-1* promoter system might provide a general means to directly measure mRNA decay rates in transgenic plants, after the midday shut-down of transcription. However, the *Cab-1* promoter is relatively weak and this may limit its utility in practice, particularly when combined with genes encoding very unstable transcripts. In our studies, nuclease protection assays were required to achieve maximal detection of *CAT* and *GUS* transcripts during the afternoon and evening, and the use of highly unstable transcripts would be expected to further exacerbate the situation. Moreover, the use of nuclease protection assays does not allow the full-length mRNA to be monitored, which may be a disadvantage for mRNA decay studies. Perhaps this problem could be overcome by using extremely stable reporter transcripts that do not cross-hybridize with endogenous plant mRNAs. It is also possible that, in its present state, the system could be used to test putative stabilizing sequences because these would be expected to slow the disappearance of a *GUS* test transcript relative to an unmodified *CAT* reference mRNA.

DST sequences as determinants of mRNA instability

In this report we show that the insertion of DST sequences into the 3′ UTR is sufficient to destabilize both the *GUS* and the globin transcripts in tobacco. The absolute half-lives measured for *GUS* and globin transcripts containing the DSTx2 inserts (32+/-9 and 33+/-2 min, respectively) were similar to those measured by Guilfoyle and co-workers for a mixture of *SAUR* mRNAs in elongating soybean hypocotyl segments (Franco et al., 1990). To achieve this marked destabilization of the *GUS* and globin transcripts, the presence of two DST sequences was required, although the *SAUR* genes contain only a single copy of the element in their 3′ UTRs. In this respect our experiments parallel transcription studies where it is not unusual to find that multiple copies of a single regulatory element are necessary for maximal function in a foreign or altered context. For example, in transgenic plants, multiple copies of Box II from the pea *rbcS-3A* gene are required to confer light regulatory properties on derivatives of a plant viral promoter (Kuhlemeier et al., 1987; Lam and Chua, 1990) even though 5′ deletion derivatives of the *rbcS-3A* gene are light regulated with only one copy of Box II present (Kuhlemeier et al., 1987). Within the simian virus 40 core enhancer, the duplication of one domain of the sequence can compensate for loss of another (Herr and Clarke, 1986). In a similar manner, the requirement for two DST sequences may indicate that other sequences in addition to DST contribute to the instability of the *SAUR* transcripts. Transcripts with multiple instability determinants have been

identified in yeast (Peltz and Jacobson, 1993) and mammalian (Peltz et al., 1991) systems. Alternatively, the limited effect of one copy of the DST sequence within the chimeric constructs could be related to its position relative to the poly(A) addition site which is likely to differ from that in the *SAUR* genes. The 3′ ends of the five soybean *SAUR* transcripts have not been reported, but poly(A) addition sites of the Arabidopsis *SAUR-AC1* and Mungbean *ARG7* transcripts were recently identified 10 bases (P. Gil and P.J. Green, unpublished results), and 87 bases (van der Zaal et al., 1991) downstream of the DST sequences, respectively. In our chimeric genes, the DST sequences were inserted about 250 bp upstream of the poly(A) addition site in order to avoid disrupting polyadenylation. Polyadenylation signals are known to be more extensive in plants than in mammalian cells, and in the *rbcS-E9* 3′ end used in our constructs, are known to extend to about 140 bp upstream of the sites of polyadenylation (Mogen et al., 1990).

Although the measurement of mRNA stability following transcription inhibition is wide spread in mammalian systems, these approaches can lead to an over estimation of mRNA half-lives in some cases. Whether or not this occurs, is dependent on the specific mRNA instability determinant that is involved. A well studied example is the 3′ AU-rich (ARE) instability determinant of c-*fos* mRNA. The half-life of a globin transcript containing this element is about 40 min in the absence of ActD and $\geq$ 110 min in the presence of the inhibitor (Shyu et al., 1989). Thus it is possible that the ARE requires on-going

transcription for maximal function. In contrast, ActD has little effect on the instability determinant within the coding region of the c-*fos* mRNA (Shyu et al., 1989). It has been suggested that transcriptional inhibition may dampen the degradation rate of several transcripts in plants (Fritz et al., 1991; Seeley et al., 1992) including the *SAURs* (Franco et al., 1990). However, in the case of the *SAURs*, this suggestion was based on a comparison of half-lives measured in experiments involving other variables in addition to the presence and absence of ActD (Franco et al., 1990; McClure and Guilfoyle, 1989). The DSTx2 element is clearly an active instability determinant when assayed following ActD treatment. At present we can not rule out a modest dampening effect of ActD since we observed a somewhat greater effect of DSTx2 in transgenic plants which were not exposed to the inhibitor. Nevertheless, the correlation between plants and NT-1 cells is quite good considering the obvious physiological differences between the two systems.

All seven *SAUR* genes identified to date have been found to contain DST sequences in their 3′ UTRs (Figure 4-1). Although six of the genes originate from leguminous plants, the presence of a DST sequence in a *SAUR* gene from a rather distantly related species, Arabidopsis, argues that the element is highly conserved among *SAUR* genes in many higher plants. The two tobacco transcripts listed in Figure 4-1 that are unrelated to the *SAURs*, also contain DST sequences but these are not as well conserved (Figure 4-1). In all cases however, the sequences TAGAT in the central region and GTA in the 3′ region

of the DST element are invariant. Additional functional studies will be required to test the relative importance of the conserved regions and their spacing. Taken together, the sequence conservation shown in Figure 4-1, and the instability function attributed to DST sequences in this report suggest that recognition of DST sequences is one mechanism targeting the *SAUR* transcripts for rapid decay. Clearly, mutagenesis of the elements within natural *SAUR* genes will be necessary to test this hypothesis. One property that all of the DST-containing transcripts in Figure 4-1 have in common is that they are induced by the plant hormone auxin (McClure et al., 1989; van der Zaal et al., 1991). For several of the *SAUR* genes, the effect of auxin has been shown to be mediated at least in part at the level of transcription (Franco et al., 1990). While these observations do not eliminate the possibility that auxin also acts to stabilize the mRNAs, direct evidence for such a model is lacking. We have not been able to detect specific stabilization of DST-containing transcripts following treatment of tobacco cells with the synthetic auxin 2,4-D (M. Ohme-Takagi and P.J. Green, unpublished results). Thus if DST sequences function to destabilize the *SAUR* transcripts in higher plants, they may do so constitutively in order to facilitate rapid adjustments in gene expression in response to changes in auxin levels or auxin sensitivities.

In any event, the identification of *cis*-acting sequences that are selectively recognized by a plant's mRNA degradation machinery is an important first step towards elucidating mRNA decay pathways in plants.

Using the DSTx2 instability determinant identified in this report, it should now be possible to identify RNA binding proteins or ribonucleases that specifically recognize this element, similar to what has been done in mammalian, bacterial, and chloroplast systems (Peltz et al., 1991; Belasco and Higgins, 1988; Stern et al., 1989). These studies, together with further analysis of DST function in vivo, should provide fundamental information about the mechanisms targeting transcripts for rapid degradation in higher plants.

ACKNOWLEDGMENTS

We thank Dr. Nam-Hai Chua, Dr. Jeff Ross, Dr. Richard Spritz, and Dr. Chris Somerville for plasmids, Dr. Gynheung An for the NT-1 cell line, Kathy Wilhelm for constructing the spacer control, André Dandridge and Carrie Climer for technical assistance, and Drs. Natasha Raikhel and Michael Sullivan for comments on the manuscript. This work was funded by grants from the United States Department of Agriculture (No. 9001167) and the Department of Energy (No. DE-FG02-90ER20021) to P.J.G. M. O.-T. was supported, in part, by a fellowship from the Human Frontier Science Program.

LITERATURE CITED

An, G. (1985). High efficiency transformation of cultured tobacco cells. Plant Physiol. **79**, 568-570.

Belasco, J.G., and Higgins, C.F. (1988). Mechanisms of mRNA decay in bacteria: a perspective. Gene **72**, 15-23.

Boutry, M., and Chua N.-H. (1985). A nuclear gene encoding the beta subunit of the mitochondrial ATP synthase in *Nicotiana plumbaginifolia*. EMBO J. **4**, 2159-2165.

Brock, M.L., and Shapiro, D.J. (1983). Estrogen stabilizes vitellogenin mRNA against cytoplasmic degradation. Cell **34**, 207-214.

Chung, S., Landfear, S., Blumberg, D., Cohen, N.S., and Lodish, H. (1981). Synthesis and stability of developmentally regulated mRNAs are affected by cell-cell contact and cAMP. Cell **24**, 785-797.

Cleveland, D.W., and Yen, T.J. (1989). Multiple determinants of eukaryotic mRNA stability. New Biologist **1**, 121-126.

Colbert, J.T., Hershey, H.P., and Quail, P.H. (1985). Phytochrome regulation of phytochrome mRNA abundance. Plant Mol. Biol. **5**, 91-101.

Cox, K.H. and Goldberg, R.B. (1988). Analysis of plant gene expression. In Plant Molecular Biology: A practical approach, C.H. Shaw, ed (Oxford: IRL Press), pp. 1-34.

Dickey, L.F., Gallo-Meagher, M., and Thompson, W.F. (1992). Light regulatory sequences are located within the 5′ portion of the *Fed-1* message sequence. EMBO J. **11**, 2311-2317.

Fang, R.-X., Nagy, F., Sivasubramaniam, S., and Chua, N.-H. (1989). Multiple *cis* regulatory elements for maximal expression of the cauliflower mosaic virus 35S promoter in transgenic plants. Plant Cell **1**, 141-150.

Fluhr, R., Moses, P., Morelli, G., Coruzzi, G., and Chua, N.-H. (1986). Expression dynamics of the pea rbcS multigene family and organ distribution of the transcripts. EMBO J. **5**, 2065-2071.

Fraley, R.T., Rogers, S.G., Horsch, R.B., Eichholtz, D.A., Flick, J.S., Fink, C.L., Hoffmann, N.L., and Sanders, P.R. (1985) The SEV system: A new disarmed Ti plasmid vector system for plant transformation.

Biotechnology **3**, 629-635.

Franco, A.R., Gee, M.A., and Guilfoyle, T.J. (1990). Induction and superinduction of auxin-responsive mRNAs with auxin and protein synthesis inhibitors. J. Biol. Chem. **265**, 15845-15849.

Fritz, C.C., Herget, T., Wolter, F.P., Schell, J., and Schreier, P.H. (1991). Reduced steady-state levels of rbcS mRNA in plants kept in the dark are due to differential degradation. Proc. Natl. Acad. Sci. USA **88**, 4458-4462.

Gallie, D.R., Feder, J.N., Schimke, R.T., and Walbot, V. (1991). Post-transcriptional regulation in higher eukaryotes: the role of the reporter gene in controlling expression. Mol. Gen. Genet. **228**, 258-264.

Gee, M.A., Hagen, G., and Guilfoyle, T.J. (1991). Tissue-specific and organ-specific expression of soybean auxin-responsive transcripts GH3 and SAURs. Plant Cell **3**, 419-430.

Giuliano, G., Hoffman, N., Ko, K., Scolnik, P., and Cashmore, A. (1988). A light-entrained circadian clock controls transcription of several plant genes. EMBO J. **7**, 3635-3642.

Green, P.J., Yong, M.-H., Cuozzo, M., Kano-Murakami, Y., Silverstein, P., and Chua, N.-H. (1988) Binding site requirements for pea nuclear protein factor GT-1 correlate with sequences required for light-dependent transcriptional activation of the *rbcs-3A* gene. EMBO J. **7**, 4035-4044.

Hadfield, C., Cashmore, A.M., and Meacock, P.A. (1986) Sequence and expression characteristics of a shuttle chloramphenicol-resistance marker for *Saccharomyces cerevisiae* and *Eschericia coli*. Gene **45**, 149-158.

Herr, W., and Clarke, J. (1986). The SV40 enhancer is composed of multiple functional elements that can compensate for one another. Cell. **45**, 461-470.

Herrick, D., Parker, R., and Jacobson, A. (1990) Identification and comparison of stable and unstable mRNAs in *Saccharomyces cerevisiae*. Mol. Cell. Biol. **10**, 2269-2284.

Hunt, T. (1988). Controlling mRNA lifespan. Nature **38**, 567-568.

Ingelbrecht, I.L.W., Herman, L.M.F., Dekeyser, R.A., Van Montagu, M.C., and

Depicker, A.G. (1989) Different 3′ end regions strongly influence the level of gene expression in plant cells. Plant Cell **1**, 671-680.

Jefferson, R.A., Burgess, S.M., and Hirsh, D. (1986) *β*-glucuronidase from *Eschericia coli* as a gene-fusion marker. Proc. Natl. Acad. Sci. USA **83**, 8447-845.

Kuhlemeier, C., Fluhr, R., Green, P.J., and Chua-N.-H. (1987). Sequences in the pea *rbcS-3A* gene have homology to constitutive mammalian enhancers but function as negative regulatory elements. Genes and Devel. **1**, 247-255.

Lam, E., and Chua, N.-H. (1990). GT-1 binding site confers light responsive expression in transgenic tobacco. Science **248**, 471-474.

Lang, K.M. and Spritz, R.A. (1987) In vitro splicing pathways of pre-mRNA containing multiple intervening sequences. Mol. Cell. Biol. **7**, 3428-3437.

Lincoln, J.E., and Fisher, R.L. (1988). Diverse mechanisms for the regulation of ethylene-inducible gene expression. Mol. Gen. Genet. **212**, 71-75.

Logemann, J., Mayer, J.E., Schell, J., and Willmitzer, L. (1988). Differential expression of genes in potato tubers after wounding. Proc. Natl. Acad. Sci. USA **85**, 1136-1140.

Marcotte, W.R. Jr., Russell, S.H., Quatrano, R.S. (1989). Abscisic acid-responsive sequences from the Em gene of wheat. Plant Cell **1**, 969-976.

McClure, B.A., and Guilfoyle, T. (1989). Rapid redistribution of auxin-regulated RNAs during gravitropism. Science **243**, 91-93.

McClure, B.A., Hagen, G., Brown, C.S., Gee, M.A., and Guilfoyle, T.J. (1989). Transcription, organization, and sequence of an auxin-regulated gene cluster in soybean. Plant Cell **1**, 229-239.

Millar, A.J., and Kay, S.A. (1991). Circadian control of *cab* gene transcription and mRNA accumulation in *Arabidopsis*. Plant Cell **3**, 541-550.

Mogen, B.D., MacDonald, M.H., Graybosch, R., and Hunt, A.G. (1990). Upstream sequences other than AAUAAA are required for efficient messenger RNA 3′-end formation in plants. Plant Cell **2**, 1261-1272.

Mogen, B.D., MacDonald, M.H., Leggewie, G., and Hunt, A.G. (1992). Several distinct types of sequence elements are required for efficient mRNA 3' end formation in a pea rbcs gene. Mol. Cell. Biol. **12**, 5406-5414.

Nagy, F., Boutry, M., Hsu, M.-Y.,Wong, M., and Chua, N.-H. (1987) The 5' proximal region of the wheat *Cab-1* gene contains a 268-bp enhancer-like sequence for phytochrome response. EMBO J. **6**, 2537-2542.

Nagy, F., Kay, S.A., and Chua, N.-H. (1988a). A circadian clock regulates transcription of the wheat *Cab-1* gene. Genes and Devel. **2**, 376-382.

Nagy, F., Kay, S.A., and Chua, N.-H. (1988b) Analysis of gene expression in transgenic plants. In Plant Molecular Biology Manual. S.B. Gelvin, R.A. Schilperoort, eds. (Dordrecht, The Netherlands: Kluwer Academic Publishers), pp. 1-29.

Peltz, S.W., Brewer, G., Bernstein, P., Hart, P.A., and Ross, J. (1991). Regulation of mRNA turnover in eukaryotic cells. In Critical Reviews in Eukaryotic Gene Expression, G.S. Stein, J.L. Stein, and J.B. Lian, eds (Boca Raton: CRC Press), pp. 99-126.

Peltz, S.W., and Jacobson, A. (1993). mRNA turnover in *Saccharomyces cerevisiae*. In Control of mRNA Stability, G. Brawerman and J. Belasc o, eds (New York: Academic Press), in press.

Puissant, C. and Houdebine, L.-M. (1990) An improvement of the single-step method of RNA isolation by acid guanidinium thiocyanate-phenol-chloroform extraction. Biotechniques **8**, 148-149.

Rogers, S.G., Klee, H.J., Horsch, R.B., and Fraley, R.T. (1987) Improved vectors for plant transformation: Expression cassette vectors and new selectable markers. Methods Enzymol. **153**, 253-277.

Rowland, L.J., and Strommer, J.N. (1986). Anaerobic treatment of maize roots affects transcription of *Adh1* and transcript stability. Mol. Cell. Biol. **6**, 3368-3372.

Savant-Bhonsale, S. and Cleveland, D.W. (1992). Evidence for instability of mRNAs containing AUUUA motifs mediated through translation-dependent assembly of a >20S degradation complex. Genes and Devel. **6**, 1927-1939.

Seeley, K.A., Byrne, D.H., and Colbert, J.T. (1992). Red light-independent instability of oat phytochrome mRNA in vivo. Plant Cell **4**, 29-38.

Shaw, G., and Kamen, R. (1986). A conserved AU sequence from the 3' untranslated region of GM-CSF mRNA mediates selective mRNA degradation. Cell **46**, 659-667.

Shirley, B.W., and Meagher, R.B. (1990). A potential role for RNA turnover in the light regulation of plant gene expression: ribulose-1, 5-bisphosphate carboxylase small subunit in soybean. Nucleic Acids Res. **18**, 3377-3385.

Shyu, A.-B., Greenberg, M.E., and Belasco, J.G. (1989). The c-*fos* transcript is targeted for rapid decay by two distinct mRNA degradation pathways. Genes and Devel. **3**, 60-72.

Siflow, C.D. and Key, J.L. (1979) Stability of polysome-associated polyadenylated RNA from soybean suspension culture cells. Biochemistry **18**, 1013-1018.

Stern, D.B., Jones, H., and Gruissem, W. (1989). Function of plastid mRNA 3' inverted repeats: RNA stabilization and gene-specific protein binding.

Taylor, C.B. and Green, P.J. (1991). Genes with homology to fungal and S-gene ribonucleases are expressed in *Arabidopsis thaliana*. Plant Physiol. **96**, 980-984

Theologis, A. (1986). Rapid gene regulation by auxin. Annu. Rev. Plant Physiol. **37**, 407-438.

Thompson, W.F. and White, M.J. (1991). Physiological and molecular studies of light-regulated nuclear genes in higher plants. Annu. Rev. Plant Physiol. Plant Mol. Biol. **42**, 423-466.

van der Zaal, E.J., Droog, F.N.J., Boot, C.J.M., Hensgens, L.A.M., Hoge, J.H.C., Schilperoort, R.A., and Libbenga, K.R. (1991). Promoters of auxin-induced genes from tobacco can lead to auxin-inducible and root tip-specific expression. Plant Mol. Biol. **16**, 983-998.

Walling, L., Drews, G.N., and Goldberg, R.B. (1986). Transcriptional and post-transcriptional regulation of soybean seed protein mRNA levels. Proc. Natl. Acad. Sci. USA **83**, 2123-2127.

Wu, C.H., Caspar, T., Browse, J., Lindquist, S., and Somerville, C. (1988). Characterization of an HSP70 cognate gene family in *Arabidopsis*. Plant Physiol. **88**, 731-740.

Yamamoto, K.T., Mori, H., and Imaseki H. (1992). cDNA cloning of indole-3-acetic acid-regulated genes: Aux22 and SAUR from mung bean (*Vigna radiata*) hypocotyl tissue. Plant Cell Physiol. **33**, 93-97.

CHAPTER 5

THE EFFECT OF SEQUENCES WITH HIGH AU CONTENT ON mRNA STABILITY IN TOBACCO

This Chapter is in press in *Proc. Natl. Acad. Sci.*

Reference: Ohme-Takagi, M., Taylor, C.B., Newman, T.C. and Green, P.J. (1993). The effect of sequences with high AU-content on mRNA stability in tobacco. P.N.A.S. (in press).

ABSTRACT

Little is known about the mechanisms that target transcripts for rapid degradation in plants. In mammalian cells, sequences with a high A + U content and multiple AUUUA motifs have been shown to cause mRNA instability when present in the 3′ untranslated regions of several transcripts. This precedent, coupled with the poor accumulation of AU-rich foreign transcripts in plants (e.g., BT-toxin mRNAs), prompted us to test whether AU sequences could destabilize transcripts in tobacco. To address this question, we made a set of constructs containing sequences with high A + U content inserted into the 3′ untranslated regions of reporter genes. The stability of the corresponding transcripts was then assayed in stably transformed cell lines of tobacco. These experiments showed that a 60-base sequence containing 11 copies of the AUUUA motif ("AUUUA repeat") markedly destabilized a β-glucuronidase reporter transcript, compared to a no-insert control or a 60-base spacer sequence (GC-control). Another sequence with an identical A + U content had little effect. The same results were obtained when each sequence was assayed within the 3′ untranslated region of a β-globin reporter transcript. In regenerated transgenic plants, the AUUUA repeat decreased the accumulation of the β-globin transcript by about 14 fold, compared to the GC-control. Taken together, our results indicate that the AUUUA repeat is recognized as an instability determinant in plant cells and that the effect is due to the sequence of the element, not simply to the high A + U content.

INTRODUCTION

An important aspect of the regulation of gene expression in all organisms is the control of cytoplasmic mRNA levels. In mammalian, plant, and yeast cells it is clear that for many transcripts, differences in mRNA stability contribute greatly to the establishment of steady-state mRNA levels and the speed at which those levels are achieved after a change in transcription rate [1-3]. In particular, proteins that are required by the cell only transiently are often encoded by very unstable mRNAs, generally with half-lives of less than 60 min. In mammalian cells, these include mRNAs encoding regulatory proteins such as c-myc, c-fos and granulocyte-monocyte colony stimulating factor (GM-CSF) [1,2]. The instability of these mRNAs allows for their rapid disappearance from the cytoplasm after a decrease in transcription.

Efforts to elucidate mechanisms of mRNA turnover in mammalian cells have led to the identification of sequences that mediate the selective decay of specific transcripts. For some transcripts, such as those encoding histone H4 [4], β-tubulin [5], and transferrin receptor [6], the mechanisms that control mRNA stability are highly specialized and therefore involve unique regulatory sequences. In contrast, there is a common sequence characteristic among many of the lymphokine, cytokine, and proto-oncogene mRNAs that are rapidly degraded. The 3′ untranslated regions (UTRs) of these transcripts generally contain a 30 to 80-base sequence that is rich in A and U residues and includes

copies of the pentamer motif AUUUA [7, 8]. When the AU domains of the GM-CSF [7] or c-fos [9] 3′ UTRs were inserted into the otherwise stable β-globin mRNA, the modified transcripts were rapidly degraded. Substitution mutations in natural and synthetic sequences indicate that AUUUA repeats in some of these transcripts can function as important determinants of mRNA instability in mammalian cells [10, 11]. However, there does not appear to be any clear relationship between the number of AUUUAs and the instability conferred by endogenous AU domains, perhaps due to differences in context [12]. AU domains in the 3′ UTR may also contribute to mRNA instability in yeast, but AUUUA sequences do not appear to be critical determinants [13]. The apparent lack of recognition of AUUUA sequences in a molecular genetic system presently limits the approaches that can be taken to address the mechanisms of AUUUA- mediated mRNA instability.

A number of plant transcripts contain AU motifs that may contribute to post-transcriptional control. For example, it has been suggested that the *par* transcript of tobacco [14] may be relatively short lived because of the presence of three AUUUA sequences in the 3′ UTR, but this possibility has not been tested. Another AU motif, AUUAA, is repeated six times in the UTRs of the soybean AUX22 mRNA [15]. The stability of this transcript has not been directly measured, but its level decreases appreciably when excised tissue is incubated in auxin-free medium for 6 h [16]. Evidence suggesting that AU-rich transcripts may be rapidly degraded in plants has also resulted from

experiments aimed at engineering plants to express bacterial insecticidal protein genes. Several groups have demonstrated that the mRNAs encoded by insecticidal protein (BT-toxin) genes of *Bacillus thuringiensis* fail to accumulate in plants even when strong plant promoters are used to drive gene expression [e.g., 17, 18]. These BT-toxin transcripts are generally AU-rich and contain multiple copies of both AUUUA and AUUAA motifs that may contribute to mRNA instability, but other mechanisms limiting transcript accumulation (e.g. anomalous RNA processing or a block in transcriptional elongation) have not been ruled out. Thus, direct evidence linking sequences having high AU content with mRNA instability in plants is lacking. The only sequences that have been demonstrated to selectively target transcripts for rapid decay in plants are DST sequences [19], elements highly conserved in the 3′ UTRs of plant Small Auxin Up RNA (SAUR) transcripts [20]. These sequences are not AU-rich, nor do they contain conserved AUUUA or AUUAA motifs [19, 20].

In this study, we compare the effect of AU-containing and control sequences on mRNA stability in plants by direct measurement of mRNA decay rates in stably transformed cells of tobacco. Our results demonstrate that a 60 base sequence consisting of 11 overlapping repeats of the AUUUA motif can target transcripts for rapid degradation in tobacco cells and decrease transcript accumulation in transgenic plants. Furthermore, we show that the effect of AU-containing sequences on mRNA stability is sequence-specific and not simply due to high A+U content.

MATERIALS AND METHODS

Gene Construction

Oligonucleotides were synthesized using an Applied Biosystems DNA Synthesizer, subcloned, and sequenced before insertion into the test genes as described in Figure 5-1. GUS test genes were inserted into a pMON505 derivative containing a 35S-CAT-3C reference gene (p847; [19]; see Figure 5-1). Globin test genes were inserted into a pMON505 derivative containing a 35S-GUS-3C reference gene (p851; [19]). All constructs were introduced into *Agrobacterium tumefaciens* GV3111SE, as described previously [19].

Tobacco transformation

An established line of *Nicotiana tabacum* (NT-1) cells [23] was transformed as described by An [23], as modified by Newman et al. [19]. After 3 to 5 weeks, individual stably transformed calli were transferred to fresh medium (NTKC; [19]), and the expression of the test and reference genes was confirmed by Northern blots [19]. For regeneration of transgenic plants, constructs were introduced into *Nicotiana tabacum* cv. Xanthi by leaf disc transformation [19]. Plants were grown and leaves were harvested for RNA isolation as described previously [19].

RNA isolation, measurement of mRNA stability, and S1 mapping

To establish suspension cultures of NT-cells for mRNA half-life measurements, stably transformed calli were transferred to liquid NT-medium and grown as described [19]. Three to four days after sub-culture, Actinomycin D (ActD) was added to a concentration of 100 μg/ml and 10 ml samples were withdrawn every 30 min for 2.5 h. Cells were harvested, frozen in liquid nitrogen, and RNA was isolated and analyzed on Northern blots [19, 24]. The blots were hybridized to random-primed probes corresponding to the coding regions of the appropriate test and reference genes [19]. Signals were quantitated with a Phosphorimager and linear regressions were plotted to calculate the half-lives of the individual mRNAs [19]. S1 nuclease protection analyses were performed as described previously [25], except that the annealing reactions contained 70% formamide and 40 μg of total RNA.

RESULTS

We were interested in determining whether sequences with high A+U content could selectively target transcripts for rapid decay in plants and, if so, whether the sequence of the element was important or simply the high AU-content. Testing the function of AUUUA sequences in plants was of particular interest because these elements can cause mRNA instability in mammalian cells, but

have not been shown to function in other systems. To this end, we prepared the series of constructs illustrated in Figure 5-1. Beginning with a 35S-GUS-E9 test gene (designated "no insert"), three derivatives were made, each containing a different 60-bp fragment inserted into the 3' UTR between the GUS coding region and the E9 3'end. As shown in Figure 5-1B, two of the constructs, the AUUAA repeat and the AUUUA repeat, contained inserts with identically high A+U content, but different primary sequences. The AUUAA repeat was designed to test whether AUUAA sequences, such as those repeated in the UTRs of the soybean *AUX22* gene [15], could destabilize transcripts in plants. To investigate specifically the function of AUUUA sequences in plants, a 60 base insert containing the maximum possible number of AUUUA motifs (11 overlapping pentamers) was made and designated the AUUUA repeat. As a size control, an insert interspersed with Gs and Cs was synthesized and designated the GC-control [7]. Each of the GUS test genes was incorporated into a transformation vector containing an identical 35S-CAT-3C reference gene that served as an internal standard to facilitate comparisons between experiments.

These constructs were introduced into a tobacco cell line in which mRNA decay rates could be directly measured. The NT-1 cell line was chosen for these experiments because it can be grown as a suspension culture in liquid media and can be stably transformed with foreign genes via *A. tumefaciens* [23]. In addition, ActD, a transcription inhibitor commonly used to facilitate the

Legend to Figure 5-1

Constructs used to test the effect of AU-containing sequences.

(A) The vectors used for plant transformation included a test gene to test the insert of interest and a reference gene to serve as an internal standard. In the first series of constructs, the test gene consisted of the β-glucuronidase (GUS) gene under the transcriptional control of the 35S promoter (-941 to +8) from cauliflower mosaic virus (35S) as described previously [19]. Polyadenylation signals were provided by the 3' end of the *RBCS-E9* gene (E9 3') [19]. Inserts were cloned into a unique BamHI site between the GUS coding region and the E9 3' end of the test gene. A 35S-GUS-E9 test gene without an insert was designated the "no insert" control. All vectors containing GUS test genes contained an identical reference gene consisting of a chloramphenicol acetyltransferase (CAT) gene flanked by the 35S promoter and *RBCS-3C* 3' end (3C 3') [19] as shown. The test and reference genes were inserted into the polylinker (between SacI and ClaI), and the HpaI site of the binary vector pMON505 [21], respectively. A second set of constructs (not shown) was made in the same manner but contained a 35S-globin-E9 test gene and a 35S-GUS-3C reference gene. Globin sequences were derived from a human β-globin cDNA clone [22] modified to remove the BamHI site within the coding region [19]. (B) Synthetic sequences inserted into the test genes contained a BglII site and a BamHI site at the 5' and 3' ends, respectively; the GATC sequences within these sites are underlined. The pentamer motifs, AUUAA and AUUUA, found within the AUUAA repeat (rpt) and AUUUA repeat (rpt) are indicated by arrows.

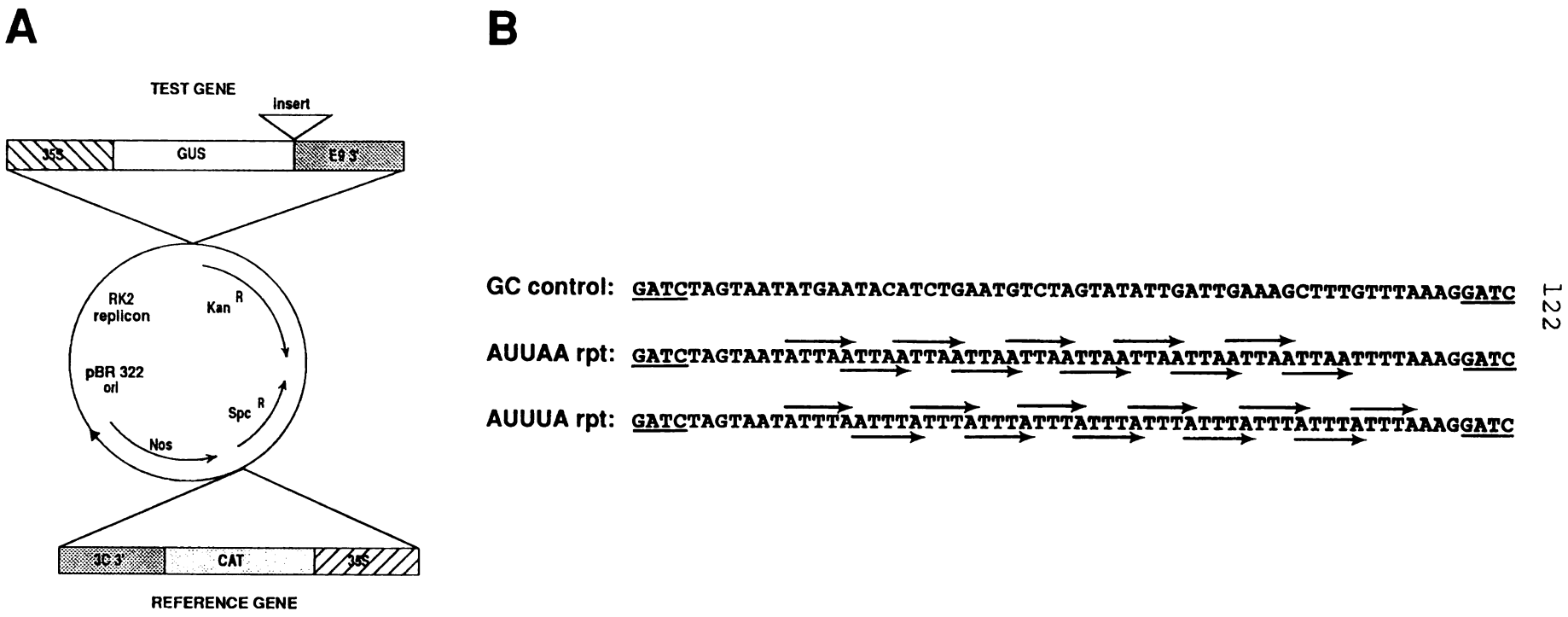

Figure 5-1. Constructs used to test the effect of AU-containing sequences.

measurement of mRNA decay rates in cultured mammalian cells (e.g., [7, 12]), is an effective inhibitor in NT-1 cells [19]. To measure decay rates in NT-1 cells, stably transformed cell lines containing each of the constructs in Figure 5-1 were grown in suspension culture and treated with ActD. RNA was isolated from samples removed from the cultures at various times as described in Materials and Methods. Degradation of the test and reference transcripts was monitored on Northern blots by hybridization to GUS and CAT probes, as described [19]. For each construct, two to four independent stably transformed cell lines were analyzed and four to eight Northern blot experiments were performed.

Among all of the sequences tested, only the AUUUA repeat was found to affect significantly the stability of the GUS test transcript in NT-1 cells. As the Northern blots in Figure 5-2 indicate, the GUS transcript containing the AUUUA repeat (GUS-AUUUA) was degraded considerably faster than the GUS transcript containing the GC-control (GUS-GC). Table 5-1 shows the mRNA half-life measurements for the test (GUS) and reference (CAT) transcripts in cell lines transformed with the GUS-AUUUA and GC-control constructs. We also normalized the half-life of test transcript to that of the reference transcript to calculate the relative half-life in each case. While the decreased stability of the GUS-AUUUA transcript can be observed by comparing either value, we have found previously [19] that the relative half-life calculations most effectively normalize for cell-line to cell-line variations that are periodically observed (e.g.,

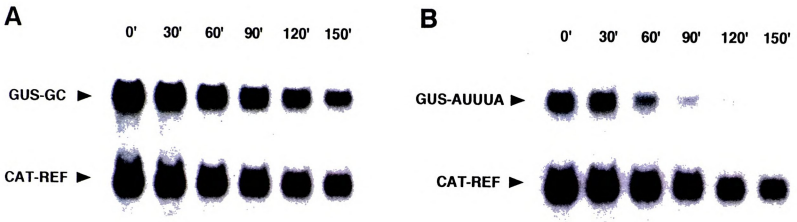

Figure 5-2. Rapid mRNA decay caused by the AUUUA repeat in tobacco cells.

Representative northern blot experiments used to measure the decay of GUS reporter transcripts containing A) the GC-Control (GUS-GC; cell line 3 in Table 5-1) or B) the AUUUA repeat (GUS-AUUUA; cell line 2 of Table 5-1) in stably transformed NT-1 cells. Each lane contained 20 μg RNA isolated from cells harvested at the indicated times after ActD treatment. Blots were hybridized to GUS and CAT coding region probes to detect the test and reference (REF) gene transcripts, respectively.

These experiments were performed by Masaru Ohme-Takagi

Table 1. Effect of the AUUUA repeat on GUS mRNA half-lives in NT-1 cells.

Insert	Cell line	no. of blots	Absolute Half-lives (min) GUS Test	CAT Ref.	Avg. Ratio*	Relative Half-life†
GC	1	1	148	176	0.84	
	2	1	180	215	0.84	0.83±0.02 (0.97±0.02)
	3	3	82±8	102±2	0.80	
AUUUA	1	3	47±3	136±32	0.36	
	2	3	47±16	102±18	0.45	0.35±0.11 (0.41±0.13)
	3	2	42±3	187±12	0.23	

Absolute half-lives of the GUS test and CAT reference (CAT ref.) transcripts measured in GUS-GC (GC) and GUS-AUUUA (AUUUA) transformants are listed. Data derived from multiple Northern blots is expressed ± the SD. *For each Northern blot, the half-life of the test transcript was divided by that of the reference transcript, and these values were averaged to calculate the Average Ratio (Avg. Ratio). †The relative half-life corresponds to the mean of the Average Ratios for each construct ± the SD. The values in parentheses are relative half-lives normalized to that of the no-insert control (as in Figure 3).

These experiments were performed by Masaru Ohme-Takagi

compare data from GUS-GC cell line 3 in Table 5-1 to that of cell lines 1 and 2). The relative half-lives show that the GUS-AUUUA transcript was degraded 2.4 times faster than the GC-control (3 times faster if only absolute half-lives are considered). Figure 5-3A compares our cumulative data for all of the inserts. In contrast to the AUUUA repeat, the other AU-rich sequence, the AUUAA repeat, had virtually no effect on mRNA stability compared to the no-insert and GC-controls.

In order to investigate the effect of the AU-containing sequences in Figure 5-1B within a different context, another set of constructs was made using a second test gene (human β-globin). We chose to engineer globin sequences for this purpose because, similar to GUS probes, globin probes do not cross-hybridize to endogenous plant transcripts. In addition, use of a globin gene would allow us to evaluate the ability of AUUUA sequences to destabilize a shorter transcript, as the globin coding region is only about one-fourth the size of the GUS coding region. Analogous to the GUS constructs, each of the AU-containing sequences and the GC-control was inserted into the 3′ UTR between the coding region and the E9 3′ end. For the mRNA stability studies with the globin test genes, an unmodified 35S-GUS-3C reference gene, co-transformed on the same vector, served as the internal standard. As shown in the histogram in Figure 5-3B, the results with the globin test genes were nearly identical to those obtained with the GUS test genes. The stability of the globin transcript containing no-insert was the same as that containing the GC-control.

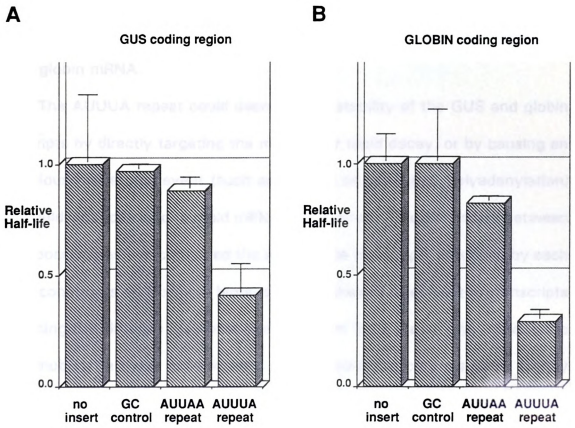

Figure 5-3. Histograms comparing the effect of the AU-containing inserts on GUS and globin mRNA stability.

Each bar represents the cumulative data from multiple mRNA half-life experiments performed on multiple transformants for each construct as described in the text. The half-lives of the test transcripts were normalized to the half-lives of reference transcripts and are expressed relative to the no-insert control (relative half-lives calculated as in Table 5-1). Error bars represent the SD.

These experiments were performed by Masaru Ohme-Takagi and Tom Newman

Insertion of the AUUUA repeat caused the globin transcript to decay 3.4 times faster (4.4 times faster if only absolute half-lives, which averaged 44 ± 4 min and 193 ± 31 min for the AUUUA repeat and the GC control respectively, are considered). The AUUAA repeat sequence had very little effect on the stability of the globin mRNA.

The AUUUA repeat could decrease the stability of the GUS and globin transcripts by directly targeting the mRNAs for rapid decay, or by causing an anomalous processing event (such as splicing or premature polyadenylation) that could indirectly lead to rapid mRNA degradation. To differentiate between these possibilities, we compared the sizes of the transcripts produced by each of the constructs on Northern blots (data not shown). Each set of transcripts containing the 60-base inserts shown in Figure 5-1B was found to have the same mobility in these experiments, and the no-insert controls were slightly smaller, as expected. To compare the 3′ ends of each of the transcripts, S1 nuclease protection experiments were performed. As shown in Figure 5-4, all of the transcripts had the same 3′ ends that corresponded to the known polyadenylation sites from the *RBCS-E9* gene [25] that were present in each construct. These data indicate that the observed effects of the AUUUA repeat were direct effects of the insertion on mRNA stability and were not caused indirectly by anomalous processing events.

The increased degradation rates of transcripts containing the AUUUA repeat observed in NT-1 cells suggested that the AUUUA repeat may also have

Legend to Figure 5-4: S1 nuclease protection analysis of the 3′ ends of the test transcripts.

A 690- base probe from the 3′ end of the *RBCS-E9* gene [25] was annealed to RNA samples from cell lines expressing each construct and digested with S1 nuclease as described in Materials and Methods. The expected protection products [19, 25] corresponding to the test (A; 230 to 249 bases) and reference (B; 81 bases) transcripts are indicated. The reference gene protection product is detected due to partial homology between the *RBCS-E9* and *RBCS-3C* 3′ ends [25].

These experiments were performed by Masaru Ohme-Takagi

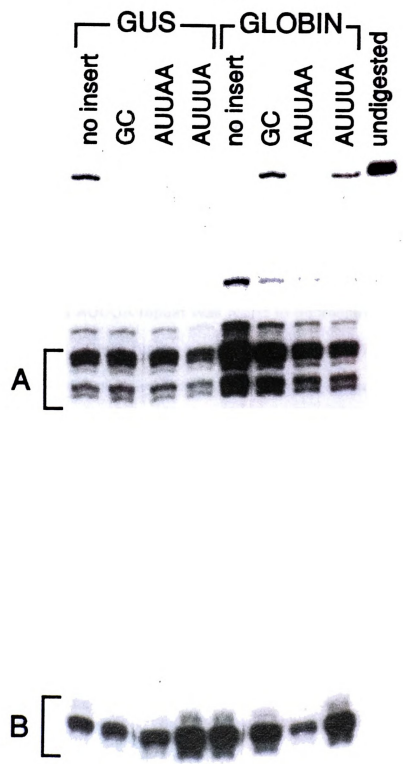

Figure 5-4. S1 nuclease protection analysis of the 3′ ends of the test transcripts.

an effect in transgenic plants. To test this hypothesis, plants containing either the globin-GC construct or the globin-AUUUA construct were regenerated. Figure 5-5 shows a Northern blot comparing the steady-state transcript levels in 12 to 13 plants per construct. All of the transgenic plants contained an identical GUS reference gene and both sets of plants produced similar levels of the GUS transcript. In contrast, the globin-AUUUA transcript was present at clearly diminished levels relative to the globin-GC control. When globin transcript levels were quantitated and normalized to the level of GUS transcript in each plant, the AUUUA repeat was found to decrease mRNA accumulation by about 14 fold. These results suggest that the AUUUA repeat also causes mRNA instability in transgenic plants, although the magnitude of the effect is somewhat greater than that observed when mRNA half-lives were measured in NT-1 cells.

DISCUSSION

In this study we have shown the AUUUA repeat can target both the GUS and globin reporter transcripts for rapid mRNA decay in tobacco. In contrast, the AUUAA repeat, an insert with the same A+U content, had little effect compared to the no insert or GC control. This indicates that the sequence of the AUUUA repeat, rather than its high AU content per se, is responsible for its

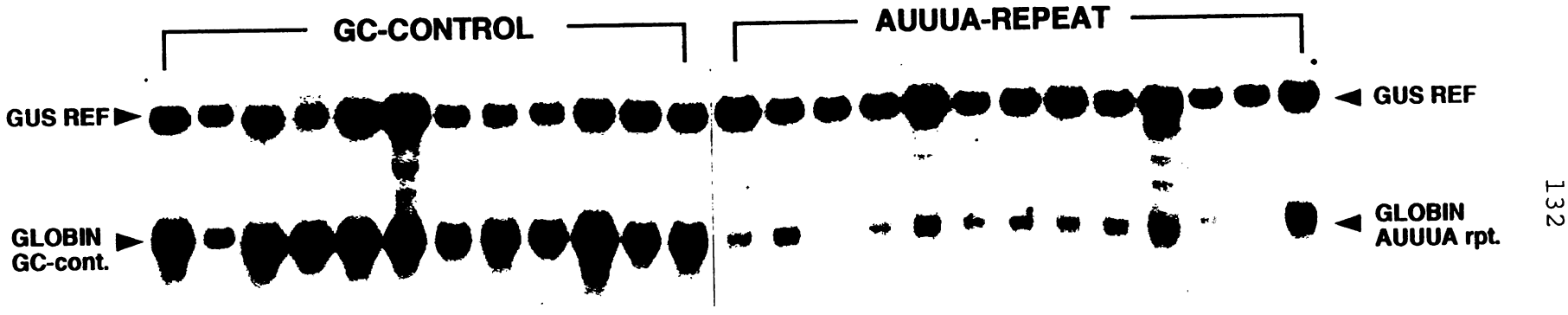

Figure 5-5. Effect of the AUUUA repeat in transgenic plants.

RNA was isolated from leaves of transgenic plants expressing the globin-GC or globin-AUUUA constructs. Each lane contained 10 μg of RNA from an independent transformant. Blots were hybridized to globin and GUS coding region probes to detect the test and reference transcripts, respectively.

recognition as an mRNA instability determinant in tobacco. Although our study is the first demonstration that AUUUA sequences can destabilize transcripts in nonmammalian systems, the magnitude of the effect that we observed was comparable to that observed for other AU-rich instability determinants examined in mammalian cells. In our experiments, the half-lives of the globin-AUUUA and globin-GC control averaged 44 min and 193 min, respectively. Similarly, when a synthetic sequence consisting of alternating AUUUA motifs and U_6 stretches was inserted into a globin transcript, the half-life decreased to about 66 min compared to the >240 min half-life exhibited by a GC control in mammalian cells [11]. In another example, insertion of sequences from the GM-CSF 3'UTR into a globin reporter transcript decreased its half-life to <30 min [7]. The GC control in this case had a half life of >120 min. Therefore it appears that the degradation kinetics of globin transcripts containing the AUUUA repeat in plants are similar to those of transcripts containing AU-rich instability determinants in mammalian cells.

Our results demonstrated that the instability caused by the AUUUA repeat in tobacco is not the result of altered polyadenylation because all the transcripts have the expected size and 3' ends. Therefore, the increase in mRNA degradation observed in stably transformed tobacco cell lines is due to the direct effect of the element on GUS and globin mRNA decay. Because the AUUUA repeat also decreases mRNA levels in leaves of transgenic plants, it is highly likely that the element acts as an mRNA instability determinant in intact

plants as well. At present it is unknown why the magnitude of the effect was greater when mRNA accumulation was measured in transgenic plants (14 fold decrease) than when mRNA half-lives were measured in cultured cells (3.4 fold decrease). This difference is not due to an inherent difference between NT-1 cells and plants because in NT-1 cells the accumulation of the globin-AUUUA repeat mRNA is also decreased more than 10 fold compared to that of the globin-GC control transcript (M.L. Sullivan and P.J.G., unpublished). A similar discrepancy (of up to 10 fold) between mRNA accumulation and mRNA half-lives was observed recently when an AU instability determinant was investigated in stably transformed mammalian cells [26]. In the latter study, it was suggested that mRNA half-lives measured in the presence of ActD may have been over estimated if recognition of the AU sequence was dependent on ongoing transcription. A similar hypothesis may also explain the discrepancy between mRNA levels and half-life measurements in tobacco. Alternatively, the difference could indicate that the AUUUA repeat limits transcript accumulation by more than one mechanism. In order to differentiate between these possibilities, we are presently developing methods for direct measurement of mRNA decay rates in intact transgenic plants and cultured cells using regulated promoters, rather than transcriptional inhibitors.

Although it is significant that both plants and mammalian cells have mechanisms to rapidly degrade transcripts containing AUUUA repeats, it should also be noted that all AUUUA sequences do not cause mRNA instability.

Several years ago Shaw and Kamen [7] and Caput et al. [8] pointed out that AUUUA sequences were common to the AU-rich 3′ UTRs of many unstable lymphokine and protooncogene transcripts, but no correlation has been established between the degree of mRNA instability and the number of AUUUA sequences that are present in these transcripts. Moreover, AUUUA sequences are not excluded from stable mammalian mRNAs [12]. Another consideration is that even when an AUUUA containing sequence is found to cause mRNA instability, it may not do so constitutively. For example, an AU-rich sequence in the 3′UTR of the β-interferon transcript mediates hormone-induced mRNA destabilization [27], and the GM-CSF 3′UTR does not cause mRNA instability in all mammalian cell lines [28]. It has been proposed that the arrangement or context of AUUUA sequences may control differential recognition of mammalian AUUUA sequences [12].

Nonfunctional AUUUA repeats are also likely to exist in plants because we have observed that some sequences that contain AUUUA motifs, albeit fewer and in a different context than the AUUUA repeat, do not cause mRNA instability in tobacco (M.O.-T., C.B.T., G. Coruzzi, and P.J.G., unpublished results). Our strategy in constructing the AUUUA repeat as a potential instability determinant was to maximize the number of elements within the insert (because multiple elements work better for some transcription factors [29, 30] and the DST instability determinant in plants [19]), and to place them in an overlapping arrangement as found in some natural mammalian elements

[7]. Although this strategy was effective, a systematic study will be necessary in both mammalian and plant cells to identify the critical parameters for AUUUA function in the two systems.

Another aspect that should be considered is whether secondary or higher order structure of AUUUA repeats can influence their function. Higher order structure is difficult to determine experimentally and that of the transcripts investigated in this study is unknown, but the secondary structure potential in each of the 60 base inserts has been analyzed (unpublished results). Only the AUUAA repeat insert has sufficient internal complementarity to form a stable stem-loop structure, and it is possible that this structure contributes to the lack of function of this insert.

Recognition of the AUUUA repeat as an instability determinant in plant cells most likely involves one or more sequence-specific RNA-binding factors or ribonucleases. In mammalian systems, several RNA-binding factors have been identified, via UV crosslinking or gel retardation assays, that interact with AUUUA multimers or U-rich sequences. These include AUBF from Jurkat cells [31, 32]; the AU-A, AU-B, and AU-C proteins from human T cells [33, 34]; a 32-kDa AU-binding protein from HeLa cells [11, 35] that may be identical to AU-A [34]; and a set of U-rich sequence-binding proteins (URBPs) from mouse 3T3 cells [36]. A cytosolic factor from human erythroleukemia K562 cells has also been identified that binds to U-rich sequences such as the 3′ end of c-myc mRNA [37]. The latter activity has been reported to destabilize c-myc mRNA

in vitro [37]. While the *in vivo* role of all of these RNA-binding activities remains unknown, several are inducible and therefore may have a regulatory function [33, 37]. AUUUA motifs have been reported to destabilize spliced and unspliced transcripts [11], and some of the RNA-binding proteins that interact with the element are nuclear [11, 33, 35]. This indicates that AUUUA elements may be recognized in both the nucleus and the cytoplasm. The identification of the AUUUA repeat as an instability determinant in tobacco should facilitate the biochemical investigation of cellular factors that interact with the element in plants. The properties of the plant factors can then be compared with AUUUA-binding proteins identified in mammalian cells.

The most significant result of this study is the demonstration that mechanisms that recognize AUUUA-containing instability determinants extend beyond mammalian cell systems, and more importantly, that such mechanisms are present in organisms that are more amenable to molecular genetics. Genetic approaches have led to major breakthroughs in our understanding of other pathways in RNA metabolism such as splicing [38, 39], but have not been applied to the study of AUUUA recognition because of the technical limitations posed by mammalian systems and the apparent lack of function of the element in yeast. However, in model plants such as *Arabidopsis thaliana* [40], it should be feasible to search for genetic determinants that mediate decay of transcripts containing AUUUA repeats. For example, it may be possible to engineer selectable marker genes with AUUUA repeats and then

select for mutants that lose the ability to degrade rapidly the corresponding transcripts. If the mechanisms for degradation of AUUUA-containing transcripts are similar in plants and mammalian cells, the characterization of such mutants should have broad significance. In addition, the ease of generating transgenic plants should allow the metabolism of AUUUA containing transcripts to be studied in fully intact organisms throughout development.

ACKNOWLEDGEMENTS

We are grateful to Dr. Nam-Hai Chua for synthetic oligonucleotides, and Drs. Michael Sullivan and Natasha Raikhel for comments on the manuscript. This work was supported by grants from the United States Department of Agriculture, the Department of Energy, and the Midwest Plant Biotechnology Consortium and matching funds to P.J.G. M.O.-T. was supported in part by a fellowship from the Human Frontier Science Program.

LITERATURE CITED

1. Atwater, J.A., Wisdom, R. & Verma, I.M. (1990) *Ann. Rev. Genet.* **24,** 519-541.

2. Peltz, S.W., Brewer, G., Bernstein, P., Hart, P.A. & Ross, J. (1991) *Crit. Rev. Eukaryotic Gene Expression* **1,** 99-126.

3. Green, P.J. (1993) *Plant Physiol.* **102,** 1065-1070.

4. Pandey, N.B. & Marzluff, W.F. (1987) *Mol. Cell. Biol.* **7,** 4557-4559.

5. Yen, T.J., Machlin, P.S. & Cleveland, D.W. (1988) *Nature (London)* **334,** 580-585.

6. Müller, E.W. & Kühn, L.C. (1988) *Cell* **53,** 815-825.

7. Shaw, G. & Kamen, R. (1986) *Cell* **46,** 659-667.

8. Caput, D., Beutler, B., Hartog, K., Thayer, R., Brown-Shiner, S. & Cerami, A. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 1670-1674.

9. Shyu, A.-B., Greenberg, M.E. & Belasco, J.G. (1989) *Genes Dev.* **3,** 60-72.

10. Shyu, A.-B., Belasco, J.G. & Greenberg, M.E. (1991) *Genes Dev.* **5,** 221-231.

11. Vakalopoulou, E., Schaack, J. & Shenk, T. (1991) *Mol. Cell. Biol.* **11,** 3355-3364.

12. Greenberg, M.E. & Belasco, J.G. (1993) in *Control of Messenger RNA Stability*, eds. Belasco, J.G. & Brawerman, G. (Academic, New York), pp. 199-218.

13. Muhlrad, D. & Parker, R. (1992) *Genes Dev.* **6,** 2100-2111.

14. Takahashi, Y., Kuroda, H., Tanaka, T., Machida, Y., Takebe, I. & Nagata, T. (1989) *Proc. Natl. Acad. Sci. USA* **86,** 9279-9283.

15. Ainley, W.M., Walker, J.C., Nagao, R.T. & Key, J.L. (1988) *J. Biol. Chem.* **263,** 1058-1066.

16. Walker, J. & Key, J. (1992) *Proc. Natl. Acad. Sci. USA* **79,** 7185-7189.

17. Vaeck, M., Reynaerts, A., Höfte, Jansens, S., De Beuckeleer, M., Dean, C., Zabeau, M., Van Montagu, M. & Leemans, J. (1987) *Nature (London)* **328**, 33-37.

18. Fischoff, D.A., Bowdish, K.S., Perlak, F.J., Marrone, P.G., McCormick, S.M., Niedermeyer, J.G., Dean, D.A., Kusano-Kretzmer, K., Mayer, E.J., Rochester, D.E., Rogers, S.G. & Fraley, R.T. (1987) *Biotechnology*. **5**, 807-813.

19. Newman, T.C., Ohme-Takagi, M., Taylor, C.B. & Green, P.J. (1993) *Plant Cell* **5**, 701-714.

20. McClure, B.A., Hagen, G., Brown, C.S., Gee, M.A. & Guilfoyle, T.J. (1989) *Plant Cell* **1**, 229-239.

21. Rogers, S.G., Klee, H.J., Horsch, R.B. & Fraley, R.T. (1987) *Methods Enzymol.* **153**, 253-277.

22. Lang, K.M. & Spritz, R.A. (1987) *Mol. Cell. Biol.* **7**, 3428-3437.

23. An, G. (1985) *Plant Physiol.* **79**, 568-570.

24. Taylor, C.B. & Green, P.J. (1991) *Plant Physiol.* **96**, 980-984.

25. Fang, R.-X., Nagy, F., Sivasubramaniam, S. & Chua, N.-H. (1989) *Plant Cell* **1**, 141-150.

26. Savant-Bhonsale, S. & Cleveland, D.W. (1992) *Genes Dev.* **6**, 1927-1939.

27. Peppel, K., Vinci, J. & Baglioni, C. (1991) *J. Exp. Med.* **173**, 349-355.

28. Schuler, G.D. & Cole, M.D. (1988) *Cell* **55**, 1115-1122.

29. Kuhlemeier, C., Fluhr, R., Green, P.J. & Chua, N.-H. (1987) *Genes Dev.* **1**, 247-255.

30. Green, P.J., Yong, M.-H., Cuozzo, M., Kano-Murakami, Y., Silverstein, P. & Chua, N.-H. (1988) *EMBO J.* **7**, 3635-3642.

31. Malter, J.S. (1989) *Science* **246**, 664-666.

32. Gillis, P. & Malter, J.S. (1991) *J. Biol. Chem.* **266**, 3172-3177.

33. Bohjanen, P.R., Petryniak, B., June, C.H., Thompson, C.B. & Lindsten, T. (1991) *Mol. Cell. Biol.* **11,** 3288-3295.

34. Bohjanen, P.R., Petryniak, B., June, C.H., Thompson, C.B. & Lindsten, T. (1992) *J. Biol. Chem.* **267,** 6302-6309.

35. Meyer, V.E., Lee, S.I. & Steitz, J.A. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 1296-1300.

36. You, Y., Chen, C.-Y.A. & Shyu, A.-B. (1992) *Mol. Cell. Biol.* **12,** 2931-2940.

37. Brewer, G. (1991) *Mol. Cell. Biol.* **11,** 2460-2466.

38. Ruby, S.W. & Abelson, J. (1991) *Trends in Genet.* **7,** 79-85.

39. Guthrie, G. (1991) *Science* **253,** 157-163.

40. Somerville, C. (1989) *Plant Cell* **1,** 1131-1135.

CHAPTER 6

IDENTIFICATION AND CHARACTERIZATION OF GENES WITH UNSTABLE TRANSCRIPTS (GUTs) IN TOBACCO

ABSTRACT

Plants and other higher eukaryotes have the ability to recognize and target specific transcripts for rapid decay from among the majority of relatively stable mRNAs present within cells. However, little is known about the nature of unstable transcripts in plants, nor the mechanisms that facilitate their rapid degradation. As a first step towards understanding how plants distinguish between unstable and stable transcripts, a novel differential screen was used to identify cDNAs for Genes with Unstable Transcripts (*GUT*s), solely on the basis of the instability of their mRNAs. cDNA probes were prepared from tobacco cells that had been depleted of highly unstable mRNAs by treatment for 90 mins with a transcriptional inhibitor, and from untreated control, untreated cells. *GUT* clones were selected on the basis of weak hybridization to the former probe relative to the latter probe. Half-life measurements performed on the mRNAs hybridizing to eight *GUT* clones indicated that each was unstable, with a half-life on the order of an hour or less. All eight of the cDNAs corresponded to new tobacco genes, and four showed sequence similarity with genes from other species, including the eukaryotic family of DNAJ homologs, a tomato wound-inducible protein, and histone H3. In addition to providing information about the types of transcripts that are inherently unstable in plants, the GUT clones should provide excellent tools for the identification of and cis- and trans-acting determinants of mRNA instability.

INTRODUCTION

Control at the level of mRNA stability is an important component in the regulation of many eukaryotic genes [1,2,3]. This is because mRNA decay rates define the rapidity with which new steady-state transcript levels can be achieved following changes in the transcription rate of a gene. Most mRNAs in higher eukaryotes appear to be relatively stable, with half-lives on the order of hours [4]. Unstable transcripts, with half-lives on the order of minutes, tend to encode proteins that are required only transiently in cells, such as those with critical roles in the regulation of cellular growth and differentiation. Rapid alterations in the synthesis of these proteins is thought to be facilitated by the instability of the corresponding mRNAs [3]. A number of sequence elements have been identified within unstable transcripts that mediate the targeting of those mRNAs for rapid degradation, presumably via their interactions with specific binding factors [1, 2]. One well characterized example is a sequence element termed the AU-rich element (ARE), which consists of a stretch of 40 - 60 A and U bases, including one or more copies of an AUUUA pentamer motif. AREs have been shown to play an important role in the rapid degradation of certain lymphokine, cytokine and proto-oncogene transcripts (reviewed in [2]). Recently, a number of proteins have been identified in mammalian cells that specifically bind to AREs (reviewed in [2]). Some of these proteins may be involved in the recognition of ARE-containing unstable transcripts and their

targeting for rapid degradation.

Control of gene expression at the level of mRNA stability is likely to be particularly important in plants [1]. As sessile organisms plants are unable to move away from adverse stimuli, and are obliged to respond by altering endogenous gene expression, often very quickly. In addition to genes with a role in growth and development, genes involved in rapid plant responses are therefore apt to encode unstable transcripts. The functions of the few plant proteins that are known to be encoded by unstable transcripts generally support this contention. These include PvPRP1, which participates in the response of bean cells to pathogen attack, (the half-life of its transcript appears to decrease markedly following the treatment of bean cells with an elicitor [5]), phytochrome, which mediates many light regulated responses [6], and the Small Auxin Up RNAs (SAURs) of soybean, which encode proteins of unknown function, but whose expression properties are consistent with a role in auxin-induced cell elongation [7, 8]. However, the mechanisms that target these and other unstable transcripts for rapid degradation in plants remain to be elucidated.

Initial insights have been provided by recent studies examining the effect of specific sequence elements on reporter transcript stability in transformed tobacco cell lines [9]. In these experiments, transcription was halted by the addition of Actinomycin D (ActD), and the subsequent decay of the reporter transcripts was monitored by Northern blotting [9]. Using this approach two

very different sequence determinants have been identified that can confer instability on mRNAs in plants. One, termed DST, is highly conserved within the 3′ UTRs of all the SAURs reported to date, it's position and sequence conservation suggesting that it may mediate post-transcriptional control of SAUR expression [10]. When present in two copies in the 3'UTRs of reporter genes, the DST element leads to a pronounced reduction in mRNA half-life [9]. The other sequence, consisting of 11 copies of the AUUUA motif, was also found to markedly decrease reporter transcript half-lives [11]. These observations demonstrate that sequence-specific recognition of unstable transcripts occurs in plants, but it is unlikely that the DST and AUUUA elements are representative of the repertoire of instability determinants present in plant transcripts. To obtain direct information about the complexity of different mechanisms by which specific endogenous transcripts are selected for rapid degradation, the identification and study of a variety of unstable plant transcripts will be necessary.

As a first step towards this goal, we report here the application of a novel differential screening approach to identify Genes with Unstable Transcripts (*GUT*s) in tobacco. The *GUT*s, which were identified solely on the basis of the instability of their transcripts, provide tools for the further characterization of mRNA decay pathways, and have yielded additional information concerning the kinds of proteins that are encoded by unstable transcripts in plants.

MATERIALS AND METHODS

Preparation of RNA and cDNA Probes

Culture conditions for the *Nicotiana tabacum* cv bright yellow 2 (NT-1) cells [12] used in this work have been described elsewhere [9]. To generate cDNA probes for the differential screen, duplicate NT-1 cultures at 3 days post sub-culture were left untreated (control cells), or treated for 90 minutes with the transcriptional inhibitor ActD at 100μg/mL (this concentration of ActD was previously shown to inhibit incorporation of[^{3}H]-uridine into poly (A)$^+$ RNA by 94% [9]). Cells were harvested by centrifugation at 700 x g and frozen in liquid nitrogen. Total RNA was isolated essentially as described [13], except that the extraction solution was buffered with 80mM Tris at pH7.5 (in the place of 25mM sodium citrate, pH 7.0) and a second CsCl cushion, at 2.8M, was included in the gradient. Poly(A)$^+$ RNA was prepared from these samples using oligo dT columns (5 Prime - 3 Prime, Inc., Boulder, Co.), and 1μg of each Poly(A)$^+$ RNA was used as template for the preparation of random-primed first-strand cDNA probes, as described [14].

Isolation of *GUT* clones

The NT-1 cell cDNA library used for screening for *GUT*s (the generous gift of Dr. G. An) was prepared from RNA isolated from NT-1 cells 3 days after subculture, and cloned into λZAP II (Stratagene) with EcoRI-NotI adaptors. The

library was plated at low density (5,000 plaques per 150 mm plate). Duplicate lifts were prepared, using standard protocols [14], and were prehybridized and hybridized as described [15]. Promising clones were plaque purified and zapped into plasmid form [16]. For Northern and Southern blot analyses, the *GUT* inserts were labelled with [^{32}P] by random priming [17].

Half-Life Determinations

Determinations of *GUT* mRNA half-lives were performed as described previously [9]. Briefly, NT-1 cell lines expressing a 35S-GUS-E9 reference gene [9] were treated with ActD at 100μg/mL, and RNA was isolated from aliquots of NT-1 cells harvested from these cultures every 30 mins thereafter. Total RNA was subjected to Northern blotting and hybridized to *GUT* and GUS probes, as described [9]. Quantitation of *GUT* and GUS mRNA half-lives were performed using a phosphorimager as described previously [9]. At least 2 Northerns were performed on at least two independent cell lines with each *GUT* probe. In some cases, absolute half-life determinations were performed using untransformed cells.

Southern Analysis

DNA was isolated from *Nicotiana tabacum*, strain SR1 using an SDS/ phenol extraction procedure [18]. DNA was dissolved in ddH_2O, and aliquots were digested with the restriction enzymes EcoRI, EcoRV or HindIII. 10μg of

digested DNA was separated on 1% agarose gels and transferred to Biotrace HP membrane, as recommended by the manufacturer (Gelman). Prehybridization, hybridization and washing of the Southern blots was performed using the same conditions as for the Northern blots.

Sequence Analysis

Sequencing was performed at the Plant Biochemistry Facility at the Plant Research Laboratory. Plasmid DNA from each *GUT* cDNA clone (or subclone) was prepared using Magic miniprep columns (Promega), and sequenced using Taq cycle sequencing and the Applied Biosystems Inc. 373A automatic sequencer, as recommended by the manufacturer. Nucleotide sequences of 250-300 bp from each end of the cloned *GUT* fragments were compared to sequences in the protein databases using the BLASTX algorithm under default parameter settings [19]. The nucleotide sequences of any *GUT* clones for which no database matches were found at the amino acid level, were compared to the nucleotide databases using the BLASTN algorithm [19]. To search for any conserved motifs among the *GUT*s, the *GUT* cDNA sequences were compared to one another using DNAsis (Hitachi).

RESULTS AND DISCUSSION

Identification of *GUT* cDNA Clones

We have found previously that treatment of NT-1 cells with actinomycin D (ActD) effectively inhibits RNA polymerase II transcription [9], and have used this inhibitor as a tool to measure the half-lives of reporter transcripts in the absence of mRNA synthesis [9, 11]. In an analogous fashion, the levels of endogenous unstable transcripts should also decrease rapidly over time following the treatment of NT-1 cells with ActD. For example, a transcript with a half-life of 20 mins would be present at only 6% of it's initial level after 90 mins of ActD treatment. In contrast, the levels of an average transcript, with a half-life of several hours [3, 4] would remain essentially unchanged after 90 mins of ActD treatment. This wide discrepancy between levels of unstable transcripts in 90 min ActD-treated and in control cells formed the basis for the differential screen described in this report. cDNA probes prepared using Poly(A)$^+$ RNA from ActD-treated cells will be essentially depleted of sequences corresponding to highly unstable transcripts, whereas control cDNA probes will contain a full complement of NT-1 mRNA sequences. The control probe should therefore hybridize to each cDNA clone in the library, but the 90 min ActD probe should hybridize weakly or not at all to plaques harboring cDNA clones derived from unstable transcripts.

A total of approximately 300,000 plaques of an NT-1 cell cDNA library

were screened with control and 90 min ActD cDNA probes, as described in Materials and Methods. To minimize the selection of false positive clones, the first lift from each library plate was hybridized to the 90 min ActD first-strand cDNA probe, and the second lift to the control probe. 150 plaques with significantly decreased hybridization to the 90 min ActD probe, relative to that of the control probe, were isolated and rescreened. Of these, 7 were designated potential *GUT*s, as the plaques continued to show a diminished hybridization signal with the 90 min ActD probe. Purified plaques of potential *GUT* clones were zapped into plasmid form, and the plasmid DNA was digested with the restriction endonuclease EcoRI. In cases where there was more than one EcoRI fragment in a clone, the fragments were subcloned. This was because, in most cases, Southern analyses indicated that the fragments were derived from a fusion of two or more cDNAs in one λ clone (see below).

Further evaluation of the potential *GUT* clones was performed by examining transcript levels in the control and ActD-treated NT-1 cells. The EcoRI fragments from the potential *GUT* clones were labelled with [^{32}P]-dCTP to similar specific activities, and hybridized to Northern blots of total RNA from control and ActD treated NT-1 cells (the same RNAs used for the isolation of poly(A)$^+$ RNA for the generation of cDNA probes for the library screening). Quantitation of the Northern blots shown in Figure 6-1 demonstrated that from the original 7 λ clones, a total of 8 clones (3-2, 7-2A, 7-2C, 8-1, 8-2A, 8-2B, 8-3, 15) could be designated *GUT*s as they exhibited a $\geq$ 2.5 fold decrease in

Legend to Figure 6-1: Preliminary Characterization of Potential *GUT*s

Northern blots of 10μg total RNA isolated from control, untreated NT-1 cells (0), and from cells treated for 90 minutes with ActD (90') or with ActD and cycloheximide (CX) were hybridized to the indicated *GUT* and *NOTGUT* probes. Transcript sizes were calculated relative to RNA molecular weight markers. *GUT* and *NOTGUT* transcript levels under each experimental condition were quantitated using a phosphorimager (see Table 6-1).

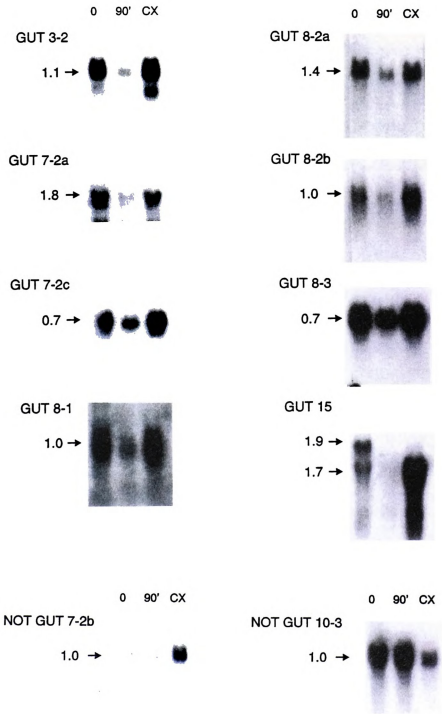

Figure 6-1. Preliminary Characterization of Potential *GUT*s

Legend to Table 6-1

---

[1] The fold increase (↑) or decrease (↓) in *GUT* and *NOTGUT* mRNA in NT-1 cells treated for 90 minutes with Actinomycin D (ActD) or with ActD plus cycloheximide (CHX) was calculated relative to the mRNA levels in control NT-1 cells using the Northern blots in Figure 6-2.

[2]The absolute mRNA half-lives are averages ± standard error of at least three half-life determinations for each *GUT* and *NOTGUT*, with one exception (*GUT* 15 half-life has been determined once). In addition to measurements made in the transformed NT-1 cell lines, these include half-life determinations in untransformed NT-1 cells.

[3]Relative half-lives were calculated by measuring *GUT* or *NOTGUT* and GUS mRNA half-lives, as described in the legend to Figure 6-3, in NT-1 cell lines transformed with a GUS reference construct used previously as an internal standard [9]. Values are average ratios of *GUT* or *NOTGUT* half-life to GUS half-life, ± standard error, calculated from at least three different Northern blots for each *GUT* and *NOTGUT*.

Table 6-1: *GUT* and *NOTGUT* mRNA Levels and Half-Lives

GUT	ActD Effect (Fold)[1]	CHX Effect (Fold)[1]	Absolute Half-Life (*GUT*)[2]	Relative Half-Life (*GUT*/GUS)[3]
3-2	3.0↓	1.5↑	65′ ± 30′	1.26 ± 0.30
7-2a	2.8↓	1.4↑	74 ± 19′	0.92 ± 0.13
7-2c	2.2↓	1.5↑	56′ ± 13′	0.84 ± 0.19
8-1	4.0↓	2.0↑	44′ ± 16′	0.77 ± 0.20
8-2a	2.6↓	1.1↓	49′ ± 20′	0.79 ± 0.10
8-2b	2.8↓	2.0↑	55′ ± 16′	0.86 ± 0.39
8-3	3.0↓	1.0 ↕	52′ ± 06′	0.79 ± 0.16
15 i) ii)	18.0↓ 5.5↓	6.0↓ 6.5↑	43′ 62′	0.36 0.52
NOTGUT				
7-2b	1.1↓	2.3↑	>3hr	>3
10-3	1.1↓	3.1↓	>3hr	>3

transcript abundance during the 90 mins of ActD treatment (see Table 6-1). This corresponds to a theoretical half-life of 72 mins or less. Two clones (7-2B and 10-3) that showed little difference in transcript levels in the cells treated for 90 mins with ActD (Table 6-1) were designated *NOTGUT*s, and used as controls for subsequent experiments.

Cycloheximide Effect

It has been suggested that treatment with the protein synthesis inhibitor cycloheximide (CHX) tends to "stabilize" unstable transcripts in yeast [20]. These data have been interpreted as either implying a role for active translation in the degradation of unstable mRNAs, or the participation of an labile factor in the degradation of these transcripts [20; reviewed in 2]. We were therefore interested in examining the effect of CHX on the accumulation of *GUT* and *NOTGUT* transcripts in NT-1 cells. The blots used for the initial characterization of *GUT* transcript levels in control and 90 min ActD-treated cells also included total RNA isolated from NT-1 cells that had been treated for 90 mins with both ActD and the protein synthesis inhibitor cycloheximide (CHX; Figure 6-1). In contrast to the data from yeast [20], our data suggest that CHX treatment does not affect all unstable transcripts similarly in NT-1 cells. Although most of the *GUT* and *NOTGUT* mRNAs showed a modest increase in mRNA levels in the presence of ActD + CHX, for the larger of the two transcripts identified by the *GUT* 15 cDNA probe, and *NOTGUT* 10-3 there was a clear decrease in mRNA

levels in cells treated with ActD + CHX, relative to the control cells (Figure 6-1; quantitation in Table 6-1). Conversely, the accumulation of the smaller *GUT* 15 transcript in the ActD + CHX treated cells increased markedly, more than that of any other *GUT* (Figure 6-1). It is interesting to speculate on the causes of the different CHX effects on the two *GUT* 15 transcripts. We have preliminary evidence that there are two different classes of *GUT*15 cDNA (CBT and Green, P.J., unpublished) and it is possible that the CHX effect is specific for one of these classes. However, there is some evidence that CHX can stimulate the transcription of some genes [21], and such an effect cannot be ruled out for *GUT* 15 (despite the presence of ActD concentrations known to halt transcription). Nonetheless, whether the CHX effect is mediated post-transcriptionally, or transcriptionally, it will be most interesting to identify any sequence differences between the two classes of *GUT* 15 that may cause this effect.

GUT mRNA Half-Lives

To confirm that the *GUT*s did indeed encode unstable transcripts, *GUT* mRNA decay rates were measured by Northern blotting of RNAs isolated from NT-1 cells at 30 minute intervals after the addition of ActD to the cultures (see Materials and Methods). Representative Northern blots of *GUT* and *NOTGUT* mRNA decay are shown in Figure 6-2. To calculate the half-lives of their transcripts, the signals of each *GUT* and *NOTGUT* probe were quantitated as

Legend to Figure 6-2: Time-Course Experiments to Determine *GUT* transcript Half-Lives

Northern blots of 10μg total RNA, isolated from NT-1 cells at the indicated times (in minutes) after the addition of 100μg/mL ActD to the cultures. Blots were hybridized to the indicated *GUT* and *NOTGUT* probes. Half-life measurements from these and additional experiments were determined following quantitation of the signals using a phosphorimager, and are summarized in Table 6-1.

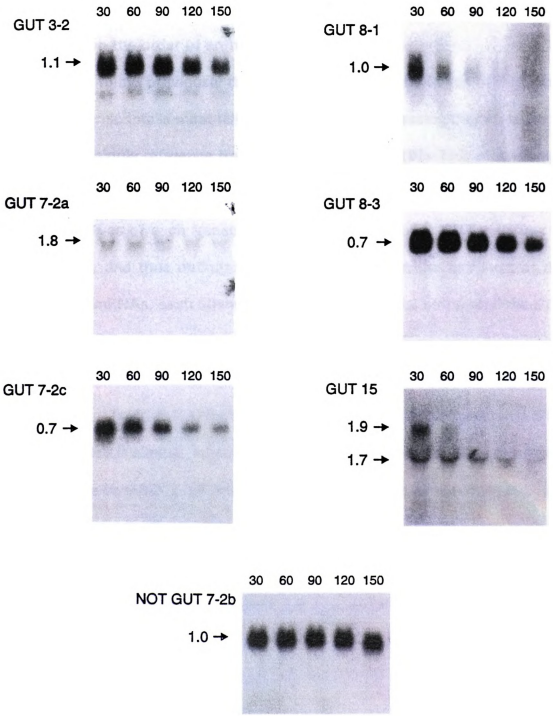

Figure 6-2. Time-Course Experiments to Determine *GUT* transcript Half-Lives

described in Materials and Methods. These experiments show that the *GUT* mRNAs all fall into the category of unstable transcripts [3, 4] as their absolute half-lives are on the order of an hour or less (Table 6-1).

We have found previously that comparison of the half-lives of a number of different transcripts is most reliably achieved by expressing them relative to half-lives of a single reference transcript, such as GUS [9]. This procedure is also an effective way of normalizing for cell line - to - cell line variations in the absolute half-lives of each transcript [9, 11]. To normalize *GUT* half-lives to those of GUS, and thus optimize comparisons between the half-lives of the different *GUT* mRNAs, each time-course Northern blot was stripped of the *GUT* probe and hybridized to a GUS probe. GUS mRNA half-lives were quantitated and ratios of *GUT*/GUS half-lives (relative half-lives) were calculated as described [9]. Table 6-1 shows that, as expected, the rank order of *GUT* mRNA stabilities is similar, whether they are expressed as absolute or relative half-lives. The two *NOTGUT* transcripts did not degrade appreciably over the 150 min timecourse (Figure 6-2 and Table 6-1).

Structural Characteristics of the *GUT*s and *NOTGUT*s

To investigate the number of genes that hybridize to the individual *GUT* probes, Southern analyses were performed as described in Materials and Methods. As summarized in Table 6-2, in most cases the *GUT* and *NOTGUT* cDNA probes hybridized to between 2 and 5 genomic DNA fragments in each restriction

digest (data not shown), demonstrating that they are encoded by single or low copy genes. The most complex pattern was obtained with the *GUT* 8-2B probe, which encodes a tobacco histone H3 cDNA (see below). This probe hybridized to >20 fragments in each digest. The copy numbers for *GUT*s 7-2A, B and C cannot be reliably determined at present from Southern blots, as these Southerns were performed with a probe from the original *GUT* 7-2 λ clone, which includes EcoRI fragments corresponding to each of the three cDNAs (CBT and Green, P.J.), unpublished observations).

In general, the sizes of the *GUT* and *NOTGUT* cDNA clones correlate with the sizes of the corresponding transcripts. With the exception of *GUT* 3-2, all of the *GUT* and *NOTGUT* cDNA probes are smaller than, or the same size as, the major transcripts they detect on Northern blots (see Table 6-2). Nonetheless, in some cases (eg *GUT*s 3-2, 7-2c and 15), it is clear that an individual *GUT* cDNA probe hybridizes to more than one transcript in NT-1 cells (see Figures 6-1 and 6-2). The origins of these smaller transcripts are unknown at present. One intriguing possibility is that some may represent *GUT* mRNA degradation products. However, it is also possible that they are full-length mRNAs derived from genes closely related to the *GUT*s. The presence of two different cDNAs in a single *GUT* probe cannot be ruled out, especially for *GUT* 3-2, where the cloned insert is about three times larger than the major transcript it detects on Northern blots (Table 6-2). The Northern blot experiments in Figures 6-1 and 6-2 also serve to demonstrate that the basal

Table 6-2: Structural Characteristics of *GUT*s and *NOTGUT*s

GUT	Transcript Size (Kb)	cDNA Size (Kb)	Estimated Gene Number[1]	Sequence Similarity[2]
3-2	1.1	2.8	2-5	none
7-2a	1.8	0.75	<10	DNA J
7-2c	0.7	0.35	<10	W.I.P
8-1	1.0	1.0	2-5	none
8-2a	1.4	1.4	1-2	E2
8-2b	1.0	0.6	5-10	Histone H3
8-3	0.7	0.6	2-5	none
15 i) ii)	1.9 1.7	1.7	1-2	none
NOTGUT				
7-2b	1.0	0.4	<10	S5 protein
10-3	1.4	1.1	1-2	Histone H1

[1]Gene numbers were estimated from Southern blots of tobacco genomic DNA.

[2]Summarized from Figure 6-3.

expression levels of the *GUT*s differ markedly from one and other, despite the instability of all their transcripts. For example, *GUT*s 3-2 and 8-3 are expressed at very high levels in control NT-1 cells, but *GUT*s 7-2a and 15 are expressed at levels that are barely detectable on Northern blots (Figures 6-1 and 6-2). Moreover, the sizes of the *GUT* transcripts range from 1.9 kb for the larger *GUT* 15 mRNA, to 0.7 kb for *GUT*s 7-2c and 8-3, indicating that transcript size per se is unlikely to influence transcript half-life.

GUT Sequence Similarities

Information about the possible functions of protein products of the *GUT*s can be obtained by identifying similarities between the sequences of the GUTs and known protein sequences. This information may suggest reasons why the corresponding transcript is unstable. A similarity search of deduced amino acid sequences from the ends of each *GUT* and *NOTGUT* cDNA clone against protein sequences in the databases using the BLAST algorithm [19] reveals that four of the *GUT*s (7-2A, 7-2C, 8-2A, and 8-2B) and both of the *NOTGUT*s (7-2B and 10-3) are similar to previously characterized proteins. These similarities are discussed further below. For the remaining *GUT*s (3-2, 8-1, 8-3, and 15), there were no database matches at the amino acid level, suggesting that these *GUT*s represent novel cDNAs. A subsequent similarity search of the nucleotide sequences of these *GUT* cDNAs against the nucleotide databases confirmed this.

As shown in Figure 6-3A, sequence from one end of the *GUT* 7-2A cDNA suggests that this GUT is a novel plant member of the eukaryotic family of *E. coli* DNAJ homologs. These proteins, which have been identified in yeast, *Drosophila* and human cells, are thought to interact with HSP70 proteins and play roles in protein sorting [22]. GUT 7-2A possesses each of the residues that define the N-terminal "DNAJ domain" of this family of proteins [22]. The C-terminal domains of the DNAJ homologs are generally less conserved, and are thought to modulate the specificity of their interactions with HSP70 proteins. It is not surprising, then, that sequence from the other end of the *GUT* 7-2A clone appears to have no similarity with previously characterized DNAJ proteins. Deduced amino acid sequence of GUT 7-2C (sequence from each end of the 350 bp EcoRI fragment overlaps in the middle) is similar to a fragment of a tomato wound-inducible protein (see Figure 6-3B). Although the kinetics of induction of the homologous tomato transcript following wounding have not been described, in general wound-inducible proteins are rapidly induced at the mRNA level in response to wounding in plants [e.g. 23], and therefore conform to our predictions as to the type of genes that are likely to have unstable transcripts. GUT 8-2A has restricted similarity to the extreme N-terminus of the ubiquitin conjugating enzyme E2 of *Drosophila melanogaster* (63% similarity over 44 amino acids - Figure 6-3C), but it's similarity to plant E2s is weaker. Whether or not the E2 enzymes have unstable transcripts has not been evaluated.

Legend to Figure 6-3

Nucleotide sequence data were obtained from both ends of each *GUT* and *NOTGUT* cDNA clone and compared to sequences on databases using the default parameters of the BLASTX algorithm as described in Materials and Methods. Shown here are the alignments produced from the BLAST similarity searches. (A) GUT 7-2A; (B) GUT 7-2C; (C) GUT 8-2A; (D) GUT 8-2B; (E) NOTGUT 7-2B; and (E) NOTGUT 10-3.

Figure 6-3. Sequence similarities of *GUT* and *NOTGUT* cDNAs

The strongest similarity between a GUT sequence and one on the databases is that of GUT 8-2B which, over the 600bp cloned fragment, is 99% identical to maize histone H3 (Figure 6-3D). Rapid changes in the levels of cell-cycle regulated histone genes is thought to be facilitated by the instability of their transcripts [24]. Indeed, the histone H3 transcript has previously been characterized as unstable in mammalian cells [25], so it is not surprising the corresponding tobacco gene was identified in our screen for *GUT*s. The sequences mediating the rapid degradation of animal cell-cycle regulated histone transcripts are thought to reside in the extreme 3′ stem-loop structure [26]. It is not clear whether this is also the case for plant histone mRNAs, which lack strong sequence conservation in their 3'UTRs and, unlike animal histone transcripts, are polyadenylated [27].

Both NOTGUTs are also similar to proteins in the databases. NOTGUT 7-2B is 97% similar to rat 40s ribosomal protein S5 over a stretch of 49 amino acids, and NOTGUT 10-3 is similar, but not identical, to *Arabidopsis* histone H1 (Figure 6-3 E and F, respectively). The stability of the *NOTGUT* 10-3 (histone H1) transcript is in contrast to the instability of the *GUT* 8-2B (histone H3) transcript in the same cells. One possibility that might explain this observation is that *NOTGUT* 10-3 encodes a "variant" or "replacement" form of histone H1. Unlike the cell-cycle regulated forms, the variant histones are expressed constituitively [28], and therefore mRNA instability would provide no particular advantage. However, sequence conservation among H1 histones is generally

less pronounced than is the case with other histones [27], making it difficult to differentiate between cell-cycle regulated and variant forms based on sequence information alone. Nonetheless the stability of the *NOTGUT* 10-3 transcript is consistent with a "housekeeping" role for the corresponding protein. A similar conclusion can be drawn by correlating the stability of the *NOTGUT* 7-2B transcript, with ribosomal protein S5, which it encodes.

Beyond their similarities to sequences in the databases, there are a number of other interesting features of the *GUT* cDNA sequences. First, a comparison of identifiable *GUT* 3'UTRs shows that there are no obvious conserved motifs. There are occasional AUUUA motifs, but these do not occur in highly AU-rich regions resembling the AREs, nor are they restricted to *GUT* 3'UTRs, occurring also in the *NOTGUT* sequences. While a contribution of these AUUUA elements in the instability of the *GUT* transcripts cannot be ruled out, previous studies have shown that the mere presence of an AUUUA element is insufficient to cause mRNA instability (e.g. see [11]). Secondly, *GUT* 8-1, which gives a diffuse signal on Northern blots, with consistently high background signals, has two curious repeats which are quite close together. An $(AC)_8$ motif (which, if translated, would correspond to a Threonine-Histidine repeat) is found some 50 bases upstream of a stretch of 13 C residues (potentially encoding four Prolines). The significance of these sequences is not clear, but interestingly, a search of the nucleotide databases with the *GUT* 8-1 sequence showed that the former motif is related to a human genomic

minisatellite sequence.

It is not surprising that we found no sequence conservation within the *GUT* 3' UTRs. With the possible exception of the AREs in mammals, little such conservation has been found between different unstable transcripts in animals. Recently, an extensive search of the nucleotide databases for broadly conserved elements in the non-coding regions of a large number of eukaryotic genes was performed [29]. Interestingly, the only similarities that were identified in this search were between pairs of homologous genes from different species, not between groups of otherwise unrelated genes [29]. An alternative reason for the apparent lack of sequence conservation among the *GUT*s is that the relatively small number of different *GUT* clones that were isolated in this initial screen may preclude an exhaustive analysis. The fact that only a single cDNA for each different *GUT* was identified suggests that the current collection of *GUT*s is not complete, and that many additional *GUT*s are likely to exist in tobacco cells.

CONCLUSIONS AND FUTURE PROSPECTS

The data presented in this paper demonstrate that the differential screening strategy is a promising approach for the isolation of *GUT*s in tobacco, which may be broadly applicable to the identification of genes with unstable

transcripts in other species. Using this approach we have identified ten previously uncharacterized tobacco cDNAs. These include eight *GUT*s, with mRNA half-lives in the range of an hour or less, on the same order as those of transcripts characterized as unstable in mammalian cells [3, 4]. Four of the *GUT*s represent completely novel cDNAs, whereas the other four *GUT*s, and both of the *NOTGUT*s, encode proteins with similarity to proteins in the databases. These similarities conform to our predictions concerning the kinds of proteins likely to be encoded by unstable or stable transcripts. The *GUT*s should constitute excellent tools for evaluating the mechanisms by which unstable transcripts are recognized and targeted for rapid degradation in plants. In particular, it will be most informative to delineate the sequences within the *GUT* transcripts that mediate their instability. Moreover, it should also be possible to investigate the roles of processes such as translation, deadenylation and protein binding in the mRNA decay pathways of a number of individual unstable transcripts. These efforts will substantially increase our understanding of the range of different mechanisms by which unstable transcripts are targeted for rapid degradation in plants.

ACKNOWLEDGEMENTS

We would like to thank Dr. Gynheung An for the NT-1 cell cDNA library, Steven Reiser, and Pat Tranel for isolating the *GUT* 15 clone, and helping with its sequencing, respectively, Dr. Michael Sullivan for maintaining the wild-type and transformed NT-1 cell lines, and André Dandridge and Linda Danhof for expert technical assistance. This work was supported by grants from the USDA (901167) and the DOE (FG02-90ER-20021) to P.J.G.

LITERATURE CITED

1. Green, P.J. (1993) *Plant Physiol.*, **102**, 1065-1070.

2. Sachs, A.B. (1993) *Cell*, **74**, 421-427.

3. Peltz, S.W., Brewer, G., Bernstein, P., Hart, P.A. and Ross, J. (1991) *Crit. Rev. Euc. Gene Expr.*, **1**, 99-126.

4. Brock, M.L. and Shapiro, D.J. (1983) *Cell*, **34**, 207-214.

5. Zhang, S., Sheng, J., Liu, Y. and Mehdy, M.C. (1993) *Plant Cell*, **5**, in press.

6. Seeley, K.A., Byrne, D.H., and Colbert, J.T. (1992) *Plant Cell*, **4**, 29-38.

7. McClure, B.A. and Guilfoyle, T. J. (1989) *Science*, **243**, 91-93.

8 Gee, M.A., Hagen, G. and Guilfoyle, T.J. (1991) *Plant Cell*, **3**, 419-430.

9. Newman, T.C., Ohme-Takagi, M., Taylor, C.B. and Green, P.J. (1993) *Plant Cell*, **5**, 701-714.

10. McClure, B.A., Hagen, G., Brown, C.S., Gee, M.A. and Guilfoyle, T.J. (1989) *Plant Cell*, **1**, 229-239.

11. Ohme-Takagi, M., Taylor, C.B., Newman, T.C. and Green, P.J. (1993) *Proc. Natl. Acad. Sci., USA*, in press.

12. An, G. (1985) *Plant Physiol.*, **79**, 568-570.

13. Nagy, F., Kay, S.A. and Chua, N.-H. (1990) in Gelvin, S.B., Schilperoort, R.A, and Verma, D.P.S. (eds), Plant Molecular Biology Manual. Kluwer, Dordrecht, pp B4/13 - B4/14.

14. Sambrock, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual. Cold Spring Harbor Laboratory, Plainview.

15. Taylor, C.B. and Green. P.J. (1991) *Plant Physiol.*, **96**, 980-984.

16. Delauny, A.J. and Verma, D.P.S. (1990) in Gelvin, S.B., Schilperoort, R.A, and Verma, D.P.S. (ed.s), Plant Molecular Biology Manual. Kluwer, Dordrecht, pp A14/1 - A14/23.

17. Feinberg, A.P. and Vogelstein, B. (1983) *Annal. Biochem.*, **132**, 6-13.

18. Dellaporta, S.L., Wood, J. and Hicks, J.B. (1983) *Plant Mol. Biol. Rep.*, **1**, 19-21.

19. Altshul, S.F., Gish, W., Miller, W., Meyers, E.W. and Lipman, D.J. (1990) *J. Mol. Biol.*, **215**, 403-410.

20. Herrick, D., Parker, R. and Jacobson, A. (1990) *Mol. Cell. Biol.*, **10**, 2269-2284

21. Ballas, N., Wong, L.-M. and Theologis, A. (1993) *J. Mol. Biol.*, in press.

22. Silver, P.A. and Way, J.C. (1993) *Cell*, **74**, 5-6.

23. Rohrmeir, T. and Lehle, L. (1993) *Plant Mol. Biol.*, **22**, 783-792.

24. Pandey, N.B. and Marzluff, W.F. (1987) *Mol. Cell. Biol.*, **7**, 4557-4564.

25. Schuler, G.D. and Cole, M.D. (1988) *Cell*, **55**, 1115-1127.

26. Harris, M.E., Böhni, R., Schneiderman, M.H., Ramamurthy, L., Schümperli, D. and Marzloff, W.F. (1991) *Mol. Cell. Biol.*, **11**, 2416-2424.

27. Nakayama, T. and Iwabuchi, M. (1993) *Crit. Rev. Plant Sci.*, **12**, 97-110.

28. Wu, R.S. and Bonner, W.M. (1981) *Cell*, **27**, 321-330.

29. Duret, L., Dorkeld, F. and Gautier, C. (1993) *Nucl. Acids Res.*, **21**, 2315.

CHAPTER 7

CONCLUSION

At its inception, one of the initial goals of the research project described in this Dissertation was to identify genes encoding RNases in *Arabidopsis*. This has been achieved with the identification of PCR products corresponding to three different *Arabidopsis* RNase genes, *RNS1*, *RNS2*, and *RNS3*, by virtue of their homology to fungal and S-gene RNases (Chapter 2). Each of the *RNS* genes is expressed in *Arabidopsis*, albeit to different levels, and each is present in one or two copies in the *Arabidopsis* genome (Chapter 2). With a view towards identifying functions for the RNS RNases, the expression properties of *RNS2*, the most highly expressed of the *RNS* genes, were explored (Chapter 3). These experiments showed that, like the S-RNase genes to which it is related, *RNS2* is most abundantly expressed in flowers of *Arabidopsis*. However, unlike the S-RNase genes, *RNS2* is also expressed at some level in all other organs that were tested (Chapter 3). The most informative data pertaining to the identification of potential functions for the RNS2 enzyme, came from experiments demonstrating that *RNS2* expression increased markedly during senescence, in both petals and leaves of *Arabidopsis* (Chapter 3). This suggests that RNS2 may play a role in the remobilization of nutrients, such as phosphate, from RNA. The increased expression of *RNS2* in *Arabidopsis* seedlings starved for phosphate supports this contention (Chapter 3).

The *RNS2* PCR product was also used as a probe to identify full length *RNS2* cDNAs. These clones were used in a heterologous expression system to demonstrate that *RNS2* encodes an active RNase (Chapter 3). Furthermore,

comparison of the deduced amino acid sequence of RNS2 to those of related fungal and plant RNases suggested that there are two distinct subclasses of RNases within the plant lineage of the T_2/S RNase superfamily. These are the S-RNases, which are involved in self-incompatibility in the Solanaceae, and the S-like RNases (including RNS2), which are expressed in self-compatible species. The placing of RNase LE (from tomato) within the S-like RNase lineage suggests that the S-like and S-RNases diverged before speciation (Chapter 3). As sequences of more S-like RNases have become available (eg RNS1, RNS3, and a second tomato enzyme, RNase LX), this conclusion has been extended, and it now seems likely that divergence within the S-like lineage also preceded speciation. This is because RNS1 appears to be more similar to LE, and RNS3 to LX, than either is to the other, or to RNS2 (Bariola, P.A. and Green, P.J., unpublished).

RNS2 was among the first senescence-inducible and phosphate starvation-inducible genes identified for which a specific function (phosphate remobilization) could be proposed. The *RNS2* cDNA therefore constitutes a novel molecular probe for the study of senescence in plants, and for the characterization of plant responses to nutrient limitation. For example, it should be possible to identify regions of the *RNS2* promoter that mediate the induction of this gene under phosphate limiting conditions, or during senescence. Understanding plant responses to nutrient limitation is of fundamental importance to agriculture. Ultimately, the study of the roles of RNS2, and of

RNS1, which is even more strongly induced by phosphate starvation (Bariola, P.A. and Green, P.J., unpublished) may lead to increased understanding of the roles of RNases in these responses.

In the long term, this work is directed towards the identification of those RNases with a specific role in mRNA catabolism. Based on their sequences, RNS1 and RNS3 are thought to be extracellular (Bariola, P.A. and Green, P.J., unpublished), and RNS2 may be targeted to the vacuole in *Arabidopsis* (Chapter 3). These presumed localizations of the RNS enzymes are difficult to reconcile with a role in the critical initial steps in mRNA degradation, which are presumably cytoplasmic. However, small fragments of RNA have been found in plant vacuoles [Abel, S., Blume, B. and Glund, K. (1990), *Plant Physiol.* **94**, 1163-1171], so it is possible that later steps in mRNA catabolism may occur in that organelle. In the future, it should be possible to adapt the PCR approach initially used to identify *RNS1*, *RNS2*, and *RNS3*, to explore the possibility that cytoplasmic S-like RNases exist in *Arabidopsis*. Such enzymes would be of great interest, as very few intracellular RNases have been identified in eukaryotes. Cytoplasmic RNases are likely to be pivotal components in the decay pathways of highly unstable transcripts in plants, either directly involved in the recognition of such transcripts, or as the means for their rapid degradation. In the later case, RNases may be recruited to unstable transcripts via their interactions with factors bound to specific sequence elements within those transcripts.

The search for, and identification of, sequence determinants that confer instability on transcripts in tobacco was described in Chapters 4, 5 and 6 of this Dissertation. An essential initial step was the development of a system for the accurate measurement of reporter transcript half-lives in stably transformed tobacco (NT-1) cells (Chapter 4). That mRNA half-lives measured in the NT-1 cells were likely to reflect those in green plants was suggested by experiments demonstrating that control transcripts disappeared with similar kinetics when expressed in transgenic plants under the transcriptional control of a regulated promoter (*Cab-1*), as they did in NT-1 cells (Chapter 4). Using the NT-1 cell system, two very different sequences that confer instability on reporter transcripts were identified. Repeats of the DST element (Chapter 4), and the AUUUA element (Chapter 5), were each found to markedly destabilize mRNAs in the NT-1 cells. Moreover, each of these sequences caused a pronounced decrease in the accumulation of the reporter transcript in the leaves of transgenic tobacco, suggesting that both also function as instability determinants in plants, as they do in NT-1 cells. (Chapters 4 and 5).

Having identified the DST and AUUUA sequences as potent determinants of mRNA instability, it is of interest to begin to address the mechanisms by which they function. An important first step is to identify and characterize components of the RNA degradation machinery that specifically recognize these elements. One approach to this end is to search for factors that bind to the AUUUA or DST elements in vitro. This strategy has proven particularly

effective in the identification of a number of mammalian ARE-binding proteins. However, it is also necessary to identify the specific sequences within each instability determinant that are necessary for them to confer instability on transcripts. Once this is achieved, it should then be possible to correlate the binding of proteins to active or inactive elements with the impact of those same elements on the half-lives of reporter transcripts.

Genetic approaches toward the identification of factors involved in the recognition of unstable transcripts are also available in plants. For example, it may be possible to identify mutant *Arabidopsis* plants that are unable to recognize a particular instability determinant. One way this could be achieved is to place an instability determinant within the transcript of a selectable marker gene. *Arabidopsis* plants transformed with this construct should be sensitive to the selectable marker, owing to the instability of the corresponding transcript. In contrast, mutant progeny of these plants in which the instability determinant is no longer recognized as such, should be able to survive selective conditions. The major advantage of this strategy is that it may lead to the identification of factors that are not directly bound to the unstable transcripts, but still participate in their recognition and/ or degradation. Moreover, genetic approaches are not yet feasible in mammalian cells, and their application in plants may therefore be of particular relevance to the characterization of factors that interact with the AUUUA repeat. This element is similar to the AREs that destabilize transcripts such as c-*myc* and c-*fos* in mammalian cells. In the long

run, it will be interesting to address the specificity of factors identified using these approaches by examining their interactions with other instability determinants. These experiments will help to determine the extent to which mechanisms targeting a variety of unstable transcripts for rapid degradation in plants are shared or unique. Future identification of endogenous instability determinants within the *GUT* transcripts (Chapter 6) will facilitate these analyses.

Concluding Remarks

Using the tools that have been described above, it may well be found that many features of mRNA degradation pathways are shared between plants and animals. Exploration of these potential similarities can only benefit from the application of genetic approaches toward identifying components of mRNA degradation pathways in *Arabidopsis*. Nonetheless, given their more fundamental requirement for rapidly regulated gene expression, it is also likely that plants will have developed unique and fascinating mechanisms to facilitate the regulation of gene expression at the level of mRNA stability.

The work described in this Dissertation provides the foundation for future efforts aimed at elucidating mRNA decay pathways in plants. Ultimately, the analysis of sequence determinants of stability will converge with efforts to

characterize plant RNases, as those RNases with a specific role in the degradation of mRNAs are identified. This may be achieved via the isolation of cytoplasmic S-RNase homologs in *Arabidopsis*, or through the continuing characterization of cellular factors that bind to specific sequence elements. Taken together, these approaches should provide fundamental information about the mechanisms that target transcripts for rapid degradation in plants.