1991

Translational Repression of Bacteriophage T4 DNA Polymerase Biosynthesis

Mark D. Andrake
Medical University of South Carolina

TRANSLATIONAL REPRESSION
OF
BACTERIOPHAGE T4
DNA POLYMERASE BIOSYNTHESIS

by

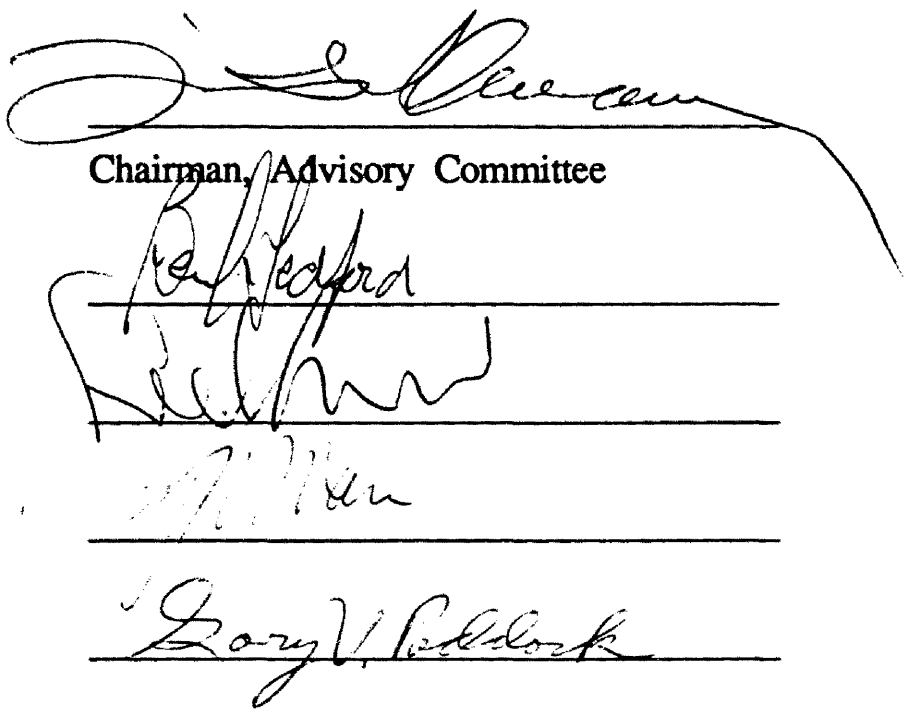
Mark D. Andrake

A dissertation submitted to the faculty of the Medical University of South Carolina in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the College of Graduate Studies.

Molecular and Cellular Biology & Pathobiology Program

1991

Approved by:



Chairman, Advisory Committee

MARK DAVID ANDRAKE. Translational Repression of Bacteriophage T4 DNA Polymerase Biosynthesis. (Under the direction of DR. JIM KARAM.)

The research described in this dissertation elucidated the mechanism by which bacteriophage T4 DNA polymerase regulates its own biosynthesis. Utilizing both *in vivo* and *in vitro* studies, I have shown that autogenous repression occurs at the level of translation. While T4 mutants defective in the structural gene for DNA polymerase (gene *43*) overproduce the protein product (gp43) *in vivo*, they do not overproduce the corresponding mRNA. *In vitro*, purified DNA polymerase *specifically* inhibited the translation of its own transcripts. Further, it was demonstrated that gp43 binds its own mRNA at a site overlapping the ribosome initiation domain. Thus, T4 DNA polymerase is a specific translational repressor that presumably inhibits initiation of translation.

The mRNA binding site (translational operator) for DNA polymerase includes 38-40 nucleotides upstream of the initiator AUG. The 5′ half of this translational operator contains a putative five base-pair stem and 8-base loop, whose existence is inferred from RNase digestion experiments and computer-assisted analysis of RNA folding. To ascertain the important RNA sequence and structural determinants for DNA polymerase binding, I carried out a mutational analysis of the translational operator via the *in vitro* construction

of several operator variants. Operator mutants were subsequently assayed for the effect of each mutation on: 1) gp43/mRNA binding, *in vitro* 2) the *in vivo* levels of gp43 biosynthesis from plasmid encoded constructs and 3) *in vivo* level of gp43 synthesis in phage infections (carried out after introducing mutant operators into the phage genome by virus-plasmid recombination).

Mutations that either disrupted the stem or altered particular loop residues, led to diminished binding of purified T4 DNA polymerase *in vitro* and to derepression of protein synthesis *in vivo.* Compensatory mutations that restored the stem pairing, with a sequence other than wild-type, restored *in vitro* binding but still exhibited a mutant phenotype *in vivo.* Results from loop substitutions suggest that the spatial arrangement of specific loop residues is a major criterion for specific binding of DNA polymerase to its mRNA operator. These studies demonstrate the effectiveness of genetic approaches in dissecting the rules that govern RNA - protein interactions.

TABLE OF CONTENTS

Illustrations

Tables

Acknowledgements

I would like to acknowledge the students and staff of the Karam lab with whom I have shared my graduate experience. While the work leading to a doctoral degree is a tremendously individual effort, the team work of the pursuit cannot be overlooked. The scientist one eventually becomes is a strange mixture of the influence of many persons. I'd like to thank all those who have been a part of the lab; in particular, I am indebted to my two predecessors Melanie Hughes and Tien Hsu, who had a prominent influence on my early work, and were the major contributors to a positive learning atmosphere.

I would like to thank each member of my advisory committee for their help, guidance, and flexibility in aiding my development as a graduate student. Of course, I am deeply grateful to Jim Karam for his continual challenge to excellence, his respect for me as a person, and genuine concern for my development and growth as a scientist.

Most of all, my heartfelt thanks goes to my patient wife, Sally Jean, for her constant love and support through the rigorous and often frustrating lifestyle of a graduate student. It is impossible to estimate the worth of the love and care that upholds you each day.

Finally, just as a composer is praised by those who listen to the beauty of his music, so I'd like to thank the Author of this wonderful, intricate creation which we all love to explore.

I) INTRODUCTION

A) ROLE OF TRANSLATIONAL REGULATION IN PROKARYOTIC GENE EXPRESSION

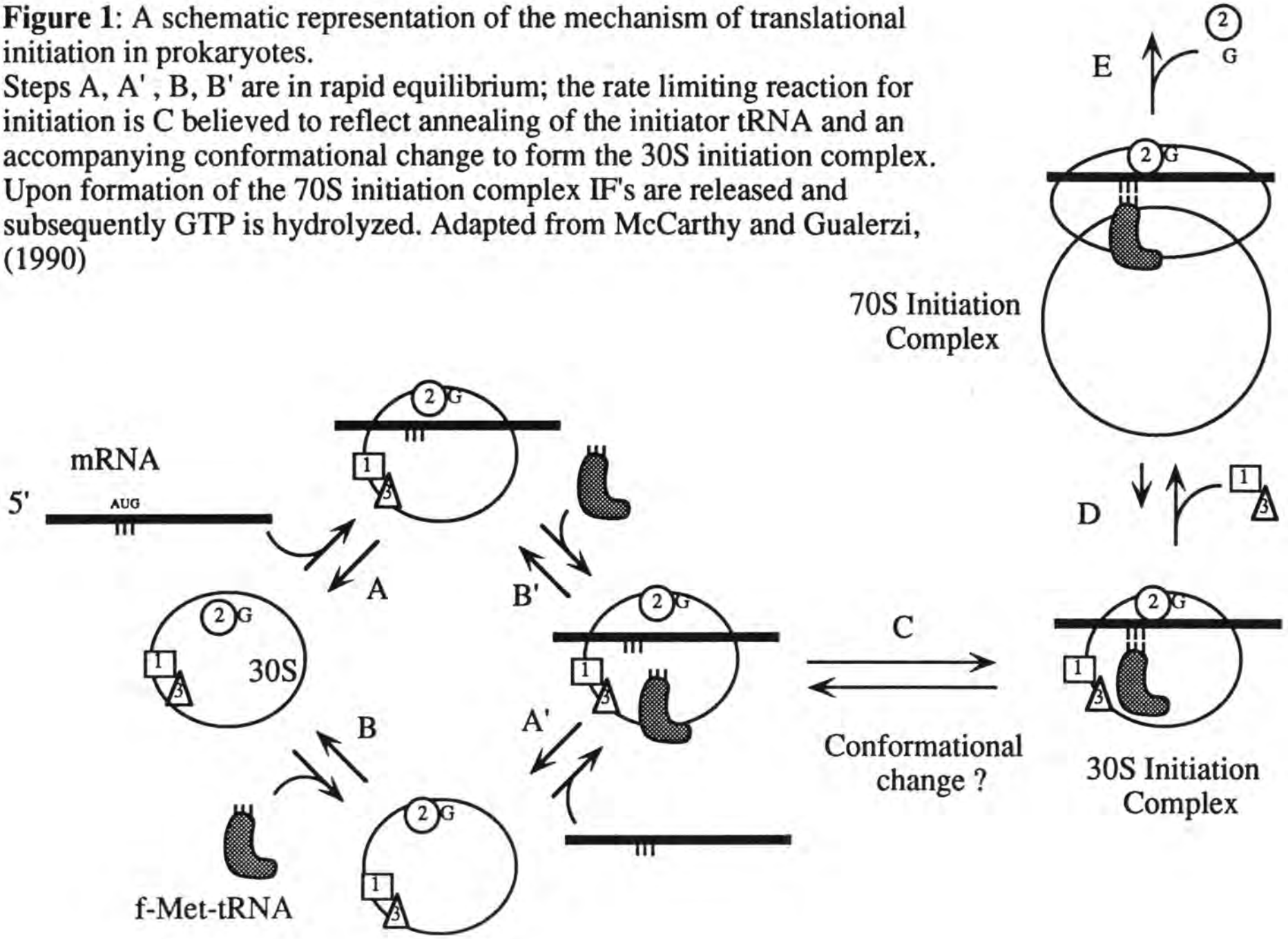
Early studies in molecular genetics established the importance of transcriptional mechanisms of gene regulation, and several aspects of transcription are now well characterized (McClure, 1985). The DNA-binding proteins that regulate transcription have been the object of intense study, and the principles underlying the binding of proteins to DNA regulatory elements have been described in detailed molecular terms (Schleif, 1988). In contrast, studies of translational regulation in prokaryotes have attracted major interest only in recent years. Despite the relatively short half-life of most prokaryotic mRNAs, it is clear from many examples that translational regulation plays a key role in prokaryotic gene expression. In *E. coli*, there are more than 250 genes that specify the macromolecules of the translational apparatus, which encompasses more than 50% of total cellular weight under favorable growth conditions. Coordinate regulation of these diverse components is of prime importance, and much of this regulation occurs at the translational level. To

discuss translational regulatory mechanisms, it is beneficial to first describe the process of translation itself. Several excellent recent reviews may be consulted for details {e.g. eukaryotic and prokaryotic translation (Kozak, 1983) and translational regulation in prokaryotes (Gold, 1988; McCarthy and Gualerzi, 1990)}. The focus of my discussion in the next two sections is a brief review of the events of translation initiation, since most forms of translational regulation occur at this stage. The remaining sections review the variety of mechanisms that are employed to regulate translation and highlight several examples of translational repressors.

B) THE MECHANISM OF TRANSLATION INITIATION

In the scheme for translation initiation proposed by Gualerzi (McCarthy and Gualerzi, 1990), the components of prokaryotic translation initiation include the 21 r-proteins and 16S rRNA contained in the small 30S subunit of the ribosome, the three known translation initiation factors, GTP, the initiator tRNA (f-MET-tRNA) and the mRNA harboring a translation initiation region (TIR) (see Figure 1). The 30S ribosomal subunit complexed with one of each initiation factor binds in random order the mRNA TIR and the initiator tRNA to form the so-called pre-ternary complex. These different bound complexes of the 30S subunit exist in rapid

Figure 1: A schematic representation of the mechanism of translational initiation in prokaryotes.
Steps A, A' , B, B' are in rapid equilibrium; the rate limiting reaction for initiation is C believed to reflect annealing of the initiator tRNA and an accompanying conformational change to form the 30S initiation complex. Upon formation of the 70S initiation complex IF's are released and subsequently GTP is hydrolyzed. Adapted from McCarthy and Gualerzi, (1990)



equilibrium (reactions A and B in Figure 1), and the rate limiting step in translation occurs when there is a first-order conformational rearrangement of the pre-ternary complex to form the 30S initiation complex (reaction C in Figure 1). It has been presumed that this conformational change reflects the base pairing of the initiator tRNA anticodon to the mRNA initiator codon, among other changes in the ribosome itself. This ternary complex either dissociates or goes on to form a 70S initiation complex by the recruitment of the 50S subunit and release/recycling of the initiation factors, with accompanying hydrolysis of GTP. Although the rate of elongation may affect the maximal translational capacity of the cell, initiation is thought to be the most important rate limiting step for the entire process (Schauder and McCarthy, 1989). Correspondingly, it is reasonable that the most common forms of translational regulation target this rate-limiting step of initiation.

Figure 2 shows a diagrammatic representation of the TIR. Note the distinction being made between the "ribosome binding site" (RBS) and the "TIR". The RBS can be defined for our purposes here as the residues that are protected from RNase digestion by the 30S subunit in the initiation complex, *i.e.* 30-35 residues surrounding the initiation codon (Steitz, 1979). The generally wider TIR may be defined as that region on the mRNA molecule which determines both the <u>site</u> *and* <u>efficiency</u> of initiation. Given the role that higher order

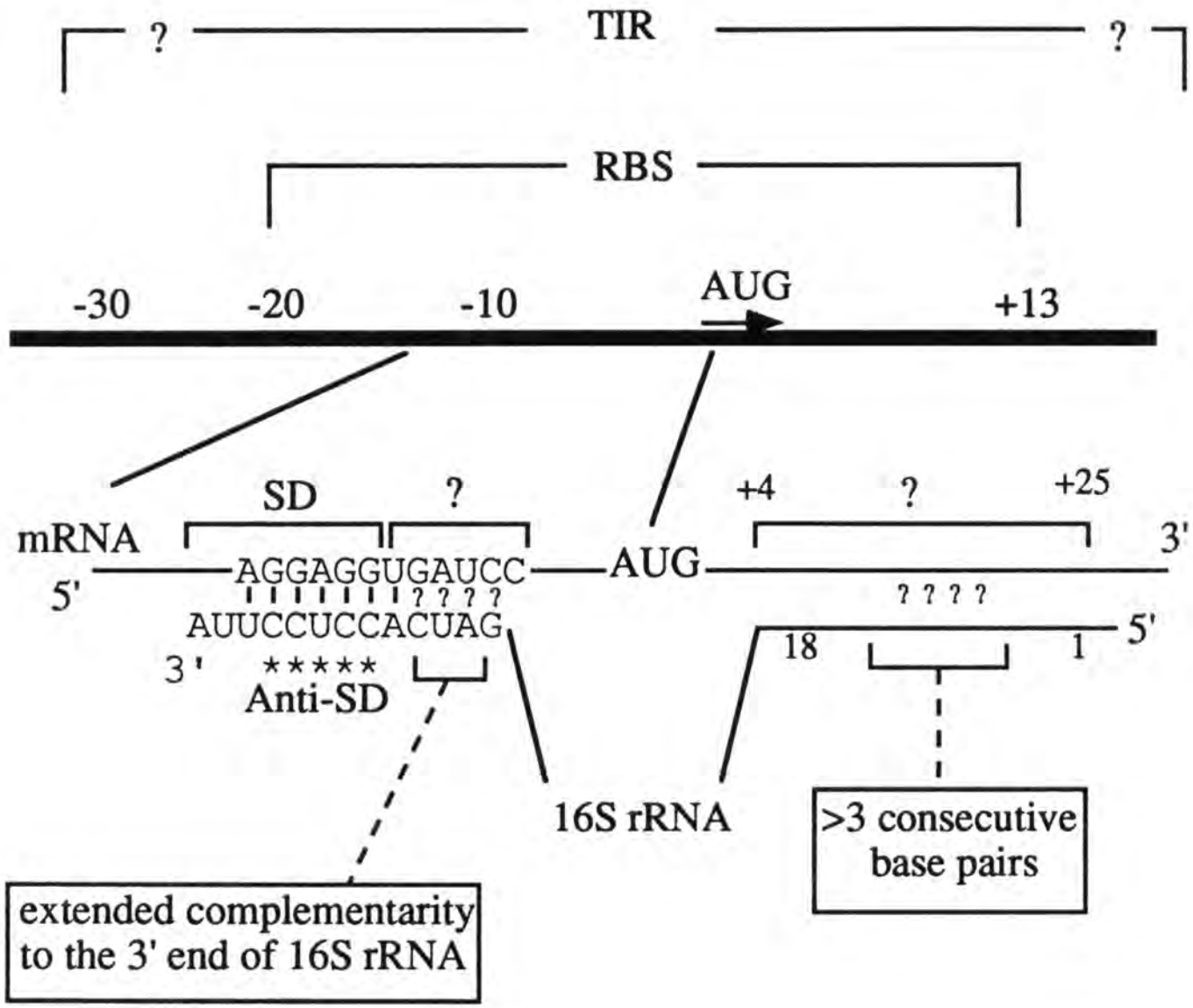


Figure 2 : Schematic representation of the TIR. Sequences important for translation initiation have been highlighted, as well as known or putative interactions between 16S rRNA and the mRNA TIR. Adapted from McCarthy and Gualerzi, (1990).

RNA structure must play in initiation, the RBS (as defined above)may or may not encompass all residues of the mRNA TIR that enhance or are essential for initiation. Another approach to defining the TIR is to look at the "relative nonrandomness" of the residues surrounding the start codon (Schneider, et al., 1986). This statistical approach yields a TIR roughly equivalent in size to the RBS; the fact that each base position in this 35-base region exhibits nonrandom sequence variance indicates that most residues play a role in contacting the ribosome. The relationship of initiation efficiency to the sequence variation that is observed in this region is yet to be clearly defined.

C) THE TIR:STRUCTURE AND FUNCTION

Sequences that are definitively involved in recognition by the 30S subunit are the initiator codon (most often AUG (~90%), sometimes GUG (~8%) and rarely UUG) and the so-called Shine-Dalgarno region (abbreviated SD). In 1974, Shine and Dalgarno sequenced the 3′ end of 16S rRNA, and proposed that this region base paired with the conserved bases found upstream of the AUG (Shine and Dalgarno, 1974). Since then, an overwhelming body of evidence has accumulated supporting this hypothesis, including elegant experimental systems that mutate a plasmid encoded 16S rRNA at the anti-SD region, and show selective translation of mRNAs that have complementary mutations in the SD region of their TIR (Hui and de Boer, 1987; Jacob, et al., 1987). Further clarification of the

role of the SD complementarity has shown that the primary function of the SD base pairing is to enhance the *kinetics of formation* of ternary complexes (Calogero, et al., 1988) and does not affect the mechanism or order of complex formation (*i.e.* there is a random order of component binding and a rapid equilibrium for each of the reactions A,B,A',B' shown in Figure 1). The most likely function for the SD sequence is to provide high affinity interactions with the 30S subunit (>10-fold higher affinity than mRNAs without the SD), so that the effective concentration of initiator codons near the peptidyl-tRNA decoding site is drastically increased.

However the initiator codon and the SD region do not specify all the information used in ribosome binding (Schneider, et al., 1986). The nonrandomness of the TIR extends to about +13 downstream and -20 upstream of the initiator codon. Gold (1988) suggests that there are two simple interpretations of this : 1) that the primary sequence of this whole region is inspected by the ribosome in binding, or 2) that sequences surrounding these key initiation sequences are nonrandom in order to keep the TIR free of competing secondary structure, false initiator codons and SD sequences.

Other proposed primary sequences that affect translational efficiency include "translational enhancers" (Olins and Rangwala, 1989; Schauder and McCarthy, 1989). These were described when special RBS sequences were inserted upstream of heterologous genes and shown to cause a pronounced

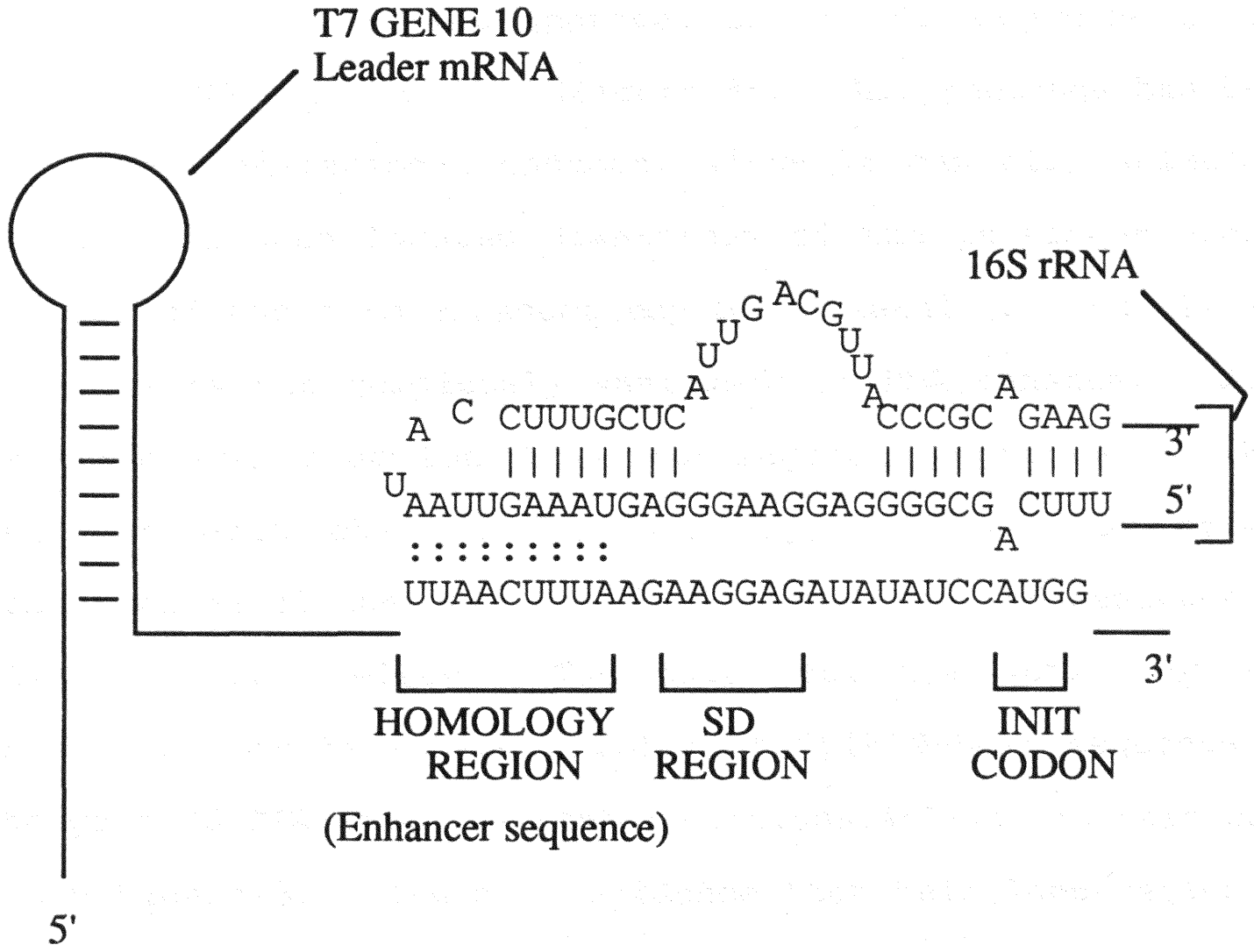


Figure 3: Model for action of gene 10 mRNA leader enhancer sequence The model reflects previously determined rRNA secondary structure (Noller and Woese, 1981). The residues of 16S rRNA portrayed are positions 437-497, and the interaction between the SD region and the anti-SD of 16S rRNA has been omitted for clarity. The enhancer-rRNA interaction may involve some disruption of the top pairs of rRNA stem. Adapted from Olins and Rangwala, (1989).

stimulation of expression. The effector portion of this RBS (described in two cistrons originally, *atpE* and T7 gene *10*) is an A/U rich segment upstream of the SD sequence in the leader of T7 gene *10* mRNA (Figure 3). This sequence has been termed a translational enhancer since it can also stimulate translation when located downstream of the initiator codon. The use of the term enhancer may be misleading, since 1) its function is not completely analogous to DNA enhancers (*i.e.* it does not function to bind regulatory proteins that interact with the translational apparatus and/or affect conformation of the RNA) and 2) the position independence is likely to be limited. The model for the action of the "enhancer" involves base pairing of this 9-base sequence in the gene *10* TIR with 16S rRNA (positions 455-462 in this case - see Figure 3). There is evidence that this loop region of rRNA is relatively exposed and could be available for interaction with mRNAs (Moazed, et al., 1986). Although direct evidence for this interaction between 16S rRNA and the enhancer sequence is forthcoming, the mechanism represents a unique form of translational regulation where increased interaction between mRNA and rRNA can provide a way for specific messages to compete more effectively for the ribosomal machinery.

Other possible recognition elements of the TIR include 1) complementarity of TIR residues to bases adjacent to the anti-SD sequence in 16S rRNA (Thanaraj and Pandit, 1989) and 2) extended complementarity of the bases surrounding the

initiator codon to the anticodon loop of $tRNA^{f\text{-}MET}$ (Ganoza, et al., 1985), and 3) complementarity of at least 3 sequential residues in the region +4 to +25 to the bases at the 5′ end (positions 1-18) of 16S rRNA. Of the *E. coli* genes that have AUG or GUG initiator codons, 98% contain a sequence of at least three consecutive nucleotides (67% have a sequence of at least four consecutive nucleotides) that are complementary to some of the 16 nucleotides at the 5' terminus of the 16S rRNA (Petersen, et al., 1988 - see Figure 2). Since the enhancer sequence described above can function 5′ or 3′ to the initiator codon, the question of the structure and flexibility of 16S rRNA in the mRNA cleft is raised. In addition, if the enhancer is located downstream of the AUG, it may overlap with the TIR region proposed to interact with the 5′ end of 16S rRNA (see Figure 2, mentioned above). Both of these interactions need not be ruled mutually exclusive, since rRNA exists in a highly folded structure where distant portions of the linear molecule are juxtaposed. However, it remains to be determined if these interactions between mRNA and rRNA represent alternate mechanisms to enhance the initiation of specific messages or possibly a more general interaction seen in most messages. The relative importance of these, and other potential interactions between mRNA and the ribosome, will be shown upon further description of the topology of 16S rRNA within the ribosome as well as the location and action of other key elements in the mRNA cleft of the ribosome. While these proposed interactions need

further experimental verification, they do represent models that explain the nonrandom sequence variation of the TIR.

The role of secondary structure of the TIR is less clearly defined. Early sequence comparisons indicated a preference for A and T residues (which are less likely to form stable base pairs compared to G-C rich regions) between the SD and initiator codon, as well as immediately downstream of the AUG (Dreyfus, 1988; Gold, et al., 1981). Some experimental results seem to indicate that stable higher order structures, which base-pair the SD or initiator codon residues, result in a decreased efficiency of translation (Schauder and McCarthy, 1989, Knight, 1987 #175). Our understanding of the role that RNA structure plays in the accessibility of initiation sequences is still preliminary. Extensive analysis is yet to be done on model RNAs that are well controlled for the presence of initiation sequences and whose structures can be verified by physical solution mapping studies. It is difficult to create an experimental design that isolates the effects of changes in initiation sequences (*i.e.* the SD) from secondary structure effects, and visa versa. Extreme care must be taken in interpretation of results, and the construction of proper controls. Some known mechanisms of translational regulation involving RNA secondary structure will be discussed below.

D) THE FUNCTIONAL ROLE OF rRNA IN OTHER ASPECTS OF TRANSLATION

In recent years, appreciation for the functional role of rRNA has grown; Dahlberg (1989) has reviewed some of these aspects in detail. In addition to the initiation tasks described above, 16S rRNA complementarity to mRNA may function also in elongation (Trifonov, 1987; Weiss, et al., 1988) and termination.(Murgola, et al., 1988). In the case of the *RF2* gene in *E. coli*, Weiss has shown that efficient frameshifting in this gene requires pairing between the anti-SD sequence of 16S rRNA and a sequence just upstream of the frameshift site, which aids the repositioning of the ribosome for the new frame. Trivonov *et.al* (1987) have postulated a role of the 3′ end of 16S rRNA in the maintenance of the proper frame. These results imply that the 3' end of 16S rRNA scans the mRNA very close to the decoding site also during elongation. Additional functional roles of rRNA include structural formation of the decoding site (Moazed and Noller, 1986) and the peptidyl transferase site, and the enhancement of 30S-50S subunit association through base pairing with 23S rRNA (Tapprich, et al., 1989). When more is known of the structure of the peptidyl transferase site, models may be formulated showing the action of rRNA in the catalysis of peptide bond formation. Since transpeptidation is not unlike transesterification (already shown to be

catalyzed by RNA), a potential catalytic role for rRNA in protein chain elongation is not unfeasible. All of these functional roles for rRNA contrast with the classical perception that relegated rRNA to a scaffolding purpose, to hold catalytic protein moieties in the proper topological distribution. The more recent insights reflect the increased appreciation of the enzymatic roles of RNA during the 1980s, and also highlight the wonderful complexity of the ribosome.

The ribosome exhibits amazing flexibility in decoding nonstandard mRNAs in a variety of ways. Some of these have been reviewed in detail (Atkins, et al., 1990). The ability of the ribosome to "hop" has been shown in several cases. For example, the sequence CUU UAG CUA (Leu stop Leu) can be decoded with 1% efficiency as a single Leu. In addition the ability of the ribosome to hop over longer distances has been shown in the case of T4 gene *60*. This gene contains an interesting intron that is not spliced but rather just skipped during translation. The requirements for skipping include a unique secondary structure, matched codon sets on the intron borders, and a stop codon at the 5′ junction of the gap. These features are diagrammed in Figure 4 (after Huang, et al., 1988). In addition, the ribosome seems to respond to a variety of signals and stimulators as it participates in frameshifting and readthrough of stop codons. In some cases involving stop-codon readthrough, the inserted amino acid (*i.e.* selenocysteine) has a specific physiological role (Atkins, et al., 1990).

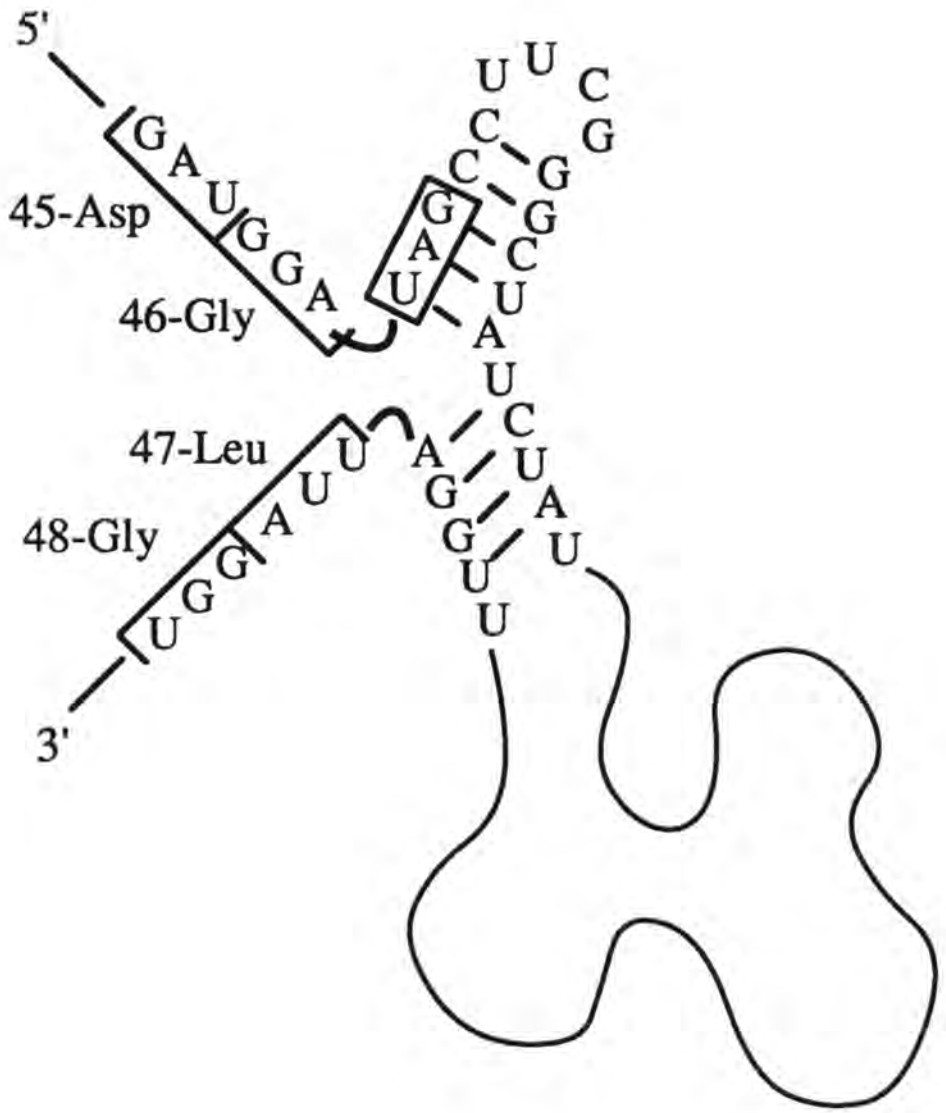


Figure 4: Bacteriophage T4 gene *60* intron
This unique intron is not spliced; rather the ribosome appears to just "hop" over the intron. The figure shows the model that accounts for the action of the ribosome in jumping from codon 46 to codon 47. Note requirements for the UAG termination codon immediately down from the 5' junction of the intron(boxed), matched Gly codons on both sides of the intron, and the proposed RNA structure. Adapted from Huang *et al.*, (1988)

In summary, the ribosome is a 2,520 kD multi-enzyme protein factory that orchestrates the action of 52 proteins and 3 rRNAs to translate a diverse set of mRNAs, and as such represents one of the most refined and intricate examples of molecular cooperation in the physiology of the cell.

E) MECHANISMS OF TRANSLATIONAL REGULATION

Given the outline of translation initiation presented above, the question arises: how is this refined cellular process regulated? Since initiation is the rate limiting step in translation, it is not surprising that most forms of regulation are directed at this stage of the process. Any mode of regulation that hinders the accessibility of any component of the 30S ternary complex can potentially hinder the kinetics of formation, and therefore, the efficiency of translation. In the following sections, I review several examples that illustrate different mechanisms known to occur in translational regulation.

1) RNA STRUCTURE INHIBITING/ENHANCING TRANSLATION

Since an important step in the initiation process involves annealing between 16S rRNA and SD residues in the TIR (and possibly the other proposed interactions with the TIR noted above), it is likely that any stable intramolecular pairing of these sequences with other regions of the mRNA will render the TIR less available for association with the 30S subunit. How significant is this inhibitory

intramolecular pairing to the rate of initiation? A structure with a $\Delta G=0$ will half the amount of time that a RBS exists in an accessible form for annealing to 16S rRNA. Each decrease of ΔG by 1 kcal/mole reduces the time of "open configuration" by at least five-fold (Gold, 1988). The decrease in ΔG obtained by the addition of a single base pair to a helix is between 1 and 3 kcal/mole. Therefore, the role of minor (in the range of -1 to -4 kcal/mole) RNA structures, which sequester initiation sequences, in initiation efficiency is likely to be more important than first imagined. This also highlights the need for care when interpreting the results of experiments with constructed TIR mutants. Often, one might misrepresent a mutation to affect changes in essential initiation sequences when in fact the primary effect is on RNA structure of the TIR. The current working hypothesis for this relationship may be explained as follows: given a theoretical maximal rate of initiation, which is dictated by a combination of a defined SD region and start codon, the actual rate of translation is a function of the degree to which the TIR higher order structure hinders initiation (McCarthy and Gualerzi, 1990).

Several examples of inhibiting structures exist in the literature; the story of antibiotic induction of methylase translation illustrates the importance of RNA structure as well as how alteration of the ribosome can affect gene expression. Normally the RBS of the erythromycin methylase gene is sequestered in a stable secondary structure, and only

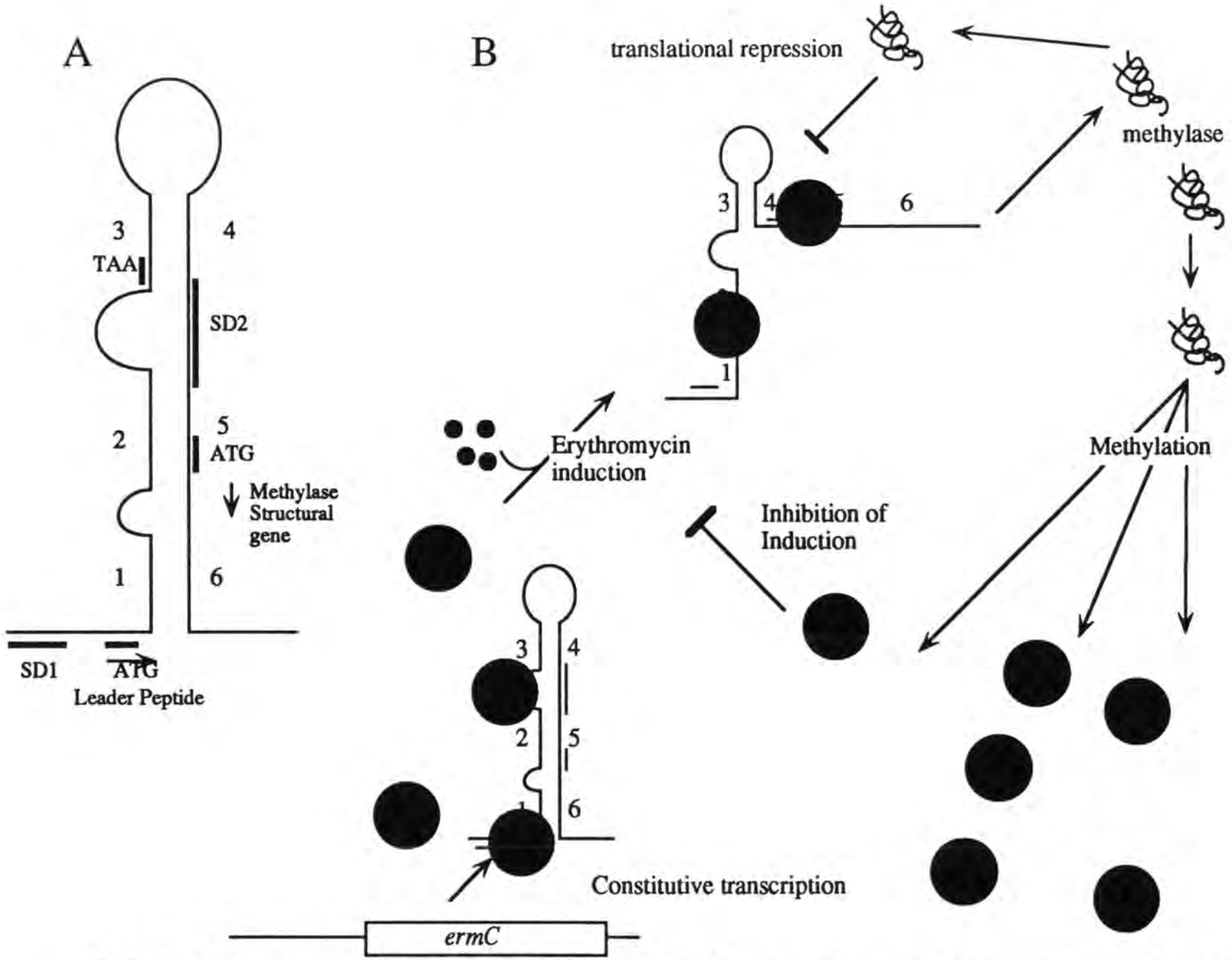


Figure 5: Summary of *ermC* regulation Shaded balls symbolize ribosome, which can be methylated (M) or bound by erythromycin(small dark circles). See text for details. Adapted from Breidt and Dubnau, (1990)

synthesized at basal levels. When the ribosome is altered by erythromycin binding, it stalls at certain amino acid codons. These particular codons are contained in a short 19 codon open reading frame (which is required for Em induction) located upstream of the methylase RBS (see Figure 5A). As the ribosome stalls in translating this ORF it opens the inhibitory secondary structure involving the methylase RBS, allowing translation of the downstream methylase gene (Hahn, et al., 1982) (see Figure 5B). In addition, two other modes of translational regulation play a role in the expression of the *ermC* methylase gene 1) autogenous translational repression (Breidt and Dubnau, 1990) and 2) inducible mRNA stability. These mechanisms are summarized in Figure 5B. The expression of this gene remains an elegant example of overlapping modes of translational regulation that have been refined to endow an organism with the ability to adjust to varying growth conditions in the presence of antibiotics.

It is also conceivable to think of secondary structure as enhancing translation. Secondary structure could rigidly hold initiation sequences of the TIR in an open conformation, for example in a large loop. Alternatively, it has been postulated that a stable stem-loop hairpin brings together the initiator AUG and the SD sequence in the proper spacing for efficient initiation, in the case of T4 gene *38* mRNA (Gold, 1988). More studies on the higher order structure of TIRs will probably unveil mechanisms by which structure can enhance initiation.

Another variation on the RNA-mediated repression of initiation sequences is the intermolecular pairing of antisense RNA (Inouye, 1988). The classic example of this occurs in the *ompF* gene of *E. coli* where divergent transcription produces an antisense RNA (called micRNA - for mRNA interfering complementary RNA) which regulates the translation of ompF RNA by sequestering the SD region of the TIR. The role of antisense RNA is also essential for the regulation of genes responsible for transposition (IS10 transposase) and plasmid maintenance (*hok/sok* system - see review by Inouye 1988).

2) TRANSLATIONAL APPARATUS-MEDIATED REGULATION

Translational coupling between adjacent cistrons is a common example of translational regulation mediated by the ribosome itself. Many of these examples also involve TIR secondary structure of the coupling-dependent cistron. Others involve inefficient initiation sequences that require "reinitiation" from ribosomes terminating at a nearby upstream cistron. Several examples have been reviewed. (Lindahl and Zengel, 1986). One interesting example that combines protein mediated repression with a long-range translational coupling involves the *E. coli rplJL* operon (Petersen, 1989). By deletion analysis, translational coupling (and expression of the downstream *rplL*) was shown to be dependent on translation of sequences over 500 bases upstream in the beginning of the *rplJ* gene. In addition to this coupling, it is known that this operon is subject to

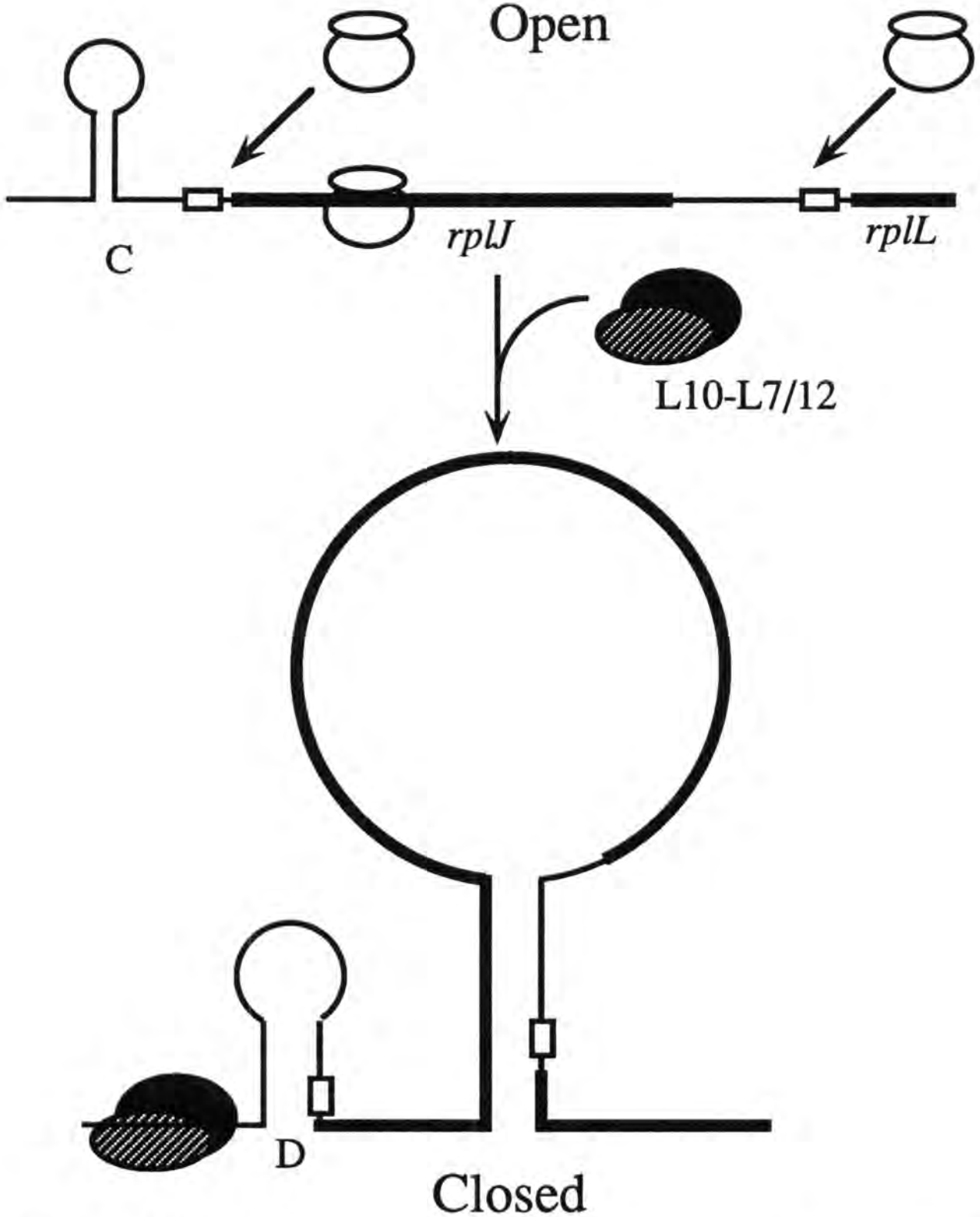


Figure 6: Model for coordinated repression of translation of the *rplJL* mRNA. In the native conformation, the operon mRNA is assumed to be in an open configuration, with both RBSs accessible for initiation of translation. The model presumes that when the L10-L7/L12 complex binds the mRNA region upstream of *rplJ*, it destabilizes the C hairpin structure and the *rplJ* RBS is blocked by a second hairpin loop, D. When the *rplJ* cistron is not being translated, the long-range RNA structure is proposed to fold up and block translation of the downstream *rplL* cistron. Adapted from Petersen, (1989)

repression by the encoded products L10(*rplJ*) and an L10-L7/L12(*rplL*) complex. Figure 6 presents the integrated model for this circuit of translational repression. This scheme accounts for the regulation of two cistrons by the binding of repressor to a single mRNA site.

The autoregulation of translation initiation factor IF3, encoded by the *infC* gene, represents an interesting example of how the primary function of an essential translation protein is utilized by the cell to maintain the proper levels of that protein. The function of IF3 in translation is to enhance ternary complex formation, but also to destabilize non-initiator aminoacyl-tRNAs binding to the 30S subunit during initiation. The start codon of the *infC* mRNA is a very rare AUU codon, which is probably very inefficient in initiation, especially in the presence of IF3 protein promoting initiator tRNA fidelity. Translation initiation on *infC* mRNA depends on extensive interaction (20 base pairs) between TIR sequences and 16S rRNA, possibly compensating for the kinetic deficit imposed by a nonstandard initiator codon. Given this odd initiator codon it was predicted and later shown that this start codon performed a necessary role in the regulation of IF3 levels. Changing the AUU to the common AUG start codon abolished autoregulation by IF3 protein (Butler, et al., 1987). The model for IF3 autoregulation predicts that when IF3 levels are low (relative to the number of 30S subunits), IF3 independent translation of *infC* mRNA will

occur preferentially over IF3 dependent translation of other cellular mRNAs, and result in an increase of the relative amount of IF3. Therefore, the regulatory mechanism utilizes the normal function of the protein to play a key role in its autoregulation, and need not require special binding of the protein to its mRNA at a specific site.

3) PROTEIN-MEDIATED TRANSLATIONAL REGULATION

Unlike the number of proteins shown to participate in transcriptional regulation, examples of proteins that function as positive regulators of translation are few in number. One example has been described in bacteriophage Mu. The upstream encoded Com protein is a positive *trans* regulator of Mom translation by specific binding to the *mom* TIR (Wulczyn, et al., 1989). The exact details of this interaction between the mRNA and protein, and the mode of action of the protein in stimulating translation have yet to be worked out. However, the authors suggest that Com binding relieves the structural inhibition of the *mom* TIR.

Endonuclease digestion presents another way in which proteins can mediate both positive and negative translational regulation. The T4 MotA protein mRNA, and a few other T4 induced mRNAs are cleaved within their SD regions (at a GAGG sequence) by a T4 encoded nuclease (Ruckman, et al., 1989; Uzan, et al., 1988), and rendered untranslatable. The mechanism of specific recognition of particular SD regions is yet to be determined. Its ability to discriminate between different SD targets may rely on secondary structure,

sequences between the SD and the start codon, or the tagging of specific SD regions by base modification. This regulatory mechanism exemplifies how a phage can irreversibly shut off a whole class of mRNAs perhaps as part of a temporal regulation of gene expression in phage development.

In contrast to the negative endonuclease action mentioned above, RNase III has been shown to cleave upstream of cistrons in polycistronic mRNAs, resulting in positive translational regulation (Belasco and Higgins, 1988). The positive regulatory effects of RNase III can even be mediated by binding alone (Altuvia, et al., 1987) and need not involve cleavage. In order to establish an integrated picture of translational regulation, more work is needed to define the susceptibility of mRNAs to digestion and the basis for differential mRNA decay.

F) EXAMPLES OF TRANSLATIONAL REPRESSOR PROTEINS

The most commonly described method of translational regulation entails blockage of the RBS by the binding of repressor proteins. Repressor proteins may bind to the TIR using different modes, but the theme of preventing ribosome interaction with the TIR in initiation seems to be prevalent. Below, several examples of translational repressors are discussed in order to emphasize basic principles in translational repression.

1) R17 COAT PROTEIN - PARADIGM OF AN RNA/PROTEIN INTERACTION

One of the most well-studied RNA-protein interactions is the binding of the RNA phage R17 coat protein to the mRNA TIR of the replicase gene. When coat protein reaches a sufficiently high concentration, it binds a defined hairpin loop in the RBS of the replicase, resulting in repression of translation initiation. *In vitro* binding studies have characterized the binding of coat protein to many mRNA operator analogs and resulted in the model shown in Figure 7. The operator hairpin is more resistant to thermal denaturation when the protein is bound, thus implicating the protein also in stabilizing an inhibitory TIR structure. Several single-stranded residues have been shown to be essential for binding to the protein, including 3 of the 4 loop residues and the bulged residue on the 5′ side of the stem. In addition to carefully characterizing the mRNA binding site through the study of many operator mutants, work has been done to elucidate the protein domains responsible for interactions with mRNA operator. Some early results implicated the role of two Cys residues in contacting a U in the mRNA operator (Romaniuk and Uhlenbeck, 1985). However, some recent results show substitution of one of these (Cys46) without loss of repressor function, but several (yet not all) substitutions of Cys101 do affect repressor function (Peabody, 1989). Although the possibility of a transient covalent contact between cys101 and RNA is not categorically

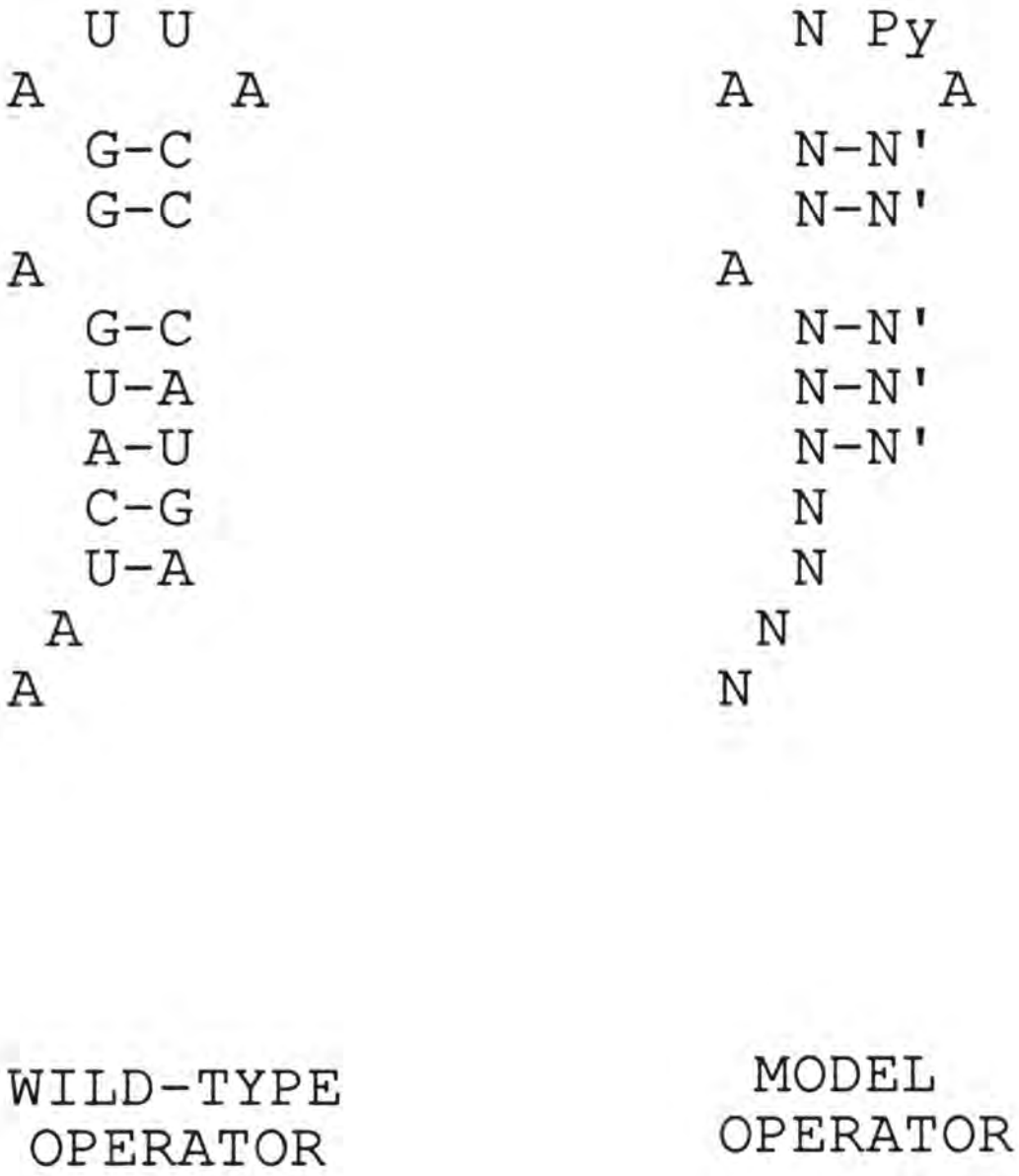


FIGURE 7: The proposed model for the R17 coat protein mRNA operator. N' denotes any base complementary to N. Note the conserved residues in the model operator. Adapted from Uhlenbeck *et al.*, 1987)

ruled out, some mutations demonstrate that cysteines are not absolutely required for translational repression.

2) f1 GENE V ON GENE II mRNA - AN RNA OPERATOR AND A DNA SITE OF ACTION

An example from the filamentous bacteriophage f1 highlights a translational repressor with both DNA and RNA binding capabilities. f1 gene *V* is the single-stranded DNA binding protein , whose activity maintains ssDNA genomes, and thereby enables DNA packaging into virions (Mazur and Zinder, 1975). In addition to its role in binding ssDNA, gpV specifically represses the translation of gene *II* mRNA (Model, et al., 1982; Yen and Webster, 1982). gpII is a site specific nuclease that nicks DNA at the origin of replication and initiates rolling circle replication. gpV specifically binds gene *II* mRNA upstream of the SD region at a 16 nucleotide region required for repression, *in vivo*. While this mRNA operator has been characterized to bind gpV with a 10-fold higher affinity than RNA without this sequence, the role of structure and specific nucleotides in gpV binding has not yet been determined. Interestingly, gpV binds the same operator sequence on DNA with higher affinity (Michel and Zinder, 1989). The sequence of the mRNA operator seems to be unique for RNA binding sites of protein repressors, and may in fact require single-strandedness (Michel and Zinder, 1989). However, translational regulation of replication proteins in phage f1, and the dual DNA/RNA binding activities of gpV are reminiscent of the translational repressors of

phage T4 (especially gp32 ssb, gp43). An important yet unresolved question for these proteins remains : do the DNA and RNA binding capabilities of these proteins reside in a single protein domain?

3) THREONYL-tRNA SYNTHETASE - MOLECULAR MIMICRY BETWEEN A PRIMARY AND SECONDARY LIGAND

Threonyl-tRNA synthetase (ThrRSase) is an intriguing example of a translational repressor that binds an mRNA site which mimics the protein's primary ligand - $tRNA^{Thr}$. The ThrRSase primary function is the charging of tRNAs with the Thr anticodon. This function must require the ability of ThrRSase to distinguish sequence and structural aspects peculiar to the Thr tRNAs. In addition to this primary function, the protein has been shown to repress its own translation, by binding to a mRNA operator that has a similar structure to the anticodon loop of Thr tRNAs (see Figure 8) (Springer, et al., 1986; Springer, et al., 1989). The structure of the wild-type *thrS* operator and several derepressed operator mutants have been confirmed by RNase mapping experiments and the data supports the contention of molecular mimicry and the necessity of this operator for regulation (Springer, et al., 1988). An insertion of 9 bases between this stem-loop hairpin and the SD region causes derepression (Springer, et al., 1986). They have also pursued characterization of the protein residues involved in RNA binding. Interestingly, mutations in several Arg

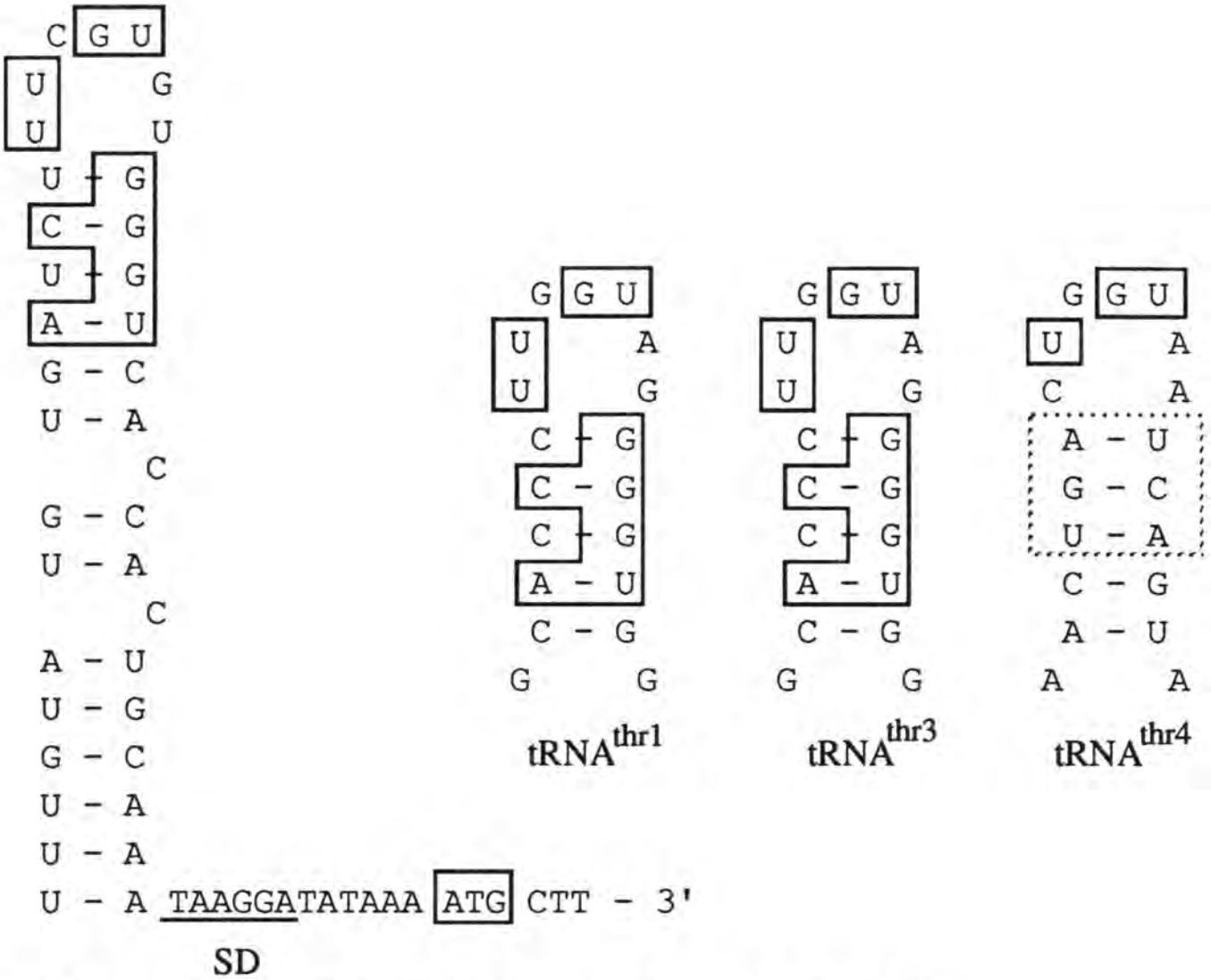


Figure 8: Comparison of *thrS* mRNA operator structure to thr-tRNAs.
Conserved bases are boxed to show the similarity of the mRNA operator to the primary ThrRSase ligand : the anticodon loops of thr-tRNAs. The SD region and initiator codon are highlighted on the mRNA sequence. Adapted from Springer, *et al.* (1986)

residues at the C-terminal of the protein render the protein unable to repress translation of wild-type operator-lacZ fusion mRNAs, *in trans* (Springer, et al., 1988). These same mutants have decreased specific activity for the ThrRSase primary function of aminoacylation, apparently the result of decreased affinity for $tRNA^{Thr}$. This feedback mechanism is reminiscent of the ribosomal protein operons (Nomura, et al., 1984) and appears to apply also to this component of the translational machinery. The TIR structure of several ribosomal protein mRNAs (which are known to be translational repressors) also mimics structure within the primary ligand of rRNA. Overall this dual utilization of binding specificity not only streamlines protein domains, but also provides a concrete connection between the physiological function and the mode of regulation of proteins. This concept may explain why several components of the translational machinery are regulated at the translational level.

4) OTHER MECHANISMS OF TIR BINDING AND REPRESSION

All of the examples I have described so far (with the possible exception of f1 gene *II* mRNA) involve mRNA binding sites that overlap the SD region; in these situations, protein obstruction of the RBS is the obvious mode of action. In contrast, the T4 regA protein binding site on mRNA does not always overlap the SD region (Liang, et al., 1988; Unnithan, et al., 1990; Winter, et al., 1987). In addition to repressing its own synthesis RegA represses the

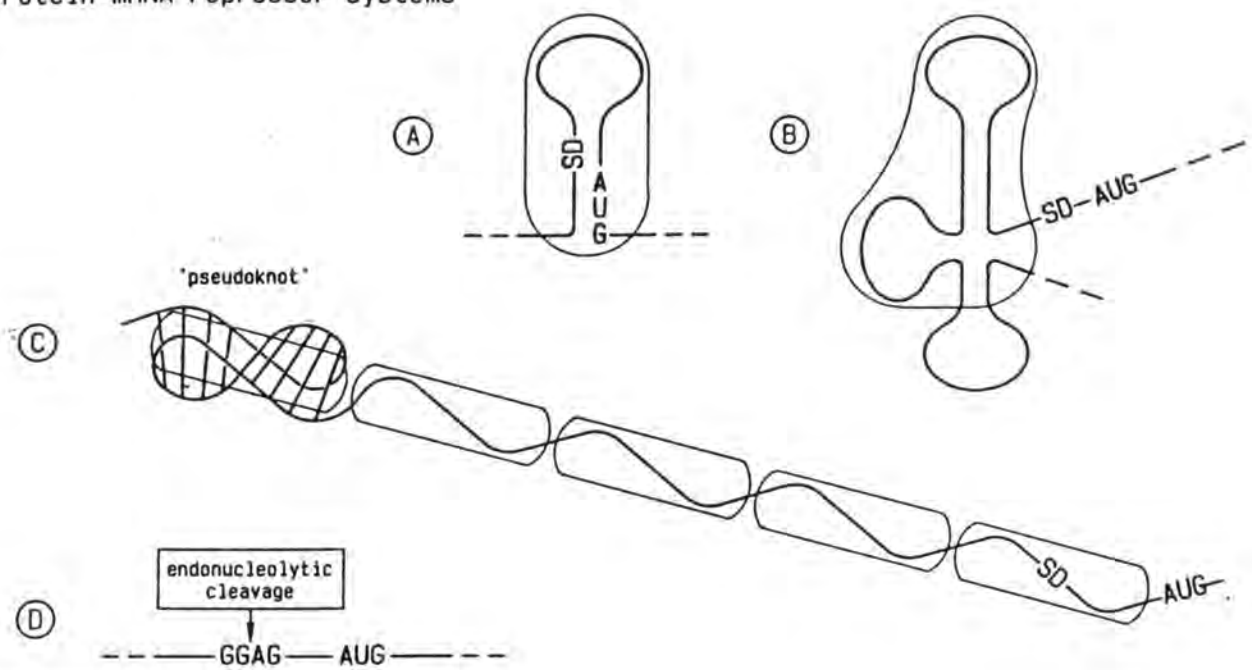
translation of several T4 encoded proteins, most involved in DNA replication. The binding site for RegA seems to be unstructured and encompasses residues surrounding the initiator AUG (-5 to +12 on one particular RBS). The initiator AUG has been shown to be necessary but not sufficient for specific binding of RegA to RNA *in vitro* (Unnithan, et al., 1990). Binding of RegA to the *rIIB* RBS blocks the formation of 30S initiation complexes (Winter, et al., 1987). Therefore, while not directly obstructing access to the SD region, RegA can inhibit initiation by preventing interaction with the initiator tRNA in initiation complex formation. The details of the other operator sequence requirements for specific binding are yet to be clarified.

Another T4 translational repressor gp32, the phage encoded single-stranded DNA binding protein, presents an additional mode of operator binding. This protein represses its own translation by binding initially upstream of the RBS; the protein binding first nucleates at a pseudoknot structure and then as the free concentration of protein increases binds cooperatively down the mRNA to the RBS. Bound gp32 can block the binding of 30S subunits (McPheeters, et al., 1988). Figure 9 schematically portrays these different modes of repressor binding, and summarizes the modes of translational regulation presented in this introduction. The interaction of gp32 with an RNA pseudoknot was one of the first examples of protein-pseudoknot binding described.

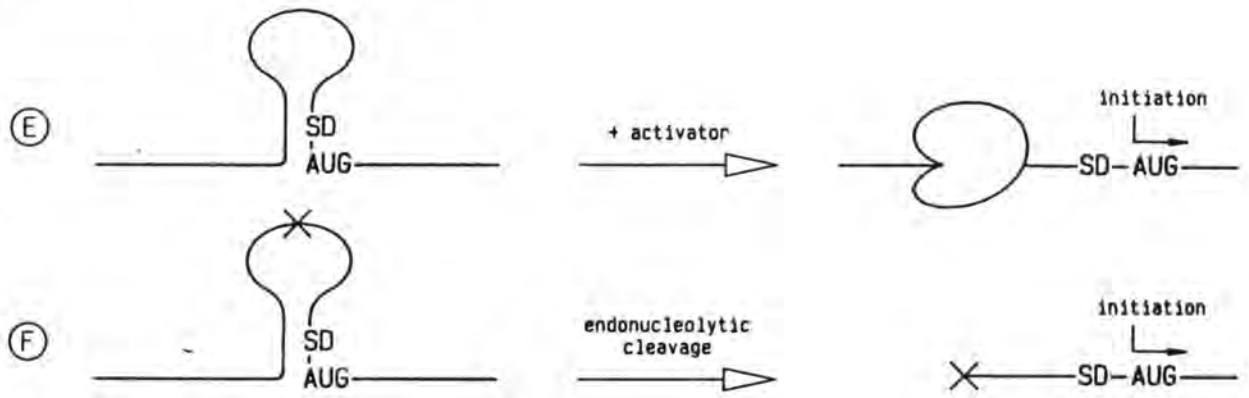
Figure 9: Modes of translational regulation

Panels A through D portray different models for protein repressor systems. The remaining panels summarize other mechanisms of post-transcriptional regulation. In each case, the SD region and initiator codon are shown schematically to illustrate the principle of regulation. Taken from McCarthy and Gualerzi, (1990).

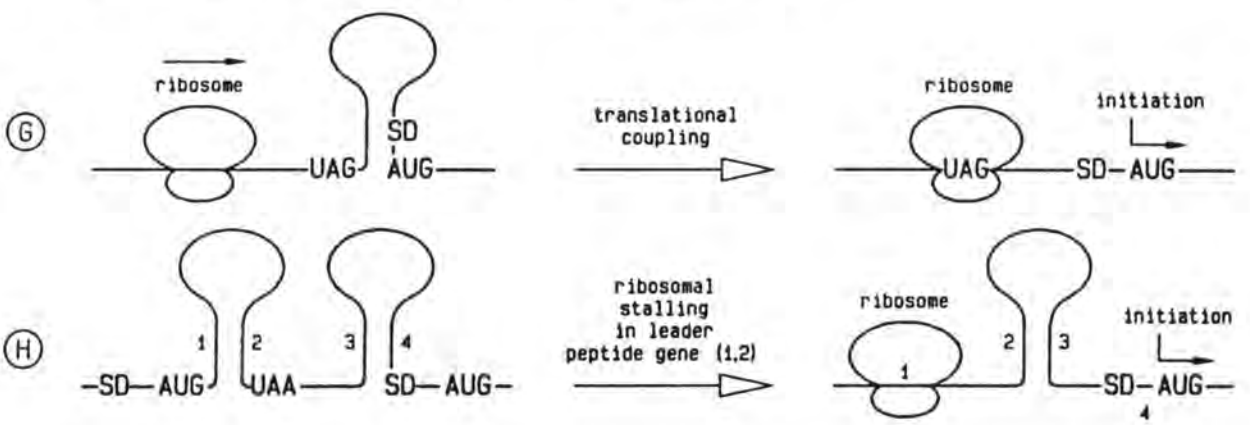
Protein-mRNA repressor systems



Protein-mRNA activation systems



Translational apparatus-mediated control / regulation



Transcriptional apparatus-mediated control / regulation

Some other examples are ribosomal protein S4 (binding to α operon mRNA)(Tang and Draper, 1989) and tRNA ligases (binding tRNAs containing pseudoknots) and RNaseP (binding viral tRNA-like RNAs)(Mans, et al., 1990). What elements of these RNA higher order structures are essential for protein recognition? Are there other tertiary RNA structures that are specifically recognized by proteins? These examples have shown the importance of considering the full three-dimensional context of RNA molecules that proteins inspect upon specific binding and recognition. The challenge remains to answer these and other questions that relate to RNA-protein interactions.

G) RECOGNITION OF RNA BY PROTEINS

The importance of understanding RNA-protein interactions extends beyond a detailed understanding of translation and its regulation. Indeed RNA-protein interactions are central to many cellular processes (see Table 1).

There is a great need to explore the basic rules that govern RNA-protein binding. With the increased recognition of the catalytic role of RNA, there is a need to define how RNA and proteins interact together to perform cellular functions; the ribosome is only one example. Translational repressors represent good model systems to begin to dissect those rules for RNA-protein binding, association, and cooperation in function. T4 DNA polymerase is a well characterized protein with several activities, making it suitable for addressing

these questions of structure and function in RNA-protein interactions.

Table 1: RNA-protein interactions important to cell/virus physiology

Examples	Function/Role
Translational Repressors	control of gene expression
Amino acid-tRNA ligases	amino acid-tRNA activation
Ribosomal proteins	ribosome biogenesis, regulation of cell growth
RNA replicases and reverse transcriptases	RNA replication, genome duplication of viruses
RNaseP, RNase III,	RNA processing
Spliceosomes, snRNPs, and scRNPs	preparation/processing of mRNA for translation, & splicing of introns
RNA modifying enzymes	capping, polyadenylation, and methylation
Ribonucleases for mRNA decay	mRNA stability/turnover
RNA-helix unwinding proteins	yet to be determined
Transcription termination and anti-termination proteins	regulation of gene expression transcriptional control
Retrovirus NC proteins HIV Rev and Tat proteins	RNA localization, genome packaging, control of gene expression

H) SUMMARY AND REMARKS

The regulation of translation in prokaryotes occurs by a variety of mechanisms, most of which culminate in the blockage of translation initiation. T4 has been the object of studies central to the elucidation of some of these mechanisms. In his review of translational regulation, Gualerzi made the following remark: "A major challenge for future research in this area will be to achieve a full understanding of the RNA structures (bearing in mind the possible significance of previously unrecognized forms) and the RNA-RNA/RNA-protein interactions responsible for the phenomena {of translational regulation}." It is in the spirit of that remark that I present the studies of this dissertation.

II) MATERIALS AND METHODS

A) MEDIA FOR BACTERIAL AND PHAGE GROWTH

The solutions and media used for growth of bacterial strains and for maintenance of phage stocks are listed below. Unless otherwise indicated, media components were dissolved in water, adjusted to the proper pH and sterilized by autoclaving at 121°C for 20-30 min.

1) LIQUID MEDIA

Nutrient broth

8.0	g	Bacto-Nutrient broth
5.0	g	NaCl
1.0	mg	thiamine

Dissolved to a final volume of 1000 ml and adjusted to pH 7.5 with 1N NaOH.

Luria broth (LB)

10.0	g	Bacto-Tryptone
5.0	g	Bacto-Yeast Extract
10.0	g	NaCl

Dissolved to a final volume of 1000 ml and adjusted to pH 7.5 with 1N NaOH.

M9-based media

Different M9 media vary with the specific components that are combined:

10 x salts

60.0	g	Na_2HPO_4
30.0	g	KH_2PO_4
5.0	g	NaCl
10.0	g	NH_4Cl

Dissolved to a final volume of 1000 ml.

Solution A

300	ml	10 x salts
3	ml	1 M $MgCl_2$
3	ml	0.1 M $FeCl_3$ (in 0.1 M HCl)
2200	ml	H_2O

Dispensed into 125 ml aliquots in bottles, autoclaved, and allowed to cool before the addition of carbon sources.

Solution B

12.0	g	glucose
60.0	mg	L-tryptophan
3.0	ml	thiamine (1 mg/ml stock)
200.0	ml	H_2O

1.1 M9

125	ml	solution A
10	ml	solution B

Amino acid mixtures for M9-based media

a) 2.5% Casamino acids--vitamin-free casamino acids (Difco) were dissolved to yield a final concentration of 2.5 g/100 ml.

b) 18 amino acid mix --solution containing a mixture of 18 amino acids (minus methionine and tyrosine) at 5 mM each.

c) Tyrosine solution at 2.5 mM.

Variations of M9-based media used in these studies were:

M9 medium

15 ml of sterile water and 0.15 ml of 1M $MgSO_4$ added to 135 ml of 1.1 M9.

M9 sulfate-free medium

15 ml of sterile water added to 135 ml of 1.1 M9.

M9S medium

15 ml of 2.5% vitamin-free casamino acids and 0.15 ml of 1M $MgSO_4$ added to 135 ml of 1.1 M9.

M9-19 amino acid medium

15 ml of 18 amino acid mix, 5 ml of 2.5 mM tyrosine, and 0.15 ml of 1M $MgSO_4$ added to 135 ml of 1.1 M9.

Phage diluent

Tris-Cl, pH 7.5 ... 10 mM

$MgCl_2$ ..... 1 mM

NaCl ..... 86 mM

Gelatin ..... 0.001% (w/v)

L-Tryptophan ..... 0.1 mM

2) MEDIA CONTAINING AGAR

The dry components listed in the recipes below were dissolved in hot (90-100°C) distilled water. The solutions were then sterilized by autoclaving and stored in congealed form at room temperature until needed. Congealed media were melted by heating in a microwave oven and kept in a 60°C water bath until needed. When pouring plates, media were cooled to 55°C before thermolabile substances (e.g. antibiotics) were added. Plates were then poured, allotting approximately 30-35 ml of agar per 85 mm petri dish and allowed to harden at room temperature before storage at 4°C in sealed plastic bags. When needed, plates were pre-warmed for 30-60 min in a 30°C incubator with tops ajar to allow condensates and surface moisture to evaporate.

EHA bottom agar (Enriched Hershey bottom agar)

10.0	g	Bacto-Agar
13.0	g	Bacto-Tryptone
8.0	g	NaCl
1.3	g	Dextrose (D-glucose)
2.0	g	Trisodium Citrate

Dissolved to a final volume of 1000 ml and autoclaved.

EHA top agar (Enriched Hershey top agar)

6.0	g	Bacto-Agar
13.0	g	Bacto-Tryptone
8.0	g	NaCl
2.0	g	Trisodium citrate
3.0	g	Dextrose (D-glucose)

Dissolved to a final volume of 1000 ml and autoclaved.

LB bottom agar

10.0	g	Bacto-Tryptone
5.0	g	Bacto-Yeast Extract
10.0	g	NaCl
15.0	g	Bacto-Agar

Dissolved to a final volume of 1000 ml and autoclaved.

B) BACTERIAL STRAINS

The *Escherichia coli* (*E. coli*) strains used, their sources, and characteristics are listed in Table 2.

Table 2. *Escherichia coli* strains.

Strain	Description	Source	Reference
K-12 Strains			
CR63	*supD60* λ^r *am* (UAG) suppressor: (serine-insertion)	R. S. Edgar	Appleyard *et al.*, 1956
CAJ70	*supU* UGA suppressor: (tryptophan-insertion)	S. Brenner	Shultz *et al.*, 1982
K802	*supE* UAG supressor: (glutamine-insertion) *hsdR hsdM*$^+$ *gal met*	N. Murray	Wood, 1966
MRE600	RNase *I*$^-$	E.T. Young	Cammack, 1965
BL21(DE3)	strain harbors a chromosomal-ly integrated phage T7 gene *1* (RNA polymerase) under the control of an inducible *E. coli lacUV5* promoter. - F$^-$ *hsdS gal* r_B^- m_B^-, *sup*^{o}	F.W. Studier	Studier, 1986
B Strains			
B^E	wild-type, *sup*^{o}	L. Gold	Gorini, 1970; Karam & O'Donnell, 1973

$B^E str^r$	stryptomycin-resistant derivative of B^E (*sup*^{o})	this lab.	Karam & O'Donnell, 1973
NapIV	*hsdR*$_k^+$, *hsdM*$_k^+$, *hsdS*$_k^+$, *thi* (*sup*^{o})	L. Gold	Nelson *et al.*, 1982
S/6*str*^{r}	T6^r, Str^r derivative of *E. coli* S/6 (*sup*^{o})	this lab.	Doermann & Hall, 1953

C) BACTERIAL PLASMIDS

Table 3 lists the bacterial plasmid vectors used in the cloning and expression studies. Other vectors that were constructed in the process of this thesis work are diagrammed in RESULTS.

Table 3. Plasmid vectors

Vector	Description	Source	Reference
pBR322	a very famous plasmid	H. Boyer	Bolivar, *et al.*, 1977
pUC8/pUC9	Multicloning site plasmids derived from pBR322	Bethesda Research Laboratories	Vieira & Messing, 1982
pSP72/pSP73	Multicloning site plasmids derived from pUC series with phage T7 and SP6 promoters	Promega Biotec	(Krieg and Melton, 1987) pEM104 plasmid derivative of pGW7 containing T4

D) TRANSFORMATION OF CELLS WITH PLASMIDS

Cells were grown with aeration in Luria Broth to about 3 x 10^8 cells/ml. 6 ml of cells (enough for 12-15 transformations) were pelleted at 3000x g for 5 min at 4°C. The pellet was resuspended in 3 ml of chilled MOPS 7/RbCl buffer (10 mM 3-[N-morpholino]propanesulfonic acid, pH 7.0 and 10 mM RbCl) and centrifuged immediately at 3000 x g for 5 min. The pellet was resuspended in 2.7 ml of chilled MOPS 6.5/RbCl buffer (100 mM 3-[N-morpholino]propanesulfonic acid, pH 6.5 and 10 mM RbCl) before adding 0.3 ml of chilled 500 mM $CaCl_2$. The cell suspension was mixed gently and incubated on ice for 30 min. The cells were then pelleted at 3000 x g for 5 min, gently resuspended in 1.5 ml chilled MOPS 6.5/RbCl buffer containing 50 mM $CaCl_2$ and returned to ice. Immediately before use, 50 µl of dimethyl sulfoxide (DMSO) was added and the cells were dispensed in 0.1 ml aliquots into sterile thin-walled 13x100 mm glass tubes and 50-100 ng of DNA was added. The mixture was then incubated on ice for 35 min before being subjected to a heat-shock by swirling in a 44°C water bath for 30 sec. The tubes were then returned to ice and the treated cells were subsequently plated on LB plates containing the appropriate antibiotic. Antibiotic-resistant colonies were picked for analysis after 15-20 hr incubation at 30°C.

E) SCREENING BACTERIAL CLONES BY COLONY HYBRIDIZATION USING OLIGONUCLEOTIDE PROBES

This procedure is a minor modification of the procedure of Paddock (Biotechniques 5(1):13-16, 1987), done by picking several copies of a master plate, rather than replica plating.

1) SOLUTIONS

10X Kinasing buffer

500 mM Tris-Cl, pH 7.5

100 mM $MgCl_2$

6X SSC

0.9M NaCl

90mM sodium citrate

20X NET

3M NaCl

20mM EDTA, pH 8.0

0.3M Tris-Cl, pH 8.0

Filter treatment solutions

0.5M NaOH (from 5M NaOH stock)

0.5M Tris-Cl, pH 7.4 (from 2M Tris-Cl, pH 7.4 stock)

2) REPLICATING COLONIES

Colonies were picked with sterile toothpicks and transferred onto LB plates containing the appropriate antibiotic and incubated at 30°C overnight. A UV-sterilized Whatman 541 filter-paper disk (8 cm in diameter) was carefully placed over each plate, flattened with a spreader and incubated at 30°C for 4 hr or overnight. The filter was carefully lifted, notched to code the source and orientation of the filter on the plate.

3) LYSING CELLS GROWN ON THE FILTER DISC

Each filter was placed (colonies-side up) in 20 ml of 0.5M NaOH and mixed intermittently at room temperature for 5 min. This was repeated by transferring to a fresh 20 ml of NaOH and mixing for 5 min. Several filters can be passed through the same two alkali baths.

The filter was then neutralized and washed by transferring to 20 ml of 0.5 M Tris-Cl, pH 7.4 (after removing the first fluid by blotting against the edge of the first dish), and mixed at room temperature for 5 min. This wash was also repeated. Each filter was subsequently incubated in two washes with 20 ml of 2X SSC for 5 min each, rinsed with 95% ethanol, and dried on filter paper.

4) RADIO-LABELING NUCLEIC ACID PROBES

The synthetic oligonucleotide probe (a minimum length of 15 bases) was labeled using γ^{32}P-ATP and T4 polynucleotide kinase as follows:

oligonucleotide probe	1 µl (~100ng - OD=3.0)
γ^{32}P-ATP	14 µl (140µCi)
10X kinasing buffer	2 µl
100mM DTT	2 µl
Kinase (10units/µl)	1 µl
Total Volume	20 µl

The kinase reaction was incubated at 37°C for 1 hr. It was usually not necessary to purify the probe from unincorporated nucleotide by passing through a Sephadex G25 spun column although this was done when background on the filters became a problem. TCA precipitable counts were determined and typically found to be in the range of 5-10x10^6 cpm/µl.

Generally, between 5 and 10 µl of the kinase reaction (25-50 ng of probe) was added to 100 µl of sonicated herring sperm DNA (at 10 mg/ml) and 0.9 ml of 6xSSC. This was sufficient probe for up to 10 filters per single hybridization experiment (usually greater than 5 x 10^6 cpms/filter); as little as 4 ng probe/filter is adequate for a good signal. Immediately before use, the probe mix was heated at 90-95°C for 3 min and then quickly cooled on ice.

5) NUCLEIC ACID HYBRIDIZATION

Dried filters were rehydrated in 6X SSC, briefly blotted on Whatman 3MM paper and placed into either a plastic bag with 5-10ml (volume depending on the number of filters) 6X NET, or more conveniently in screw-cap glass jars (8.5 cm diameter) with 20 ml of 6xNET. The probe mixture was added directly to the fluid in the bag or jar, which was then well sealed. Hybridization proceeded for 16-24 hr with constant shaking in a water bath. Although the hybridization temperature (T_H) can be calculated according to the formula below, successful hybridizations were usually obtained by using a temperature of 42°C for short probes (<20 nucleotides) and a maximum of 52°C for longer probes.

$T_D = 2°C(A+T)^* + 4°C(G+C)^*$ where * = the number of respective bases in the probe

$T_H = T_D - 5°C$

The next day, filters were removed and washed at room temperature four times with 200ml-500ml of 6X NET for 10 min. They were subsequently air-dried, packaged between mylar sheets (9/1000 inch thick) for autoradiography. When an enhancing screen was used, exposure times of only a few hours were sufficient to yield good signals.

F) PHAGE STRAINS

The phage T4 strains used in this work are described in the Table 4.. The final, purified phage stocks were suspended in a buffer containing 10 mM $MgCl_2$, 1% gelatin, and 10 mM Tris-Cl, pH 7.5.

Table 4. Phage T4 strains

Strain	Description	Source	Reference
44*am*N82	Gene 44 UAG mutant	W. Wood	Epstein *et al.*, 1963
43tsYN13	Gene *43* ts mutant	this lab	Hughes, 1987 & Hsu and Karam, 1990
43amE4301-E4322	Gene *43* double mutant	this lab	O'Donnell, 1971
43tsYN15	Gene *43* ts mutant	this lab	Hughes, 1987 & Hsu and Karam, 1990

G) PREPARATION OF PHAGE LYSATES

1) PREPARATION OF M9S LYSATES

E. coli CR63 was used as host in the preparation of phage lysates of *am* (UAG) mutants and CAJ70 was used for *ts* (temperature-sensitive) and UGA mutants. *E. coli* Nap IV were also used for preparation of M9S lysates for screening

operator recombinant phage. All phage stocks were derived from single plaques. A plaque was lifted using a sterile Pasteur pipette and phage particles were suspended in 5 ml of Nutrient Broth by vortexing. A few drops of chloroform were added to sterilize against bacterial contamination. Approximately 10 µl of phage suspension (10^4-10^5 phage total) and 0.3 ml of dense cell culture (containing ~$5x10^8$ cells) were added to 5 ml M9S medium and the infection mixture was aerated at 30°C overnight. About 0.3 ml of chloroform was added to complete the lysis and the crude lysate was centrifuged for 10 min at 10,000x g to pellet cell debris. The supernatant was then transferred to a sterile 5 ml screw-cap polypropylene tube and stored at 4°C. Phage titers ranged between $5x10^{10}$ and $1x10^{11}$ plaque-forming-units(PFU)/ml.

2) PREPARATION OF HIGH-TITER LYSATES

An *E. coli* culture (25 ml in M9S) was grown to $5x10^8$ cells/ml, infected with phage at a multiplicity of infection (m.o.i.) of 5 and aerated at 30°C for 30 min. The infected cells were then pelleted at 5000 x g for 5 min and resuspended in 5 ml of M9S. After overnight aeration at 30°C, pancreatic DNase I and RNase A were added to give a final concentration of 5 µg/ml each and the lysing culture was incubated at 30°C for 15 min before 0.3 ml of chloroform was added to complete the lysis. Cell debris was removed by centrifugation at 10,000x g for 10 min. The supernatant containing the phage was made 0.5 M in NaCl and 10% (w/v) in

polyethylene glycol 6000 (PEG6000), and the suspension incubated at 4°C overnight to precipitate the phage particles. The resulting visible precipitant was collected by centrifugation at 3000x g for 5 min and resuspended in 2 ml of sulfate-free M9 containing 0.001% gelatin. PEG was removed by gently extracting the phage suspension with 0.3 ml of chloroform followed by centrifugation at 3000x g for 5 min. The aqueous phase was then transferred to a sterile screw-cap tube and stored at 4°C. Phage titers ranged between $8x10^{11}$ and $2x10^{12}$ PFU/ml.

H) GENETIC METHODS FOR ANALYSIS OF PHAGE MUTANTS

1) MARKER-RESCUE ASSAYS

T4 mutants bearing conditional lethal mutations in the regions of interest were used to infect non-permissive *E. coli* cells carrying recombinant plasmids. Plaque formation by the phage mutants indicated the acquisition of wild-type sequence from the plasmid-borne T4 DNA via homologous recombination. The same principle was adopted to perform targeted mutagenesis in which mutations introduced to cloned T4 DNA segments could be transferred to T4 genome by homologous recombination events (see below).

In a typical assay, recombinant plasmid-bearing cells were grown at 30°C in nutrient broth containing the appropriate antibiotic to approximately $3x10^8$ cells/ml. The cells were then washed free of antibiotic and diluted to the

proper density by pelleting 0.5 ml of cell culture and resuspending in 1.5 ml of antibiotic-free medium. Infecting phages used in the assay were diluted in nutrient broth to approximately $1x10^9$ PFU/ml. One drop each of the cell and the phage dilutions were mixed in the wells of a microtiter plate and the mixture was incubated at 30°C for 30 min before being spotted on a lawn of *E. coli* restrictive to the infecting phages. Incubation was carried out overnight at the appropriate temperatures (30°C for nonsense mutants and 42°C for *ts* mutants). Production of plaques under the restrictive growth condition was indicative of recombination between infecting phage and cloned T4 DNA.

2) INTRODUCING MUTATIONS BY PLASMID-PHAGE RECOMBINATION

The T4 gene *43* mutation *43tsYN13* (Glu->Lys at codon 26) is closely linked to the gp43 translational operator region. This mutant phage was used to infect *E. coli* hosts harboring recombinant plasmids which contained the mutant operators. Phage lysates were then plated on plasmid-free *E. coli* CAJ70 at 42°C to recover temperature-resistant recombinants, *i.e.* rescue of the wild-type allele for *43tsYN13* from the plasmids. The recombinants were subsequently assayed for co-rescue of mutant operator sequences by two methods (Andrake, *et al.* 1988): measurement of gp43 synthesis after phage infection, and sequencing of gene *43* mRNA isolated from infections with the recombinants.

3) COMPLEMENTATION TESTS

Complementation tests were used to test the *in vivo* functionality of gene *43* plasmid clones. Double-amber gene *43* phage mutants were used to infect non-permissive hosts carrying the gene *43* plasmid clone. Expression of gp43 from the plasmid could complement the gene *43* defect of the infecting phage mutant when the plasmid clone encoded a functional wild-type protein.

Briefly, 0.2 ml of dense cells (containing the gene *43* plasmid clone to be tested) were mixed with about 50 µl of a phage mutant (double-amber E4301/E4322 - diluted to 1000 phage/ml) and added to 2.5 ml of EHA top agar, equilibrated in a 37°C thermolyne dry bath. The mixture was immediately vortexed and poured over a pre-warmed EHA plate. The top agar was allowed to solidify before transferring to a 30°C incubator for plaque formation overnight. Plaques were counted and plating efficiency determined relative to plating on an amber supressor strain CR63/5.

4) DNA SYNTHESIS ASSAYS

E. coli NapIV cells were grown in M9S at 30°C to a density of 3 x 10^8 cells/ml. 1 ml of cells was transferred to 0.5 ml of phage at 6 x 10^9 phage/ml (m.o.i. = 10) and aerated at 30°C for 10 min. At 10 min post-infection, 0.2 ml of infection mix was transferred to 1.8 ml of M9S containing 20 µCi ^{3}H-Thymidine/ml and 1 µg/ml thymidine (thus, 20 µCi ^{3}H/µg of unlabeled nucleoside). At 20 min post-infection,

and at 10 min intervals thereafter, 100 µl of labeled infection mix was removed and added immediately to 5 ml of chilled TCA (5%) with a drop of albumin as carrier for nucleic acid precipitation. After 45 min incubation, TCA precipitates were collected by filtration and the amount of precipitable ^{3}H determined by scintillation counting.

5) BURST SIZE MEASUREMENTS

E. coli NapIV cells were grown in M9S at 30°C to a density of 3 x 10^{8} cells/ml. 1 ml of cells was transferred to 0.5 ml of phage at 6 x 10^{9} phage/ml (m.o.i. = 10) and aerated at 30°C for 10 min. At 10 min post infection, 1.0 ml of infection mix was transferred to 5 ml of pre-chilled M9S in a 15 ml Corex tube, in order to dilute unabsorbed phage. The diluted cells were pelleted at 12,000 x g for 5 min (SS34) and the supernatant thoroughly aspirated. The cells were then resuspended in 1 ml of M9S and aerated for 5 min, after which a sample was removed, diluted, and plated for infective centers, *i.e.* the number of viable bacteria infected). The remainder of the infection mix was aerated at 30°C for 2-3 hrs, then lysed by the addition of 0.2 ml of chloroform. Dilutions were made for plating on *E. coli* CAJ70 to determine the total phage produced. The burst size was then calculated as the total number of phage divided by the number of infective centers.

I) MEASUREMENT OF GENE EXPRESSION IN PHAGE-INFECTED OR PLASMID-BEARING CELLS

1) RADIOACTIVE LABELING OF PROTEINS

Protein labeling in phage-infected cells

E. coli NapIV cells were grown at 30°C in M9 to a concentration of $3x10^8$ cells/ml, diluted to $2x10^8$ in the same medium, and infected with phage at m.o.i. = 10. The infection mixture was aerated at appropriate temperatures, and at specific time intervals, 1 ml aliquots of infected cells were transferred to 0.1 ml of pre-warmed M9 containing 5 µCi of ^{35}S-methionine. The mixture was aerated for 5 min before protein labeling was terminated by chilling on ice.

Protein labeling in plasmid-bearing cells

E. coli NapIV cells bearing the λ*c*I857pL-directed expression plasmid (e.g., pGW7) were grown in M9-19 to $2x10^8$ cells/ml at 30°C in the presence of 20 µg/ml of ampicillin. The plasmid-directed gene expression was induced by aerating 0.5 ml cell culture at 42°C for at least 10 min. 50 µl M9-19 containing 5 µCi of ^{35}S-methionine were added to the heat-induced cell culture and the culture was labeled with aeration at 42°C for at least 15 min. The labeling was terminated by chilling the culture on ice. Subsequent extraction of labeled proteins and measurement of radioactivity in cell extracts were as described (O'Farrell and Gold, 1973; Karam and Bowles, 1974).

2) Analysis of phage-specific proteins by SDS-polyacrylamide gel electrophoresis

The SDS gel electrophoresis procedures used to analyze phage proteins utilized the buffer system described by Laemmli (1970) and have been detailed previously (Karam and Bowles, 1974) except that the Bio-Rad slab electrophoresis cell (PROTEAN II) was used. Autoradiography and densitometry of protein bands on X-ray films were as described (Karam and Bowles, 1974) except that Kodak X-OMat AR X-ray films were used in autoradiography. Quantitation of protein bands was done on one of several densitometers: Joyce-Loebl Microdensitometer (model 3CS), Bio-Rad Video Densitometer model VD620, or Hoeffer model GS300 densitometer.

J) PLASMID DNA PREPARATION

1) Alkaline miniscreen

Reagents

Solution A

Lysozyme	2	mg/ml
Glucose	50	mM
Na_2EDTA	10	mM
Tris-Cl, pH 8.0	25	mM

Solution B

NaOH	0.15	M
SDS	1.0	%

Solution C

KOAc, pH 4.8	3.0	M

5x TE buffer

Tris-Cl, pH 8.0	50	mM
Na_2EDTA	5	mM

RNase A solution

10 mg/ml of bovine pancreatic RNase A (from Boehringer-Mannheim Biochemicals) in 10 mM KOAc, pH 4.8. The solution was heat-treated in boiling water bath and was allowed to renature overnight at 4°C. The heat-treatment inactivates residual activity of any DNase contaminants.

Procedure

All centrifugation steps were carried out in Eppendorf centrifuges at 12,000x g. Plasmid-bearing cells were grown dense (overnight) in LB containing antibiotic and 1.5 ml of the dense culture were harvested by centrifugation. The cell pellet was resuspended in 100 µl of Solution A followed by the addition of 10 µl of RNase A solution. After a 30 min incubation on ice, 200 µl of Solution B was added and the suspension was mixed gently. After 5 min on ice, 150 µl of Solution C was added, and the suspension was mixed gently by tube inversion, and incubated on ice for an additional 15 min. The lysed mixture was centrifuged for 15 min in the

cold and the supernatant was transferred to a chilled Eppendorf tube, mixed with 1 ml of cold 95% ethanol, and incubated at -20° C for 30 min. The precipitated nucleic acids were collected by centrifugation in the cold for 15 min. The pellet was dissolved in 100 µl of 5x TE buffer and re-precipitated by the addition of 1/10 volume of 2.5 M LiOAc/0.5 M $MgOAc_2$ and 3 volumes of cold 95% ethanol. The mixture was incubated at -20°C for at least 30 min. The nucleic acids were pelleted by centrifugation for 15 min in the cold, dried under vacuum and resuspended in 200 µl of 1x TE buffer. The crude DNA solution was then extracted with 1x TE-equilibrated phenol followed by chloroform extraction and the nucleic acids were re-precipitated by ethanol as described above. The precipitate was re-dissolved in 50 µl 1x TE and a small sample (5-10 µl) was analyzed by agarose gel electrophoresis.

2) MEDIUM SCALE DNA PREPARATION

A 50 ml culture of plasmid-bearing *E. coli* cells was grown overnight (2x25ml) to high density (>10^9 c/ml) with aeration in LB media containing the appropriate antibiotic. Cells were harvested in 30 ml Oakridge tubes at 5900 x g for 7 min (7,000 rpm in the SS34 rotor). Each pellet was washed with 12 ml of TES buffer(10 mM Tris-Cl, pH 8.0 / 10 mM NaCl / 1 mM Na_2EDTA) and the resuspended cells were pooled and re-pelleted. The washed cell pellet was suspended in 2 ml of Solution A and vigorously mixed with 50 µl of heat-treated

RNase A solution. After a 30 min incubation on ice, 4 ml of Solution B was added and incubated for 5 min to clear the solution and reduce viscosity by breaking bacterial DNA. After mixing well with 3 ml of Solution C, and incubation for 15 min on ice, denatured proteins and cell debris were pelleted at 12,000 x g (10,000 rpm in Sorvall SS34). The supernatant was drawn off with a sterile plastic pipet to note the total volume. The DNA was precipitated by the addition of 0.6 volumes of 2-propanol, mixed well and incubated at room temperature for 15 min. Then the DNA was pelleted for 30 min at 15,000 x g in a swinging-bucket rotor (HB4 at 10,000 rpm). The supernatant was aspirated, the DNA pellet dried briefly under vacuum, and resuspended in 1 ml of 1X TE buffer. The dissolved DNA was extracted twice with an equal volume of phenol (equilibrated with 1X TE buffer), followed by a single extraction with chloroform to remove trace amounts of phenol.

The DNA was subsequently precipitated by the addition of 1/10 volume of 2.5 M $LiAc_2$/0.5M $MgAc_2$ and 3 volumes of cold absolute ethanol. The DNA was again pelleted at 15,000 x g, dried and resuspended in 920 µl of 1X TE buffer. CsCl (1g) was dissolved in the DNA solution and 80 µl of ethidium bromide (10 mg/ml) was added for detection of DNA bands after centrifugation in the gradient. The suspension in CsCl was transferred to heat-seal tubes(11x32 mm) and centrifuged in a Beckman table-top centrifuge (TL100 in the verticle rotor TLV100) either at 80,000 rpm overnight(17 hrs), or at 100,000

rpm for 4 hrs. The plasmid DNA band was visualized by UV (302 nm) illumination and carefully collected by needle puncture of the tube. The collected band was extracted 4 times with CsCl saturated 2-propanol to clear ethidium bromide, and the aqueous phase was then loaded directly to a gel filtration column, as described below, to remove salt and oligonucleotides.

3) A5M COLUMN CHROMATOGRAPHY

The method for separation of plasmid DNA from contaminating oligonucleotides and excessive salt used a simplified gel filtration chromatography procedure with the A5M matrix(BIO-RAD agarose-based Bio-Gel A-5m, 200-400 mesh), which has an operating range of 10,000-5,000,000 MW. Columns were gravity packed in a 10 ml disposable column (Biorad cat# 731-1550) and equilibrated in 50 mM Tris-acetate, pH 8.0 / 5 mM Na_2EDTA. The CsCl solution was applied to the drained A5m bed and eluted with the same 5X TE buffer; 1 ml fractions were collected and checked for OD at 260 nm. The plasmid DNA eluted in the unexcluded fractions (in about 2 ml). Fractions containing plasmid DNA were chilled on ice (from collection onward), pooled, extracted, and precipitated overnite according to normal procedures. The dried pellet was subsequently dissolved in 1X TE buffer at a concentration of 0.5-1 mg nucleic acid/ml.

K) DEPHOSPHORYLATION OF DNA 5'-ENDS WITH ALKALINE PHOSPHATASE

The standard procedure used was optimized for 10-100 µg of linearized DNA. The DNA was ethanol-precipitated, pelleted by centrifugation, dried under vacuum, and resuspended in 20 µl of 0.1 M Tris-Cl, pH 8.0 / 5 mM $MgCl_2$, to which was added 20 µl of 50 mM NaCl, 20 µl of 1 mM $MgSO_4$ / 1 mM $ZnSO_4$, and 140 µl of 0.1 M Tris-Cl, pH 8.0.

A 10 µl aliquot of the mixture was transferred to another tube to be used as a control. Both tubes were pre-warmed to 60°C. The phosphatase reaction was carried out at 60°C by adding calf intestinal alkaline phosphatase (Boehringer Mannheim Biochemicals cat# 713-023) to the pre-warmed DNA solution in 3 x 1 µl aliquots at 15 min intervals, *i.e.*, the total incubation time with enzyme was 45 min. The reaction was terminated by phenol extraction. The phenol extraction was repeated 3 additional times to remove all traces of alkaline phosphatase activity. The phenol was removed from the aqueous phase by chloroform extraction followed by ethanol precipitation. The efficiency of dephosphorylation was determined by the extent of recircularization of the linearized DNA in the presence of DNA ligase. The extent of recircularization was measured by the efficiency of transformation (see "TRANSFORMATION OF CELLS WITH PLASMIDS) after the DNA ligase treatment.

L) PURIFICATION OF RNA FROM PHAGE-INFECTED OR PLASMID-BEARING CELLS

In these and in other methods involving RNA manipulations, reagents were either treated with diethyl pyrocarbonate (DEPC; an irreversible inhibitor of RNase A) or made with DEPC-treated H_2O. The solutions or H_2O were first made 0.1% DEPC, allowed to stand at room-temperature overnight, and the excess DEPC was then evaporated away by autoclaving at 121°C for 30 min.

1) RNA FROM PHAGE-INFECTED CELLS

E. coli NapIV cells were grown at 30°C in 40 ml M9 medium to a density of $3x10^8$ cells/ml and concentrated 10-fold by pelleting at 3,000x g / 5 min and resuspending in 4 ml fresh M9. The cell suspension was pre-warmed for 3 min at the desired temperature, and phage infection was initiated by the addition of 2 ml pre-warmed M9 media containing 10^{11} plaque-forming particles (m.o.i. = 10) with vigorous aeration during incubation. At 2 min before harvesting cells for RNA, 0.1 ml of infection mixture was transferred to 10 µl pre-warmed M9 containing 2 µCi ^{35}S-methionine to label phage-induced proteins for a 5 min period as described in "MEASUREMENT OF GENE EXPRESSION." Such samples served to monitor the success of infections and production of the appropriate functional mRNA. At the desired time after infection, the infected culture was chilled in ice-ethanol bath for 5-10 sec and the cells were collected by

centrifugation at 3,000 x g for 5 min and resuspended in 4 ml ice-cold lysis buffer (10 mM Tris-Cl, pH 7.5 / 1 mM Na_2EDTA / 10 mM sodium azide / 1 mg/ml lysozyme). The suspension was subjected to two rapid freeze-thaw cycles in the presence of 50 units of RNasin (RNase inhibitor purchased from Promega Biotec). $MgCl_2$ was added to a final concentration of 10 mM followed by the addition of 150 units of RNase-free DNase I (Boehringer Mannheim Biochemicals cat#100-411). The lysing suspension was incubated on ice for 30-40 min before the addition of 1/10 volume of 0.04 M acetic acid and 1/20 volume of 10% SDS and incubation at room-temperature for a few min to complete the lysis. The cell lysate was extracted twice with acid phenol (equilibrated with 20 mM NaOAc, pH 5.2 / 20 mM KCl / 10 mM $MgCl_2$) followed by chloroform extraction to remove residual phenol. The nucleic acids were precipitated by the addition of 3 volumes cold 95% ethanol in the presence of 0.2 M NaOAc, pH 5.2. The ethanol precipitation was repeated once and the final RNA pellet was dissolved in 200 µl H_2O. The quality of the RNA preparation was examined on a standard DNA agarose gel (see "ELECTROPHORETIC METHODS") and the concentration was determined from optical density readings (1 O.D. unit at 260 nm = 45 µg/ml). When compared, different RNA preparations were examined for relative intensities of 16S and 23S rRNA bands on agarose gels after ethidium bromide staining.

2) RNA FROM PLASMID-BEARING CELLS

E. coli NapIV cells carrying lambdoid P_L-directed expression plasmids were grown to $3x10^8$ cells/ml in 20 ml M9-19 medium containing 20 µg/ml ampicillin. The cells were concentrated 5-fold by harvesting at 3,000x g / 5 min and resuspending in 4 ml fresh M9-19. The concentrated cell culture was aerated for at least 20 min at 42°C to induce P_L-directed gene expression from the plasmid. At 10 min before harvesting cells, a 0.1 ml aliquot of the cell culture was transferred to 10 µl pre-warmed M9-19 containing 5 µCi ^{35}S-methionine to label protein (usually for 20 min) as a control for the RNA preparation.

M) ELECTROPHORETIC METHODS

1) DNA AGAROSE GEL ELECTROPHORESIS

The standard method in our laboratory utilizes 3-5 mm thick 1% agarose horizontal slab gels prepared in 40 mM Tris-OAc, pH 8.0 / 20 mM NaOAc / 1 mM Na_2EDTA. 1/2 volume of DNA loading buffer* was added to a 10-20 µl solution containing DNA to be examined. After electrophoresis, the gels were stained in 1 µg/ml ethidium bromide solution for 30 min and destained in electrophoresis buffer for 30 min and DNA bands were visualized and photographed on a 302 nm UV trans-illuminator (Ultra-Violet Products, Inc.).

DNA loading buffer

(autoclaved solution)

Sucrose	25.0%	(w/v)
Bromphenol Blue	0.1%	(w/v)
Xylene Cyanole	0.1%	(w/v)
Na_2EDTA, pH 8.0	40.0	mM

2) PROTEIN SDS-POLYACRYLAMIDE GEL ELECTROPHORESIS

Reagents

Acrylamide stock

30.0	g	acrylamide
0.8	g	bis-acrylamide

Dissolved to a final volume of 100 ml, deionized by stirring at 4°C for 1 hr with 1 g / 30 ml of Analytical Grade Mixed Bed Resin (20-50 mesh; Bio-Rad), filtered over glass-fiber or 0.45 µm nitrocellulose filters, and stored at 4°C in the dark.

Lower gel buffer

0.4 g SDS dissolved in 100 ml of 1.5 M Tris-Cl, pH 8.8.

Upper gel buffer

0.4 g SDS dissolved in 100 ml of 0.5 M Tris-Cl, pH 6.8.

Electrode buffer

0.025 M	Trizma base
0.192 M	Glycine
0.1%	SDS

The pH of the solution usually ranged between 8.2-8.3.

Recipes for discontinuous gradient gels

	Polyacrylamide gel (%)			
	15%	12.5%	10%	Stack gel(4.5%)
Acrylamide	6.0 ml	5.0 ml	4.0 ml	0.75ml
Lower gel buffer	3.0 ml	3.0 ml	3.0 ml	-.--
Upper gel buffer	-.--	-.--	-.--	1.25 ml
H_2O	3.0 ml	4.0 ml	5.0 ml	3.00 ml
Glycerol	-.--	-.--	-.--	0.30 ml
Ammonium persulfate	0.045 ml	0.045 ml	0.045 ml	0.015 ml
TEMED	0.008 ml	0.008 ml	0.008 ml	0.008 ml

Procedure

The step-gradient gel was prepared as previously described (Karam *et al.*, 1977) with minor modifications using a Bio-Rad slab electrophoresis cell (Protean II). 9 ml of 15% gel solution, 9 ml of 12.5% gel solution, and 5 ml of 10% gel solution were sequentially poured into a 16 x 15.5 x 0.1 cm glassplate sandwich and 5 ml of 1/4 dilution of lower gel buffer was overlayed on the gel solution. The gel was allowed to polymerize at room-temperature overnight. Next day, the overlaying buffer was drained and the stacking gel was poured with sample comb inserted. The electrophoresis was carried out at 24 mA/gel.

3) ELECTROPHORESIS OF NUCLEIC ACIDS ON DENATURING POLYACRYLAMIDE GELS

Either 6% or 8% gels were used depending on the sizes of the DNA fragments studied. For analyzing sequencing reactions, a BRL apparatus Model S2 (gel dimension 395 x 315 x 0.4 mm) was used. For other purposes, e.g., S1-mapping, a home-made apparatus, gel dimension 230 x 138 x 1 mm, was used. Electrophoresis was performed in 1x TBE buffer (see below).

Buffers and Solutions:

10x TBE buffer

Tris-Cl, pH 8.3	0.5 M
Boric acid	0.5 M
Na_2EDTA	10.0 mM

Gel solution

	6%	8%
Acrylamide	5.8 g	7.76 g
Bis-acrylamide	0.2 g	0.24 g
Urea (7M)	42.0 g	42.0 g
10x TBE	10.0 ml	10.0 ml
H2O	50.0 ml	50.0 ml
Final volume	100 ml	100 ml

4) ACRYLAMIDE-AGAROSE COMPOSITE GELS AND RNA GEL SHIFT ANALYSIS OF GP43-MRNA BINDING

The 20 cm Protean II gel apparatus was employed for these assays. Care was taken to use only baked glassware and disposable plastic pipets for all operations in pouring and running these gels. The gel apparatus, gel plates, combs, spacers, and clamps were soaked overnight in a 2% solution of Absolve™(NEN cat#NEF971), rinsed with DEPC H_2O, and the plates were dried by baking for 0.5 to 1 hr at 160°C. The gel sandwich was assembled and warmed in the 42°C incubator until immediately before casting the gel. All handling of apparatus components was done while wearing (pre-rinsed) gloves.

0.2 g of GTG agarose(FMC) was dissolved in 32 ml of DEPC H_2O by heating in a microwave oven and transferred to a 60°C bath. While equilibrating, 4 ml of 10X Tris-glycine buffer (500 mM Tris base, 500 mM glycine) and 4 ml of 10X acrylamide stock (40% with 80:1 acrylamide to bis-acrylamide ratio) were added and the gel solution was briefly de-gassed (15 seconds) and returned to 60°C. When both the gel casting unit and the gel solution were ready and equilibrated, 167 µl of freshly made 10% ammonium persulfate and 12 µl of TEMED were added to the solution with gentle swirling. Immediately, the solution was poured into the casting sandwich using a sterile 60 ml syringe, and a 20-space comb inserted as quickly as possible. The gel was allowed to congeal for at least 3 hrs before beginning an hour-long pre-electrophoresis with 1X Tris-

glycine buffer. Samples (5 µl of the binding reactions) were loaded in wells under the buffer using long micro-capillary tips. RNA-protein complexes were subsequently resolved at 13 volts per cm for 2 or 2.5 hrs at 8°C and the gels were dried and packaged for autoradiography with pre-flashed Kodak XAR-5 X-ray film. Densitometric tracings of autoradiogram bands were carried out on a Bio-Rad Video Densitometer VD620.

N) PLASMID DNA SEQUENCING USING THE CHAIN-TERMINATION METHOD

Plasmid constructs that were suspects for operator duplex insertions were screened by sequencing small-scale DNA preparations by the method of Kraft (1988). All procedures were closely followed as described in this reference.

O) RNA SEQUENCING USING REVERSE TRANSCRIPTASE IN THE CHAIN-TERMINATION METHOD

1) REAGENTS AND BUFFERS

10x kinase buffer: 0.5 M Tris-Cl, pH 7.5 / 0.1 M $MgCl_2$

10x PE buffer: 0.5 M Tris-Cl, pH 8.3

5x dNTP mix: mixture of dTTP/dGTP/dATP/dCTP at 2 mM each in 1x PE buffer

5x ddNTP solutions: each one of the four dideoxynucleotides at 0.6 mM in 1x PE buffer

60 mM $MgCl_2$

Stop buffer:

0.1% xylene cyanol / 0.1% bromphenol blue /
10 mM Na_2EDTA, pH 7.0 / 95% formamide

2) 5'-END LABELING OF SEQUENCING PRIMERS

The following were mixed in an Eppendorf tube:

primer(16-20mer)	1.4 µl	(10 ng/µl or ~2.4 pmoles)
γ-^{32}P-ATP	8.3 µl	(6000 Ci/mmole; 10 mCi/ml)
10x kinase buffer	1.2 µl	
DTT (100 mM)	0.6 µl	
kinase (10 unit/µl)	0.5 µl	
Total	12.0 µl	

The reaction was carried out at 37°C for 1 hr and stopped by extracting with phenol and chloroform. The labeled primers were dried under vacuum and resuspended in 10 µl H_2O.

3) ANNEALING AND SEQUENCING REACTIONS

The following was prepared in an Eppendorf tube and referred to as the "primer-template mix":

RNA (dried)	15.0	µg
Labelled primer	2.5	µl
10x PE buffer	1.2	µl
H_2O to a final volume of	12.0	µl

The primer-template mix was heated at 60°C for 3 min to denature RNA and primer and immediately submerged in a dry-ice/ethanol bath until frozen. The frozen mixture was then placed on ice to thaw (about 20 min).

The following enzyme solution (termed "reverse transcriptase mix") was prepared:

AMV Reverse transcriptase (~25 unit/µl; from Boehringer-Mannheim Biochemicals)	2.0 µl
10x PE buffer	2.5 µl
$MgCl_2$ (60 mM)	10.0 µl
H_2O	10.5 µl
Total	25.0 µl

An equal volume of each of the following components was added to one of a set of 4 tubes (termed "reaction mixtures") designated G, A, T, and C:

--- 5x dNTP mix

--- reverse transcriptase mix

--- one of the four 5x ddNTP solutions.

The following were added to another set of 4 different tubes:

--- Primer-template mix 2 µl

--- One of the four "reaction mixtures" 3 µl

The reactions were carried out at 48°C for 15 min before an additional 1.5 µl of "reaction mixture" was added to each of the four reaction tubes and the incubations were continued at 48°C for another 15 min. The reactions were then terminated by the addition of 10 µl stop buffer, and 2.5 µl samples were analyzed on a sequencing gel.

P) *IN VITRO* RNA SYNTHESIS: T7 AND SP6 PHAGE RNA POLYMERASE-DIRECTED TRANSCRIPTION

1) HIGH-SPECIFIC ACTIVITY LABELING OF RNA PRODUCTS

Solutions:

5x buffer

0.2 M	Tris-Cl, pH 7.5
30 mM	$MgCl_2$
10 mM	spermidine
50mM	NaCl

Nucleoside triphosphate stocks

Solutions of each of ATP, CTP, UTP, and GTP at 10 mM, pH 7.0; either purchased from Promega Biotec. or as 100 mM Stocks from Pharmacia.

100 mM dithiothreitol (DTT)

Procedure for ^{32}P-labeled RNA probes:

The following were mixed in the order shown:

5x buffer	4 µl
100 mM DTT	2 µl
RNasin	20 units
Mix of 2.5 mM each of ATP, CTP,and GTP	4 µl
1 mM cold UTP	0.5 µl
linearized DNA	0.2 µg
α-^{32}P-UTP (10 mCi/ml; 800 Ci/mmole)	8 µl
SP6 or T7 RNA polymerase	10 units
H_2O added to make the final volume	20 µl.

The reaction mixture was incubated at 37°C for 1 hr before 1 µg (about 3 units) of DNase I was added to hydrolyze the DNA templates at 37°C for 15 min. The mixture was then extracted twice with acidic phenol and once with chloroform. The labeled RNA probes were ethanol precipitated in the presence of 0.2 M NaOAc, pH 5.2, and 0.5 µg of tRNA used as carrier. The RNA pellet was dissolved in 50 µl H_2O and centrifuged through a Sephadex G-50 pre-spun column (RNA grade; Boehringer-Mannheim Biochemicals) to remove unincorporated nucleotides. The radioactivity was determined by liquid scintillation counting and was usually about $1x10^6$ cpm/µl.

2) LOW-SPECIFIC ACTIVITY LABELING OF RNA PRODUCTS

This was the method employed when RNA was synthesized for use in gp43-RNA binding studies, *in vitro*.

Linearized plasmid DNA templates containing a P_{T7} promoter were transcribed *in vitro* using methods based on the studies of Melton and Milligan (Melton, *et al.* 1984; Milligan, *et al.* 1987). We obtained optimal RNA synthesis in mixtures (usually 100 µl) containing 40 mM Tris-Cl, pH 7.5, 6 mM $MgCl_2$, 2 mM spermidine, 10 mM NaCl, 10 mM dithiothreitol, 200 units RNasin (40 Units/µl; Promega, cat #N2512), four ribo-NTPs at 1.25 mM each (prepared from 100 mM stocks pH 7.0; Pharmacia, cat #27-20X6-01), 6 µCi ^{32}P-UTP (10 mCi/ml; 800 Ci/mmole; Dupont NEN, Boston cat# NEG007X; final specific

activity 480 mCi ^{32}P/mmole UTP), 4% polyethylene glycol (PEG MW6000), 10 units T7 RNA polymerase, and 20 nM linearized DNA. Incubations were at 37°C for 2 hrs before RNase-free DNase I (60 units, Boehringer-Mannheim Biochemicals; cat #776785) was added to hydrolyze the DNA templates at 37°C for 30 min. After phenol (pH5.2):chloroform (1:1) extraction, the ^{32}P-labeled RNA was ethanol-precipitated in the presence of 0.2 M sodium acetate, pH 5.2 and 0.5 µg tRNA carrier. The RNA pellet was dissolved in 50 µl diethylpyrocarbonate (DEPC)-treated H_2O and centrifuged through a Sephadex G-50 spun column (RNA grade; Boehringer-Mannheim Biochemicals, catalog #100411) to remove unincorporated nucleotides. Acid-precipitable radioactivity was determined by liquid scintillation counting and was usually 6-10x10^4 cpm/µl. The RNA concentration ranged between 1 and 2 pmoles/µl. *In vitro* transcription products of all operator constructs consisted of 56 nucleotides of vector sequence at the 5' end followed by the 151-nucleotide T4 sequence from position -54 to position +97 relative to the initiator AUG.

Q) ANALYSIS OF GENE EXPRESSION BY *IN VITRO* TRANSLATION

1) REAGENTS

Reagents were prepared exactly according to Pratt (1984). The details of isolating the S30 extracts are detailed below. The solutions needed for this procedure and for the cell-free translation assays (*i.e.* LM/Mg supplemental

mix) are carefully detailed in the reference. All solutions and glassware were treated to prevent RNase contamination (Blumberg, 1987).

2) PREPARATION OF S30 *E. COLI* EXTRACTS.

E. coli strain MRE600 was used in the preparation of S30 extracts for *in vitro* transcription and translation. This strain is the favored source of extracts for these applications since it lacks the major RNase activity of *E. coli*. Protocols closely followed the detailed methods of J. M. Pratt (1984). The following steps are an outline of procedures used; further details can be found in the above reference:

50 - 100 ml of cells were grown for 8 hrs to an optical density of 7 at 450 nm, to be used as inoculum for the large scale liter cultures. Each liter (10 liters total) was inoculated with 10 ml of this dense culture to give an A_{450} of 0.07, and aerated vigorously in an environmental shaker at 37°C overnight. In the morning, the final A_{450} ranged from 1.5 to 2.0. The cells were harvested at 10,000 x g for 15 min at 4°C. The cells were then washed with a total of 500 ml of ice-cold S30 buffer containing 0.5 ml of 2-mercaptoethanol/liter. The wash was repeated twice and the wet weight of cells was estimated. Pellets were then stored at -70°C after freezing on ice-ethanol, overnight (2 days maximum storage).

The cell pellet was thawed on ice for 30-60 min, while fresh S30 buffer was made. 100 ml of S30 buffer was added for each 10g of cells (from wet weight estimate of cell pellet), and the pellet resuspended. Cells were subsequently centrifuged at 16,000 x g for 30 min at 4°C. The supernatant was removed and the wet weight was again estimated. Cells were transferred to a baked side-arm flask and 63.5 ml of S30 buffer was added per 50g of cells. The cells were resuspended with a Teflon pestle with frequent evacuation of the flask (as anaerobic as possible), until all cell clumps were dispersed. With a plastic pipet the volume of cell suspension was noted, and transferred to a French pressure cell. The cells were lysed at 8400 psi, and the lysate collected on ice in Corex centrifuge tubes. DTT(100 µl of 0.1M solution) was added for every 10 ml of lysate. The lysate was then immediately centrifuged at 30,000 x g (16,000 rpm in SS34) for 30 min at 4°C. The upper four-fifths of the supernatant was transferred to a new centrifuge tube and centrifuged again at 30,000 x g for 30 min at 4°C.

The upper four-fifths was transferred to a 50 ml conical tube noting the volume. 7.5 ml of pre-incubation mix was added per 25 ml of supernatant, and subsequently incubated at 37°C with gently swirling for 80 min. The pre-incubated S30 lysate was then dialyzed against 50 volumes (500 ml beaker) of S30 buffer. The dialysis solution was changed 3 times, allowing 45 min for equilibration in each case. The extract was then centrifuged at 4,000 x g for 10 min at 4°C, and the

supernatant was aliquoted, frozen in dry-ice-ethanol and stored at -70°C or under liquid nitrogen. S30 extracts were always thawed a minimum number of times and returned to -70°C as soon as possible.

3) METHODS FOR *IN VITRO* TRANSCRIPTION AND TRANSLATION ASSAYS.

The concentration of S30 extract and magnesium ion concentration was optimized for the specific S30 lysate prepared. A sample reaction consisted of:

2.5 µl	DNA & protein (or storage buffer)
10.0 µl	LM/Mg mix (see Pratt, 1984)
2.5 µl	S30 extract
15.0 µl	total volume

Nucleic acid (2µg DNA, 50µg total cellular RNA) was added to a 0.5 ml microfuge tube and dried in a speed vacuum, for greater than 15 min. The nucleic acid was then resuspended in DEPC H_2O and either protein or protein storage buffer. If desired, the protein and RNA were pre-incubated for 10 min (pre-binding), prior to the addition of the LM/Mg mix and S30 extract. The reaction was mixed well and incubated at 37°C for 20 - 30 min. Subsequently, proteins were extracted by the addition of 100 µl of extraction buffer, boiled 1.5 min, vortexed, and stored at -20°C prior to quantitative analysis by SDS-PAGE, autoradiography, and densitometry. Purified T4 DNA polymerase was initially provided kindly by Navin Sinha (Rutgers University). Later

studies used gp43 purified from overproducing clones that I had constructed.

R) CONSTRUCTING AND CLONING MUTANT TRANSLATIONAL OPERATORS

All synthetic deoxyribonucleotides were prepared in an Applied Biosystems (Foster City, CA) Model 380B DNA synthesizer, and were deprotected and desalted according to manufacturer's instructions. To anneal duplexes for cloning purposes, equal amounts of complementary strands (at O.D.= 2.0 each) were mixed in a microfuge tube, placed in boiling water (500 ml in a beaker), which was subsequently allowed to slowly cool to room temperature. Ligations were carried out with a 35:1 molar excess of annealed duplex to appropriately digested cloning plasmid (100 - 200 ng DNA) in 10 µl mixtures containing 100 mM Tris-Cl, pH 7.5, 10mM $MgCl_2$, 10 mM dithiothreitol, 1 mM ATP, and 9 Units of T4 DNA ligase (Boehringer-Mannheim, catalog #799009). After overnite incubation at 14 - 16°C, 5-10 µl of the ligation mixture was used to transform *E. coli* K802. Ampicillin-resistant colonies from transformations were screened for successful cloning of the synthetic DNA by a colony hybridization assay (Paddock 1987-see above) and the nucleotide sequence of the inserts was subsequently confirmed by double-stranded plasmid DNA sequencing (Kraft, *et al.* 1988). A summary of all mutant

translational operators and the details of their construction can be found in RESULTS.

S) PROTEIN - mRNA BINDING ASSAYS

The purified DNA polymerase (gp43) used in this study was prepared from the overproducing clones described in RESULTS by using a purification scheme involving anion-exchange, phosphate-affinity and polyribonucleotide-affinity chromatography *(manuscript in preparation)*. The binding of gp43 to ^{32}P-labeled *in vitro* RNA was analyzed by using an RNA gel mobility shift assay. All solutions used for the assay were pre-treated with DEPC as described (Blumberg 1987). Binding reactions were carried out in 10 µl of a solution containing 42 mM Tris-acetate pH 7.8, 200 mM potassium acetate, 5 mM dithiothreitol, 1 mM $MgCl_2$, 50 ng tRNA (in addition to the tRNA used as carrier in ethanol precipitation after *in vitro* transcription), and 35% (v/v) glycerol. Equal volumes of an RNA mix (containing labeled RNA, buffer, salts, and tRNA) and a protein mix (containing purified T4 DNA polymerase at the desired concentration in 70 % glycerol) were combined, and incubated for 1 hr at room temperature. Reaction mixtures generally contained 0.5 pmole of ^{32}P-labeled RNA (50 nM) while gp43 concentrations ranged from 10^{-12} to 10^{-6} M. In each case, 5 µl of binding reaction was subjected to electrophoresis in a composite gel consisting of

4% polyacrylamide (80:1 ratio of acrylamide to bis-acrylamide) and 0.5% agarose.

Densitometric tracings of autoradiogram bands were carried out on a Bio-Rad Video Densitometer VD620. The relative amounts of "free" (electrophoretic mobility unaffected) and "bound" (electrophoretic mobility retarded) RNA were determined from the densitometric scans and used in the calculation of dissociation constants (K_d values) for the gp43-RNA complexes, assuming a bi-molecular interaction between gp43 and its RNA substrate. K_d values were obtained by using an algebraic rearrangement of the defining equation for the K_d of the reaction: gp43-RNA complex ↔ free gp43 + free RNA, *i.e.*

$$K_d = \frac{[\text{free gp43}]\ [\text{free RNA}]}{[\text{gp43-RNA complex}]}, \text{ or}$$

$$K_d = \frac{([gp43]_t - F[RNA]_t)\ ([RNA]_t - F[RNA]_t)}{F[RNA]_t}$$

where F is the fraction of total RNA bound as determined from the gel assay, *i.e.* F= [bound RNA]/[total RNA], and since [gp43-RNA complex] is equal to the [bound RNA], then [gp43-RNA complex] = $F[RNA]_t$; $[RNA]_t$ and $[gp43]_t$ are the input concentrations of operator RNA and gp43, respectively. A K_d was calculated for each data point from a binding curve. Then, an average K_d was determined for each curve and an

overall average K_d derived for each operator from the results of several experiments.

T) RAISING ANTISERA AGAINST CONJUGATED PEPTIDES

1) Conjugation of peptides to Bovine serum albumin

Five mg of oligopeptide was dissolved in 0.5 ml H_2O. Depending on peptide solubility, it was necessary to dissolve initially in acid (70% formic) or base (30% ethylmorpholine) and subsequently to dilute and re-concentrate the peptide solution (oligos Karam-3 and Karam-4 used 1-methyl 2-pyrrolidinone). Sodium phosphate buffer (0.4 ml of 0.25M pH 7.2 - 7.5) containing 12.5 mg/ml purified BSA was added and the conjugation reaction was initiated by the addition of 10 µl of 12.5% glutaraldehyde. The reaction was incubated at room temperature for 30 min before it was stopped by the addition of 100 µl of 1M ethanolamine (pH 8.0). The conjugate was immediately purified by chromatography on a Sephadex G-50 column, monitoring the absorbance at 280 nm.

2) Raising antisera in New Zealand-White rabbits

Since excess protein immunization can result in the production of low affinity antibodies, we aimed at injecting 0.5 mg of protein/injection. Typically, the concentration of the conjugate was 5 mg/ml, and was diluted 1:10 with H_2O (100 µl + 900 µl dH_2O) and chilled on ice in preparation for

emulsion procedures. Likewise, a 1 ml aliquot of Freund's incomplete adjuvant (complete adjuvant was used on first injection, only) was chilled in 15 ml Corex tube for each conjugate to be injected. The adjuvant was firmly held during vortexing on the fastest setting, while the conjugate dilution was added drop-wise with a Pasteur pipet. Slow addition is essential in the preparation of a good emulsion, which has the consistency of stiff whipped cream. The emulsion was slowly aspirated into a 5 ml sterile syringe with an 18-gauge needle. The needle was then changed to 22-gauge for sub-cutaneous injection. Care was taken to remove air from the syringe and maintain sterility. Typically, the rabbit was injected in 4 places in dorsal areas where lymph nodes drain. Injections were given once a month (or less) until titers reached their maximum level. Depending on the rabbit and the peptide, this rarely exceeded 4 months. Injections thereafter, were only performed as needed due to dropping titers ("booster shots").

Blood (25 - 50 ml) was collected by ear bleeds, about once per month, (two weeks following the booster injections). More frequent bleeding was avoided in order to not induce unwanted scarring of ear veins. The blood was incubated at 37°C for 0.5 to 1 hr to induce clotting. The clot was freed from the walls of the tube with a sterile pipet and the blood centrifuged at 6000 RPM in the HS-4 rotor . The serum (*i.e.* the supernatant) was removed with sterile technique and aliquoted for storage, and titered for antibody levels.

3) TITERING ANTISERA

REAGENTS

20XPBS - (phosphate buffered saline)

115.0 g Na_2HPO_4 (di-basic, anhydrous)

29.6 g NaH_2PO_4 (mono-basic, di-hydrate)

850.0 g NaCl

Made up to 5 liters with dH_2O.

Washing Solution (PBS/Tween-20 pH 7.8)

10 ml 10% Tween-20

100 ml 20XPBS

Made up to 2 liters with dH_2O

Blocking solution (5% gelatin in washing solution)

A volume of 50 ml/plate was needed to complete the assay procedures.

16.6 ml gelatin (heated in microwave, carefully)

83.4 ml PBS Tween-20 (wash solution)

Primary antibody sera

The primary antibody (rabbit serum sample to be tested) was diluted in blocking solution. When a low titer was expected a 1/100 dilution was used; high titer serum was diluted 1/1000 or 1/10,000. To decrease non-specific binding in the assay, the dilutions were done 15 min before use, in

order to pre-adsorb to fish proteins in the blocking solution.

Secondary Conjugate antibody

Goat-anti-Rabbit Peroxidase Conjugate(Jackson Labs or Biorad)

1/1000 dilution in blocking solution

20 ml will coat 2 plates

like primary antibody, it is best to dilute 15 min prior to use

Citrate Buffer (0.1M)

1.387 g citric acid (free acid)

1.000 g sodium citrate (tri-sodium salt, dihydrate)

made up to 100 ml with dH_2O

ABTS Stock (1.5 %; make fresh every 2 weeks)

2,2'-azinobis(3-ethylbenz-thiazoline sulfonic acid - Sigma cat#A-1888)

15.0 mg ABTS

1.0 ml dH_2O

Substrate Solution

9.8 ml Citrate buffer (0.1M)

0.2 ml ABTS Stock (0.03%)

10.0 µl H_2O_2 (use 30% stock)

after addition to plate, the OD was read at 414 nm

PROCEDURE

Coating plates

On the preceding day, ELISA grade microtiter plates (NUNC cat#2-69620) were coated with peptide according to the procedure detailed below. Typically, 2 dilutions of each serum sample were tested (1:1000 and 1:10,000), and thus each sample required at least 2 columns on the microtiter plate. It is advisable to test some serum samples for cross-reactivity with other peptides. Therefore, each plate can test about 5 serum samples. The protocol is as follows:

1) One mg of peptide was weighed out and dissolved in 100 µl of stock ammonium hydroxide(N-term gp43 peptide), or 100 µl of 30%acetic acid (C-term peptide).

2) The peptide was diluted (1:100,000) in 1XPBS to 0.1 µg/ml and 100 µl (10 ng/well) added to each well using a multi-channel pipetor. The first column of each plate is omitted from coating so as to provide a reference control for the microtiter plate reader.

5) The plate was incubated at 37°C overnight uncovered.

6) If wells were not completely dry in the morning, plates were incubated at 42°C while solutions for the assay were prepared fresh.

ELISA procedure

Washes (repeated 4 times at each stage) and incubations were all done at room temperature. Coated plates were washed with PBS + 0.05% Tween, pH 7.8 (washing solution), and then

the remaining binding sites in the wells were blocked by an hr incubation with the blocking solution. After washing, the plate was incubated at room temperature for at least 1 hr with varying dilutions of rabbit serum. After washing again, the plate was incubated for 1 hr at room temperature with a goat-anti-rabbit peroxidase conjugate antibody, diluted 1:1000 in the blocking solution. The non-specifically bound secondary antibodies were removed by several washes. The relative concentration of peptide-specific antibody was determined colorimetrically at 414 nm wavelength by a multi-titertek scanner after the addition of the substrate ABTS 0.03% (w/v) in 0.1M citrate buffer with 0.1% (v/v) hydrogen peroxide. Nonspecific binding was determined, by carrying a plate not coated with the peptide through all procedures, and the background subtracted from the values obtained with the peptide coated plate.

U) QUANTITATION OF mRNA BY S1 ANALYSIS

REAGENTS

Hybridization buffer

80% (v/v)	formamide
20 mM	PIPES (1, 4 - piperazine - diethane sulfuric acid, sodium salt)
0.4 M	NaCl
2.0 mM	Na_2EDTA, pH 7.0

S1 nuclease solution

800 units of S1-nuclease (Boehringer-Mannheim Biochemicals) per 300 ml of a solution of 0.1 M NaCl / 60 mM NaOAc, pH 4.5 / 2 mM ZnSO4.

Loading buffer

0.1%	Xylene cyanole
0.1%	Bromphenol blue
10 mM	Na_2EDTA, pH 7.0
95%	Formamide

PROCEDURE

A mixture of 5 mg total RNA and $2x10^6$ cpm of complementary RNA probes (see "*IN VITRO* RNA SYNTHESIS") were lyophilized and the dried pellet was resuspended in 30 ml of ice-cold hybridization buffer. The mixture of mRNA and protection probe was heated at 90°C for 5 min and immediately transferred to the pre-set water-bath (42-45°C for protected fragments over 250 bp in length and 30-35°C for protected fragments less than 100 bp in length). After overnight incubation, 300 ml of chilled S1 nuclease solution were added and S1 nuclease digestion was carried out at 20°C for 0.5 hr and at 4°C for an additional 15 min. The reaction mixture was then extracted once with acidic phenol and once with chloroform. The nucleic acids were precipitated by adding 3 vol of cold 95% ethanol in the presence of 0.2 M NaOAc, pH

5.2, and 0.5 mg tRNA at 1 mg/ml as carrier. The nucleic acids were collected by centrifugation at 10,000x g for 30 min and resuspended in 20 ml loading buffer. 2-3 ml of the final samples were analyzed on a 6% or 8% denaturing polyacrylamide gel (see "ELECTROPHORETIC METHODS")

V) COMPUTER PROGRAMS USED

The alignment of T4 DNA polymerase with Herpes Simplex virus (HSV) DNA polymerase was done with the program CLUSTAL (Higgins and Sharp, 1988), most often used for multiple sequence alignments. This program produces a maximal alignment while introducing a minimal number of gaps. Default values were used for gap penalties, and residue homologies are designated as conservative substitutions according to the current Dayhoff table (internal to the program).

RNA folding was done with either of two programs 1) the GCG FOLD program (Zuker and Stiegler, 1981), part of the set of sequence analysis programs from the Genetics Computer Group (formerly Univ of Wisconsin), 2) the newer program MULFOLD (Jaeger, et al., 1989; Jaeger, et al., 1990; Zuker, 1989), which has several improvements including the ability to calculate sub-optimal RNA structures. A version of the latter program has been compiled for the Macintosh line of computers and can fold up to 300 bases, available via anonymous ftp to iubio.bio.indiana.edu.. A VAX version was

compiled by Richard Gadsden and myself on the MUSC VAX 11/785, and performs calculations much faster, and can fold more than 1500 bases. A Macintosh program for viewing RNA structures was used in some figures (Gilbert, 1990).

In the course of these studies, various programs from the GCG sequence analysis package were employed (Devereux, et al., 1984). In addition, the search for protein motifs in gp43 was aided by the program MacPattern (Fuchs, 1990) which searches a protein sequence for any protein motifs that are contained in the PROSITE database (Bairoch, 1990). A similar analysis was performed with the PROFILESCAN program of the GCG package.

III) RESULTS

A) PART I. THE MECHANISM OF AUTOGENOUS REGULATION OF T4 DNA POLYMERASE BIOSYNTHESIS

1) RATIONALE AND SIGNIFICANCE

The process of genome duplication is central to the physiological program and life cycle of any organism. My studies represent a subset of this laboratory's general interest in the control of bacteriophage T4 DNA replication. Our previous studies have focused on a cluster of genes that encodes several proteins essential for DNA replication in bacteriophage T4. Figure 10 diagrams this genetic cluster, summarizes the function of these different gene products, and depicts some of the transcriptional and post-transcriptional events involved in the regulation of their expression.

The assorted mechanisms of gene regulation utilized by phage T4 are abundantly demonstrated within this particular set of genes:

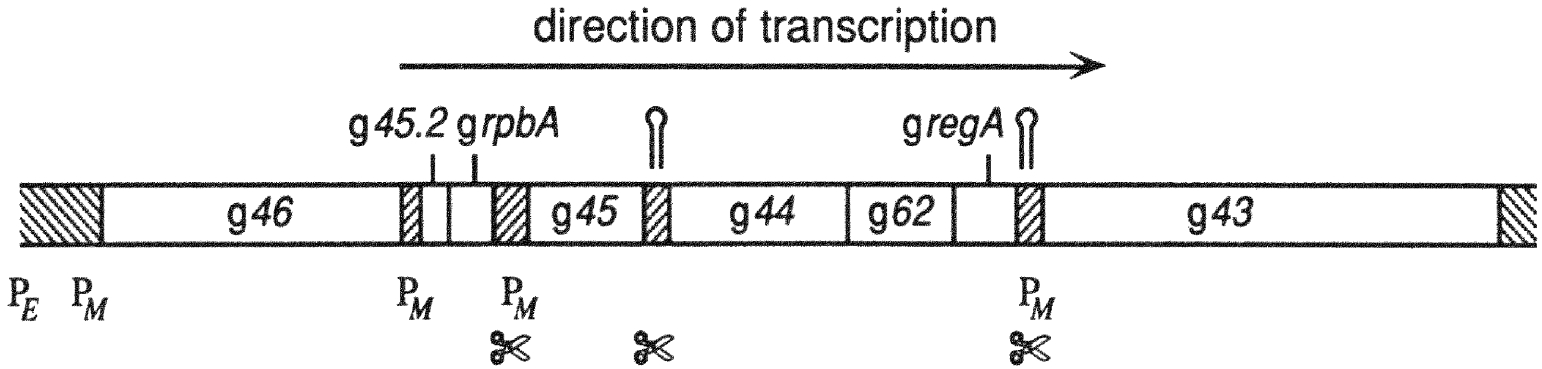


Figure 10 Key:

P_M denotes a T4 middle promoter that is dependant on T4 MotA protein

P_E denotes a T4 early promoter that is E. coli-like

✂ denote places of mRNA processing by RNA endonucleases

The stem and loop hairpin (seen above genes) signifies mapped transcription termination sites

The intercistronic regions are stripped

Gene	Function/Function	Aspects of Transcription	Aspects of Translation
43	DNA polymerase	Activated by MotA protein; also encoded in transcripts that readthrough termination signals	Autogenous translational repression
regA	Translational repressor of replication accessory proteins and other T4 proteins	Co-transcribed with upstream genes	Autogenous translational repression
62	Subunits of ssDNA dependent ATPase; function together to enhance DNA polymerase processivity	Co-transcribed with upstream genes	Translation coupled to gp44 translation
44		Co-transcribed with upstream gene 45 and dependent on read-through of upstream termination signals	Translational repression by regA protein
45	Phage modification of host RNA polymerase; also enhances 44/62 ATPase	MotA dependant promoter immediately upstream; also co-transcibed from other upstream promoters	Translational repression by regA protein
rpbA	Host RNA polymerase binding protein of unknown biological function	Co-transcibed from upstream promoters	Translational repression by regA protein

-- *Differential promoter utilization.*
Transcriptional activation of the genes in this cluster by a transcription factor, the T4 MotA protein, which modifies the host RNA polymerase to recognize T4 middle promoters (see promoter symbols in Figure 10 - (Guild, *et al.*, 1988; Hsu and Karam, 1990)).

-- *Termination of transcription* at various sites within this cluster (Hsu and Karam, 1990 - see hairpin loops in Figure 10).

-- *Processing* of polycistronic mRNA's (Hsu and Karam, 1990 - see scissors in Figure 10).

-- Differential efficiency of *translation initiation* sites, representing a spectrum of ribosome binding domains (*i.e.* initiation sequences for gp62 translation are weak compared to others in this cluster - (Trojanowska, *et al.*, 1984)).

-- *Translational coupling*: gp62 translation is coupled to gp44 translation (*i.e.* amber mutations in gene 44 have polar effects on gp62 translation), (Gerald and Karam, 1984).

-- *Translational repression* (of gp44, gp45, and gp rpbA synthesis) by the RegA protein, via binding to different albeit specific mRNA target sequences.(Webster, *et al.*, 1989; Winter, *et al.*, 1987).

-- *Autogenous translational repression* by the regA protein via binding to its mRNA in the translation initiation region (Liang, *et al.*, 1988).

These examples show this gene cluster's suitability as a model system not only for the study of gene regulation but also the coordinate biosynthesis of a multi-enzyme complex. Each component of this phage "replisome" needs to be made available in a temporally controlled fashion, and maintained at proper molar ratios for optimal growth. In addition, the isolation, purification, and characterization of each of the replication proteins make T4 an ideal candidate for structure-function studies toward a molecular understanding of genome replication and regulation of gene expression.

As outlined in Figure 10, translational repression is a major mechanism by which bacteriophage T4 regulates DNA replication. In addition to the gp43 studies being highlighted in this dissertation, both gp32 and gp regA have been well characterized as autogenous translational repressors. The RegA protein also acts as a global translational repressor of several genes involved in DNA replication. While both appear to function by inhibiting translation initiation at the ribosome binding site (RBS), each seems to recognize a characteristic RNA structure.

My studies, to elucidate the mechanism by which gp43 acts in autogenous repression, sought to further our understanding in this pivotal area of the regulation of DNA

replication. Previously, the mechanism of gp43 autogenous repression had attracted some interest but not been elucidated. Some early experiments were cited to support a transcriptional mechanism for feedback repression by this known DNA binding protein. However, these experiments were carried out before assays for the quantitation of gene-specific RNA became available. In the context of the variety of regulatory mechanisms observed in this gene cluster, studies to define the mode of regulation for this key replication protein were well justified. The approach of the studies presented in this section was to first characterize protein and mRNA synthesis *in vivo*, using wild-type and protein overproducing phage mutants. These studies then focused on developing *in vitro* assays to delineate the mechanism of gp43 action in autogenous repression.

2) AUTOGENOUS REPRESSION OF T4 DNA POLYMERASE *IN VIVO*

Evidence for gp43 autogenous repression was initially derived from experiments that examined protein synthesis in infections with missense and nonsense mutants in the structural gene. The results shown in Figure 11 provide examples of the type of observations initially made by Russel (1973) that T4 gene *43* nonsense mutants overproduce gp43 amino-terminal fragments, and further demonstrate that wild-type gp43 can repress the synthesis of mutant gp43 *in trans*. Lane 1 shows a typical level of gp43 synthesized in wild-type infections. The overproduction ("derepression") of two amber fragments is seen in lanes 4 and 5. In lane 2, the wild-type

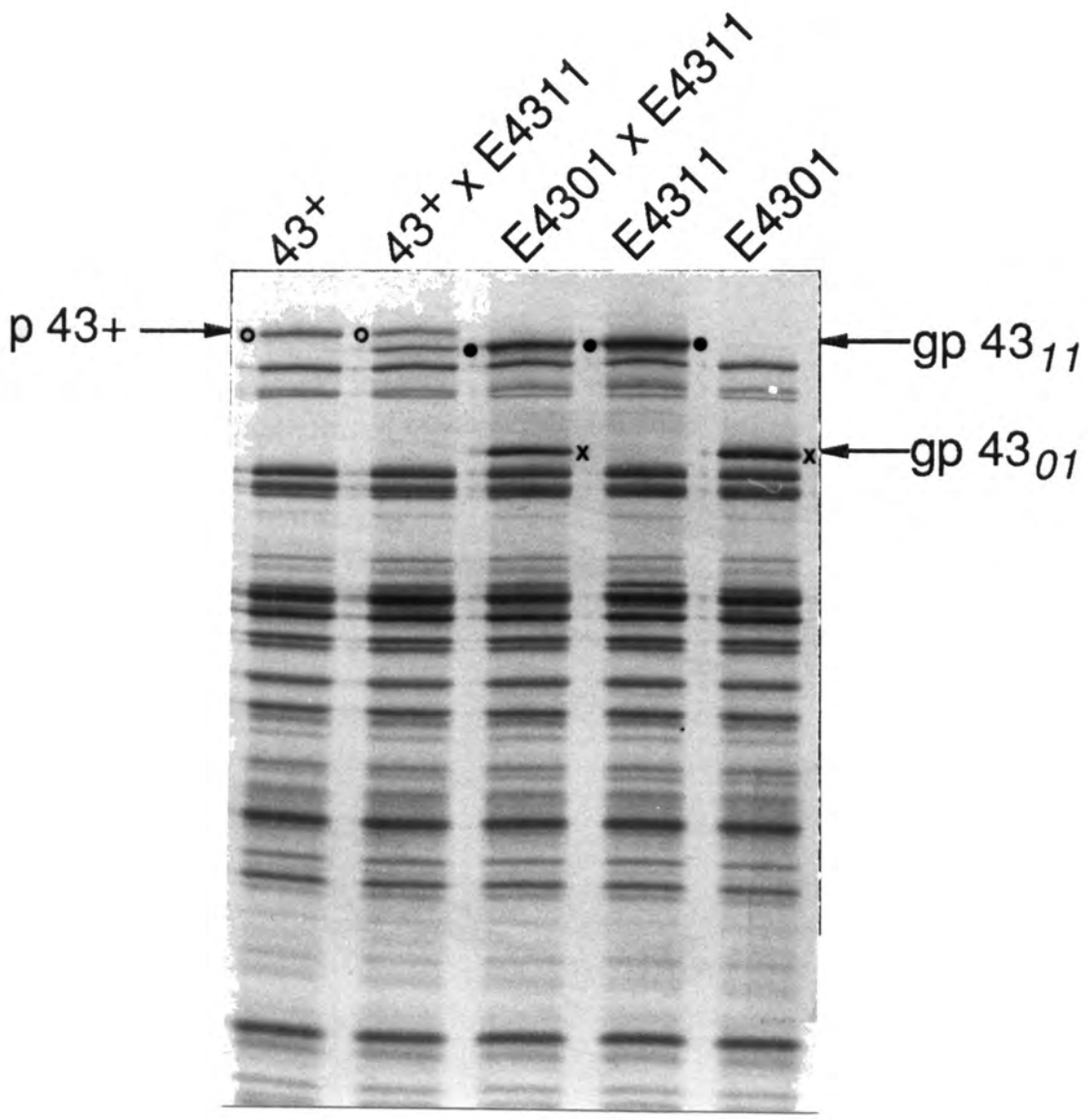
gp43 synthesized is able to repress *in trans* the biosynthesis of the mutant polypeptide, $gp43_{11}$. In contrast, co-infection of cells with two gene *43* mutant phage did not result in repression of either polypeptide.

Autogenous regulation of gene *43* expression resembles other systems of prokaryotic gene regulation in which diffusible repressor substances bind to nucleic acid target sequences.

While demonstrating the $gp43^{+}$-mediated repressor function, these and other similar experiments do not address the mechanism by which gp43 acts as a repressor. Several plasmid constructs, which were previously studied in our lab, prompted us to consider the possibility that gp43 could repress at the level of translation. These plasmids encoded the regA protein and the amino terminal portion of gp43 fused to ß-galactosidase on a single mRNA transcript, transcribed from the heat-inducible Lambda P_L promoter. Although transcribed on the same mRNA, the gp43/ß-gal fusion repressed its own synthesis but not that of regA protein (Hughes, Ph.D. Thesis). This result called into question a transcriptional mode of autogenous regulation. In the context of other known modes of regulation within this gene cluster, it was reasonable to consider the possibility that gp43 regulated its own translation. As a first step in examining the mechanism of autogenous repression, I quantitated the amount of gene 43-specific mRNA in the same infections used in the experiment shown in Figure 11.

Figure 11: Autogenous repression of T4 DNA polymerase synthesis *in vivo*

An autoradiogram showing the separation of ^{35}S-labeled T4-induced proteins by SDS-gel electrophoresis, and the ability of a wild-type gene *43* allele (*43+* to repress the expression of a mutant allele (*43amE4311*) in trans. Log-phase *E. coli* cultures (at $2x10^8$ cells/µl) were infected with phage strains carrying the designated alleles of gene *43*. All the T4 strains used also carried the gene *44* nonsense mutation *44amN82*, which disallows T4 DNA replication and late gene expression in the NapIV host. The multiplicity of infection was 10 (5 of each phage in coinfections) and infection mixtures were labeled with 5µCi ^{35}S-methionine/ml for 10 min beginning at 15 min postinfection. Extracts were then prepared and analyzed as detailed elsewhere. Note the overproduction of mutant gp*43* nonsense fragments gp43_{11} (*E4311* lane) and gp43_{01} (*E4301* lane) relative to the wild-type gp*43* (*43*$^+$ lane). Note also the reduction in gp43_{11} synthesis in the *43*$^+$ x *4311* relative to the *E4301* x *E4311* experiment.

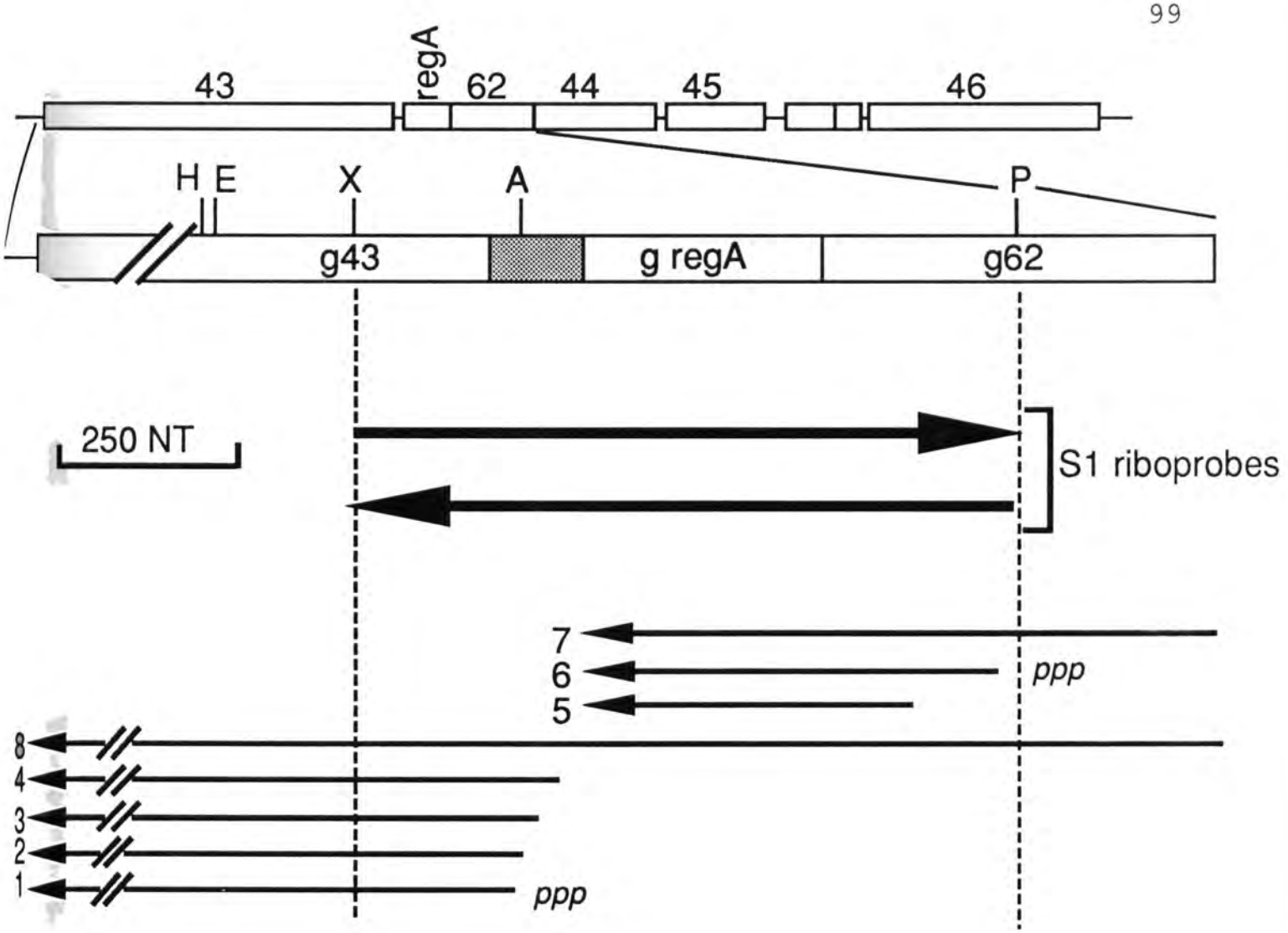

43+
43+ x E4311
E4301 x E4311
E4311
E4301
p 43+
gp 43 11
gp 43 01

3) AUTOGENOUS REPRESSION OF T4 DNA POLYMERASE SYNTHESIS OCCURS POST-TRANSCRIPTIONALLY.

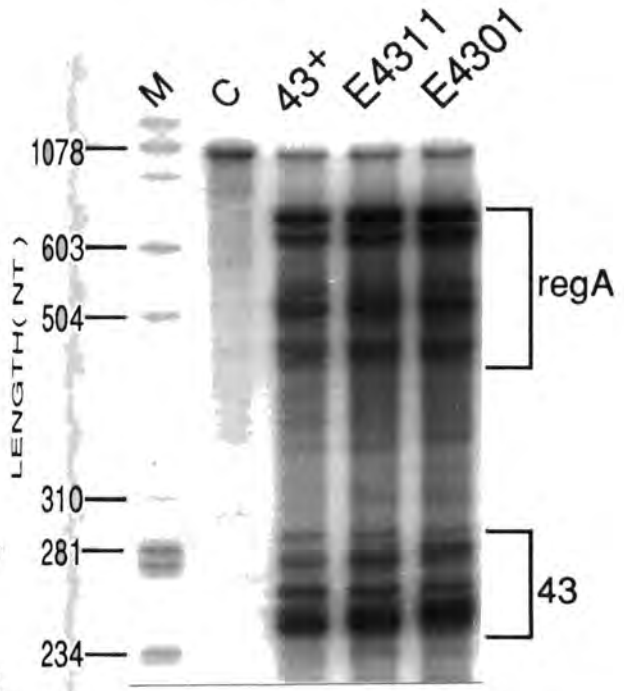
To test if gene *43* mutant infections that overproduce 43 protein also overproduce gene *43*-specific mRNA, I used an S1 mapping assay with RNA probes spanning the region shown in Figure 12; the length of the probe enabled the measurement of levels of both gene *43* RNA and *regA* RNA. The levels of *regA* RNA provided a control, since it is known to be regulated independently of T4 gene *43* (Karam and Bowles, 1974; Karam, *et al.*, 1977). As shown in Figure 12, whereas gp43 was overproduced 6-10 fold in the mutant infections relative to wild-type, the levels of gene *43* mRNA were not similarly affected. In the S1 analysis, the multiple RNA bands that were detected for genes 43 and regA reflect overlapping modes of transcription and post-transcriptional processing that characterize these two cistrons *in vivo* (Hsu and Karam, 1990). Although, the relative intensities of individual RNA bands vary slightly, the overall amount of gene 43-specific RNA remained unaffected in 43^- versus 43^+ infections. This is especially obvious when one normalizes the level of gene *43* mRNA synthesized to the level of gene *regA* mRNA synthesized. Similar quantitation of *43* mRNA in *ts* gene *43* mutants gave identical results shown in the panel on the right in Figure 12. Therefore, the observed derepression of gene *43* expression was not due to increased transcription.

Figure 12: Quantitation of gene 43 mRNA in wild-type and mutant infections

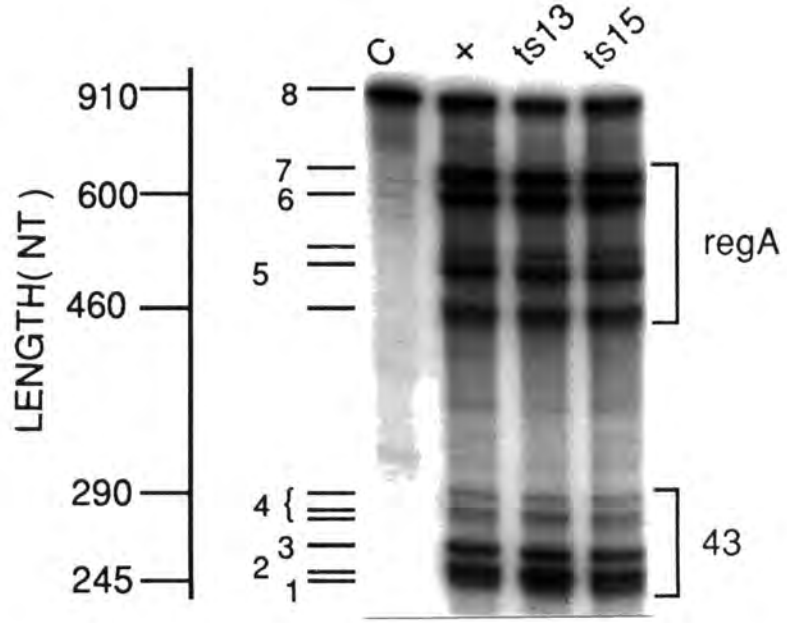
RNA analyses showing the insensitivity of T4 gene *43* mRNA synthesis to mutations in the structural gene. The drawing on top depicts the T4 genomic segment for which the analysis were carried out. The autoradiograms show the results of RNA quantitations by S1-nuclease analysis of RNA-RNA hybrids with 43 mRNA from amber mutant infections (Amber Mutants panel) and temperature-sensitive mutant infections (*ts* Mutants panel). The RNA used for S1 mapping was from infections with T4 *43amE4311-44amN82* (*E4311* lane), *43amE4301-44amN82*(*E4301* lane), *44amN82* (*43+* lane). The RNA for *ts* mutants was from infections with T4 *43tsYN13-44amN82 and 43tsYN15-44amN82* (43^- lanes) and *44amN82* (43^+ lanes). The RNAs were isolated at 15 min postinfection. The sizes and polarities of the riboprobes used for S1 mapping are shown in the drawing. The control (C lane) in the S1 mapping experiment was an S1-digested RNA-RNA hybrid that was prepared by annealing the two complimentary *in vitro* transcriptional products (S1 riboprobes) of the DNA cloned between the SP6 and T7 promoters of the pGEM4 riboprobe vector, *i.e.* it represents the size of full-length RNA spanning the region of interest. The M lane contained size markers obtained by ^{32}P-end labeling of HaeIII digested ϕX174 RF DNA. The bands numbered in the *ts* Mutant panel correspond to the mRNA species diagrammed above.



am Mutants



ts Mutants

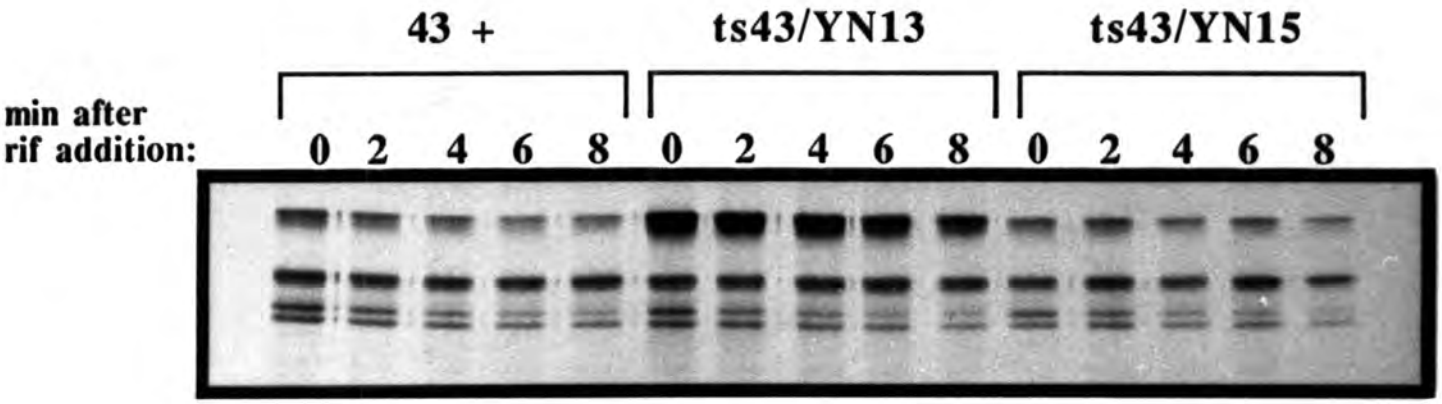


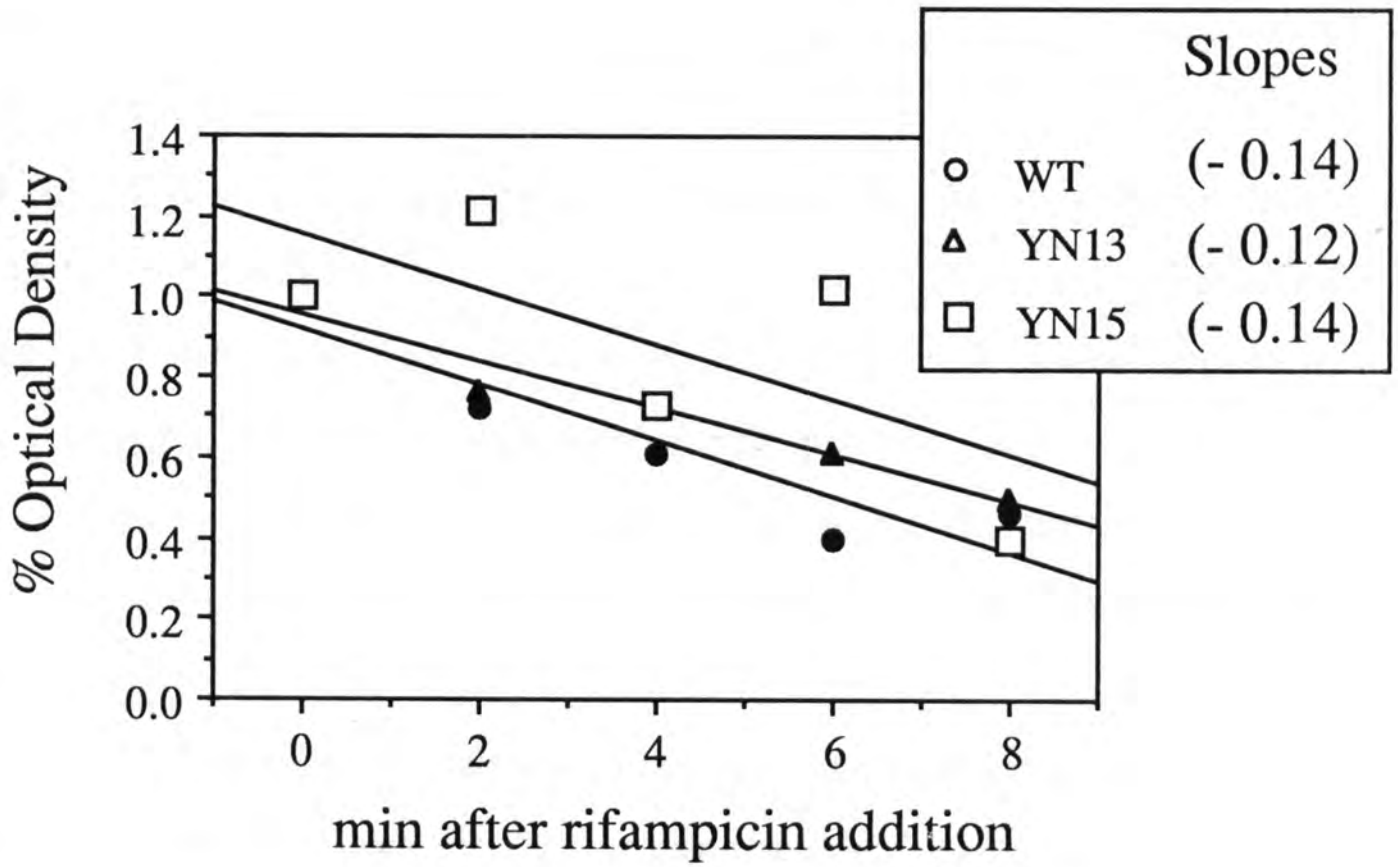
4) RATE OF DECAY OF WILD-TYPE AND MUTANT GENE 43 MRNAS

In order to rule out the possibility that alterations in protein expression were caused by a differential stabilization of mutant gene *43* mRNA, I measured the rate of functional decay of gp43 mRNA in vivo, for each of the *ts* 43 infections shown in Figure 12. Cells were infected with phage and after 12 minutes they were treated with rifampicin to block transcription. The infected cultures were then pulse-labeled with ^{35}S-Met for overlapping 5-min intervals, and proteins extracted and analyzed on SDS polyacrylamide gels. The amount of gp43 synthesized in a given time pulse is proportional to the amount of 43 mRNA present; therefore, several pulses in a time course experiment reveal the rate of decay. Figure 13 compares results of this analysis for some *ts* gene 43 mutants to wild-type. The mutation in *ts*YN13 is a single base change of G -> A in the first position of codon 26, which results in an amino acid substitution of Lys for Glu. The tsYN15 mutation is a C -> T change in the first base of the 35th codon, resulting in a change from Pro to Ser. It is unlikely that a single base changes in a long mRNA would affect stability. Although protein synthesis of *ts*YN13 is derepressed (same as amber fragment polypeptides), the rate of decay remains the same between gene 43^+ and gene 43^- infections. I concluded that the protein derepression in these mutants was not due to increased transcription or stability of mRNA.

Figure 13: Gene 43 mRNA decay in *ts* 43 mutants

Analysis of mRNA rate of decay in wild-type and 43*ts* mutants. Cells were infected with phage at m.o.i.=10 and 12 min later were treated with rifampicin to block transcription. Samples were pulse-labeled with ^{35}S-Met for 5 min intervals starting at the specified time points after rifampicin addition. Incorporation of label was quenched by adding a cold chase, chilling cells on ice, and proteins were extracted and analyzed on SDS polyacrylamide gels. Numbers above each lane are minutes post-rifampicin addition. The stability of a given mRNA can be inferred by examining the rate of decay of labeled proteins. The graph below the autoradiogram plots the percent optical density (time 0 = 100%) against the min after rifampicin addition. The panel shows a similar rate of decay for both wild-type and ts mRNA's.


43 +
ts43/YN13
ts43/YN15
min after
rif addition:
0 2 4 6 8 0 2 4 6 8 0 2 4 6 8


Slopes
WT (- 0.14)
YN13 (- 0.12)
YN15 (- 0.14)
% Optical Density
1.4
1.2
1.0
0.8
0.6
0.4
0.2
0.0
0
2
4
6
8
min after rifampicin addition

5) TRANSLATIONAL REPRESSION OF GP43 SYNTHESIS *IN VITRO*

a) DNA directed gp43 synthesis in an S30 system

In order to directly test the effects of T4 DNA polymerase on its own translation, I analyzed the effects of added protein on gp43 synthesis in cell-free extracts. Having the pre-requisites of purified gp43 and plasmid constructs available for study, several questions regarding gp43 repression could be addressed by this approach. The aim of these experiments was to reconstruct the regulation of gp43 biosynthesis *in vitro*. First step was to optimize this expression system and characterize gene expression from our plasmid DNA clones.

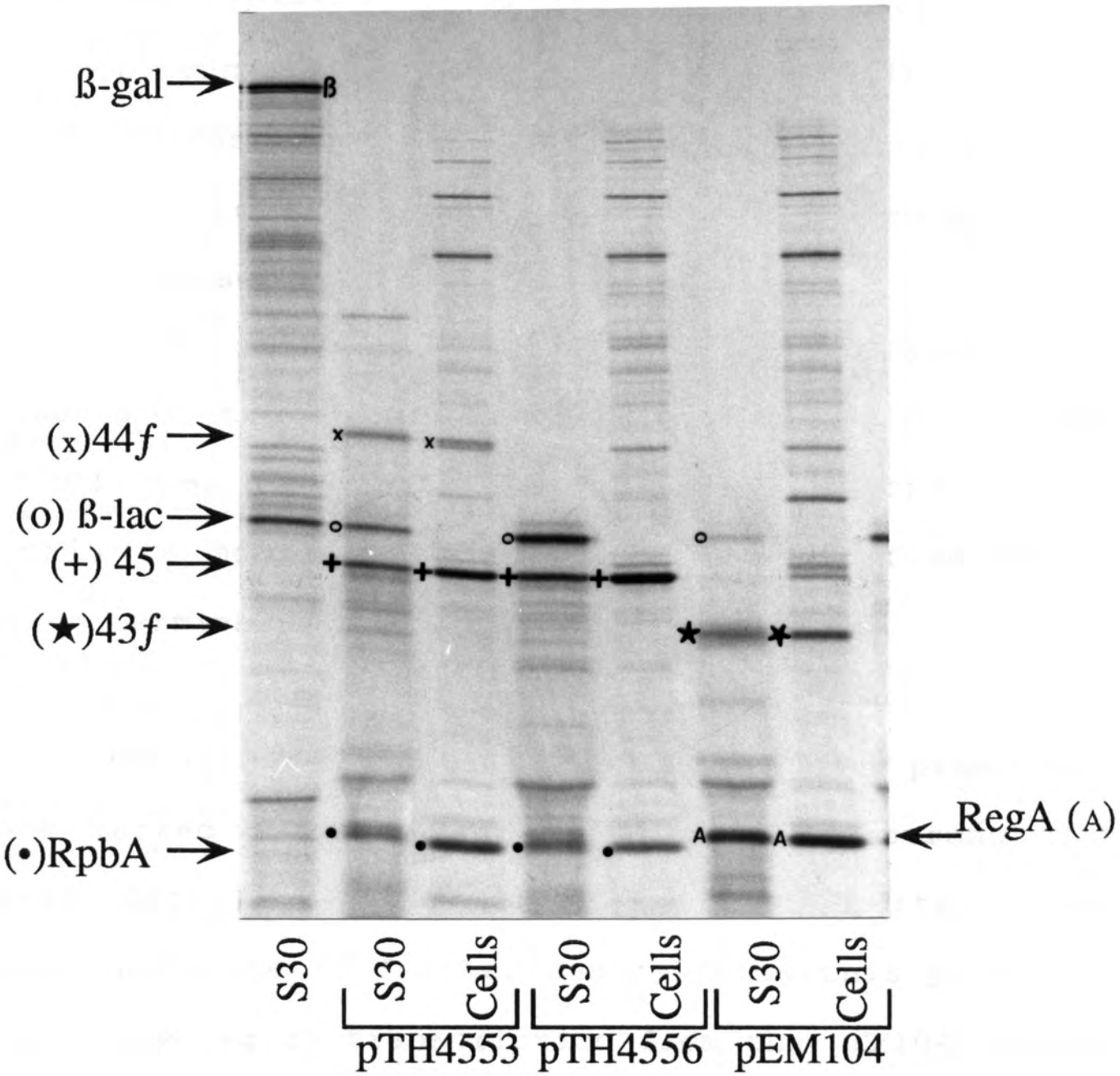
Figure 14 compares the synthesis of labeled proteins from this *in vitro* S30 system to the labeled proteins extracted from induced cells harboring the same plasmid clone. Several observations are worth noting:

1) The major products of each T4 DNA fragment cloned downstream of the P_L promoter are synthesized correctly both *in vivo* and *in vitro*.

2) The ß-lactamase (ampicillin resistance protein for each plasmid), which is transcribed from an *E. coli* promoter, is synthesized at a much higher level *in vitro*; there is little promoter competition for the transcriptional machinery *in vitro* relative to the cellular situation, *in vivo*.

Figure 14: Plasmid expression *in vivo* vs. *in vitro*

Comparison of gene expression: plasmid expression *in vivo* vs expression, *in vitro*. Lanes labeled "S30" are protein extracts from the *in vitro* S30 expression system, and lanes labeled "Cells" are cellular protein extracts from induced cells (see MATERIALS AND METHODS). Various plasmid constructs, encoding different T4 genes from this cluster of replication proteins, were tested. The positions of specific proteins are marked on the left and each protein given a code symbol.


ß-gal
(x)44f
(o) ß-lac
(+) 45
(★)43f
(•)RpbA
RegA (A)
S30
S30
Cells
S30
Cells
S30
Cells
pTH4553
pTH4556
pEM104

3) The synthesis of large proteins *in vitro* may result in the presence of several premature translation termination products (see pGAL85 lane of Figure 14). Perhaps translation processivity of the *in vitro* machinery has been compromised due to isolation procedures; nevertheless, some full length ß-galactosidase is made.

4) There may be some peptide size heterogeneity in the synthesis of the amino terminal gp43 fragment synthesized from pEM104 (see Figure 14 compare lanes S30 and Cells). Again, this is most likely due to a lack of fidelity in translation termination.

As previously mentioned, weak or cryptic promoters, which are masked *in vivo* by competition with strong host promoters, often show strong activity *in vitro*. The experiment in Figure 15 further demonstrates this point. I wanted to characterize transcription from the pEM104 plasmid and see if expression of T4 genes *regA* and *43* were dependent on the heterologous P_L promoter. Figure 16 diagrams in detail the plasmid pEM104 including relevant restriction enzyme sites. Digestion with *Bam*H1 isolates T4 sequences from the P_L promoter and the control digestion with *Sca*1 cuts in the middle of the ß-lactamase. Plasmids were thoroughly digested and subsequently assayed in the S30 *in vitro* system. Unexpectedly, transcription of the T4 genes was not dependent

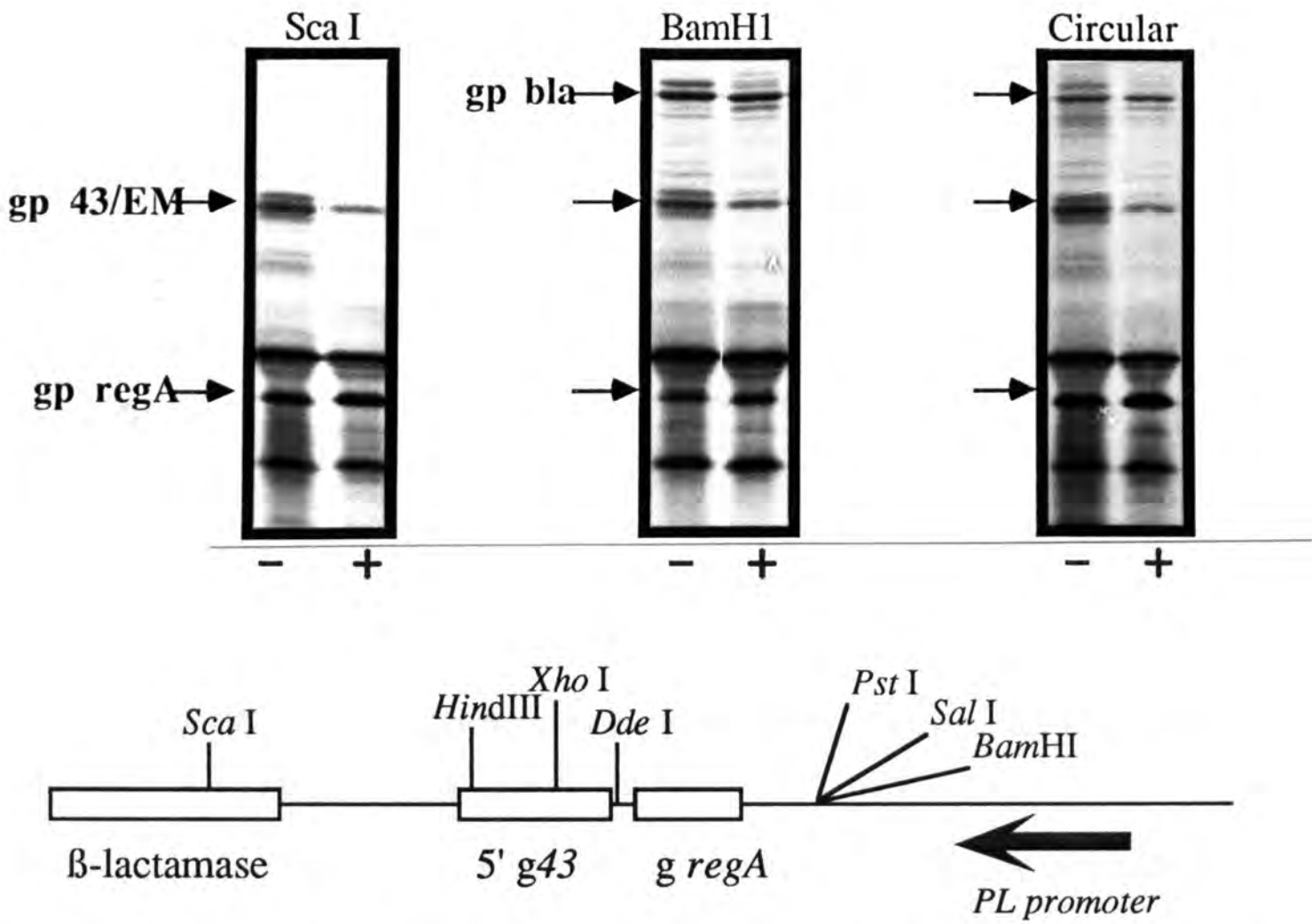


Figure 15: Gene 43 transcription is not dependent on the exogenous lambda promoter

Plasmid DNA, either digested or circular, was used to direct protein synthesis in the coupled S30 *in vitro* transcription/translation system. The positions of gpregA and the 43 N-terminal fragment are marked by arrows. The + or - denote the presence or absence of purified T4 DNA polymerase in the assay for that lane.

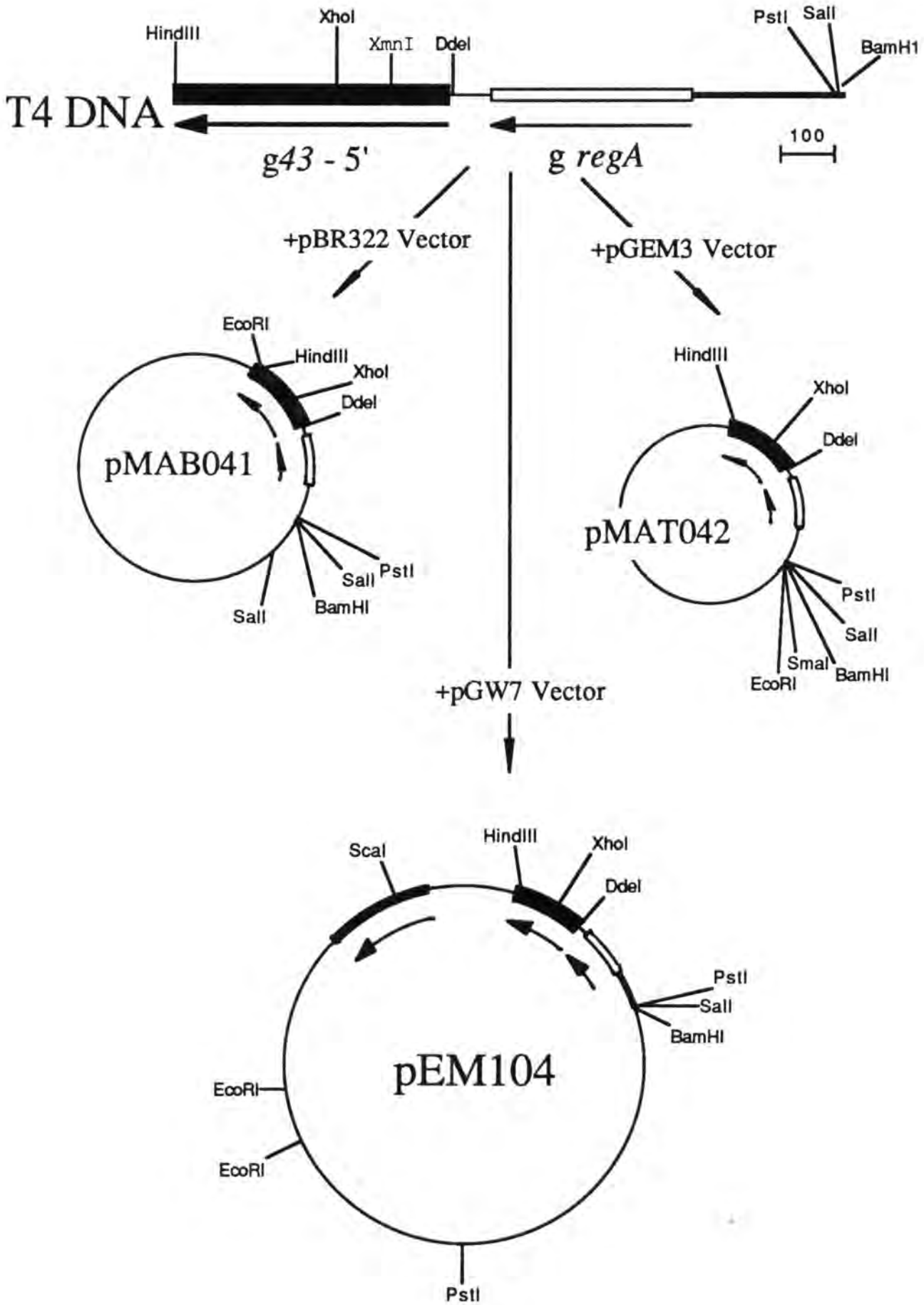
on P_L (*Bam*H1 lanes Figure 15). Most likely, efficient *in vitro* transcription initiated at a cryptic promoter in T4 sequences upstream of *regA*, which is not utilized effectively *in vivo*. Previous experiments have indicated the possible existence of an inefficient *regA* promoter (Hsu and Karam, 1990). The control digestion of pEM104 with *Sca*1 resulted in the loss of expression of the ß-lactamase, as expected. Also, the addition of purified gp43 specifically inhibited the expression of the amino terminal gp43 fragment (see more detailed studies below), irrespective of the promoter utilized. This was another indication that regulation of gp43 synthesis does not occur at the level of transcription.

b) DNA directed expression from different plasmid clones

Figure 16 diagrams several plasmid clones that were constructed by insertion of the T4 fragment encoding the C-terminal portion of gp62, gp regA, and the N-terminal portion of gp43 into various cloning vectors. Due to the presence of differing vector sequences downstream of the T4 gene *43* fragment, various lengths of read-through peptides were synthesized from these plasmids, *in vitro* (Figure 17 - gp 43 fragment noted with a dot). This allowed us to definitively assign which protein bands represented the N-terminal gp43 fragment. In addition, note that the synthesis of each gp43 fragment was repressed by the addition of purified T4 DNA polymerase, indicating that the determinants for repression

Figure 16 Clones of T4 *regA-43* segment in various plasmid vectors.

The fragment at the top portrays the genetic segment cloned in the various expression plasmids shown below. Relevant restriction sites and gene boundaries are shown. The *Bam* H1 to *Hin*d III fragment was inserted into the plasmids as diagrammed below. Restriction sites important to the work of this dissertation are shown.


HindIII
XhoI
XmnI
DdeI
PstI
SalI
BamH1
T4 DNA
100
g43 - 5'
g regA
+pBR322 Vector
+pGEM3 Vector
EcoRI
HindIII
XhoI
DdeI
pMAB041
SalI
PstI
SalI
BamHI
HindIII
XhoI
DdeI
pMAT042
PstI
SalI
SmaI
BamHI
EcoRI
+pGW7 Vector
ScaI
HindIII
XhoI
DdeI
PstI
SalI
BamHI
pEM104
EcoRI
EcoRI
PstI

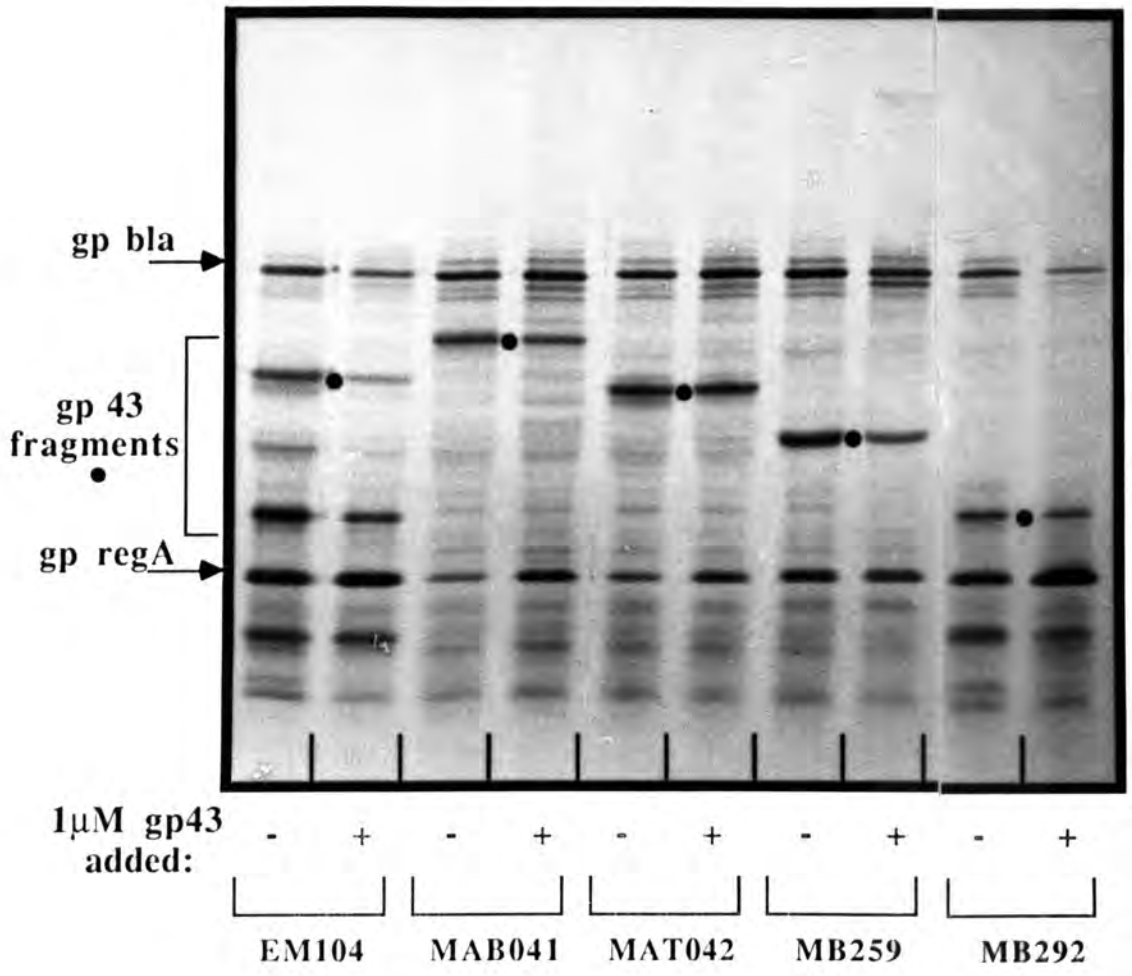


Figure 17: Expression of 43 N-terminal peptides, *in vitro*

Plasmid clones, diagrammed in Figure 7, were used in coupled S30 expression system, and the protein products analyzed on SDS-PAGE. The positions of relevant proteins are marked on the left, and the plasmid used below each lane. The + or - denote the presence or absence of purified T4 DNA polymerase is the assay for that lane. Dots mark the position of various gp43 N-terminal fragments.

pMB259 was constructed by Melanie Hughes by ligating T4 DNA segment (*Bam*H1-*Hin*dIII) into the pBH20 vector; and pMB292 inserts the *Bam* H1 - *Xho* I fragment into a pGW7 derivative and was also constructed by Melanie Hughes (Ph.D. Thesis).

most likely reside in the sequences common to all of these constructs, and also do not depend on mRNA sequences transcribed from the 3′ half of gene *43* (not encoded by these clones). The experiments above demonstrate the ability of purified gp43 to repress the synthesis of N-terminal gp43 fragments, but do not distinguish the level at which that regulation takes place. To determine if T4 DNA polymerase was able to repress translation of its message, I utilized these same S30 extracts to examine the regulation of translation *in vitro*.

c) RNA directed synthesis of gp43 in an S30 system

The heat-inducible lambda $cI^{857}P_LN$ expression plasmid pEM104 (Trojanowska, *et al.*, 1984), which is diagrammed in Figure 16, was placed in *E. coli* NapIV and used to produce a polycistronic mRNA capable of synthesizing T4 gp*regA* and an amino-terminal T4 gp*43* fragment (designated $gp43_{JM}$). Identity of the pEM104-generated polycistronic RNA was verified by RNA-RNA hybridization and primer extension assays similar to those described in Figure 12. The purified RNA was subsequently used with S-30 cell-free extracts in translation assays.

In order to be certain that *in vitro* biosynthesis of proteins arose solely from translation of exogenously added polycistronic mRNA, I tested the sensitivity of protein synthesis to inhibitors of transcription.

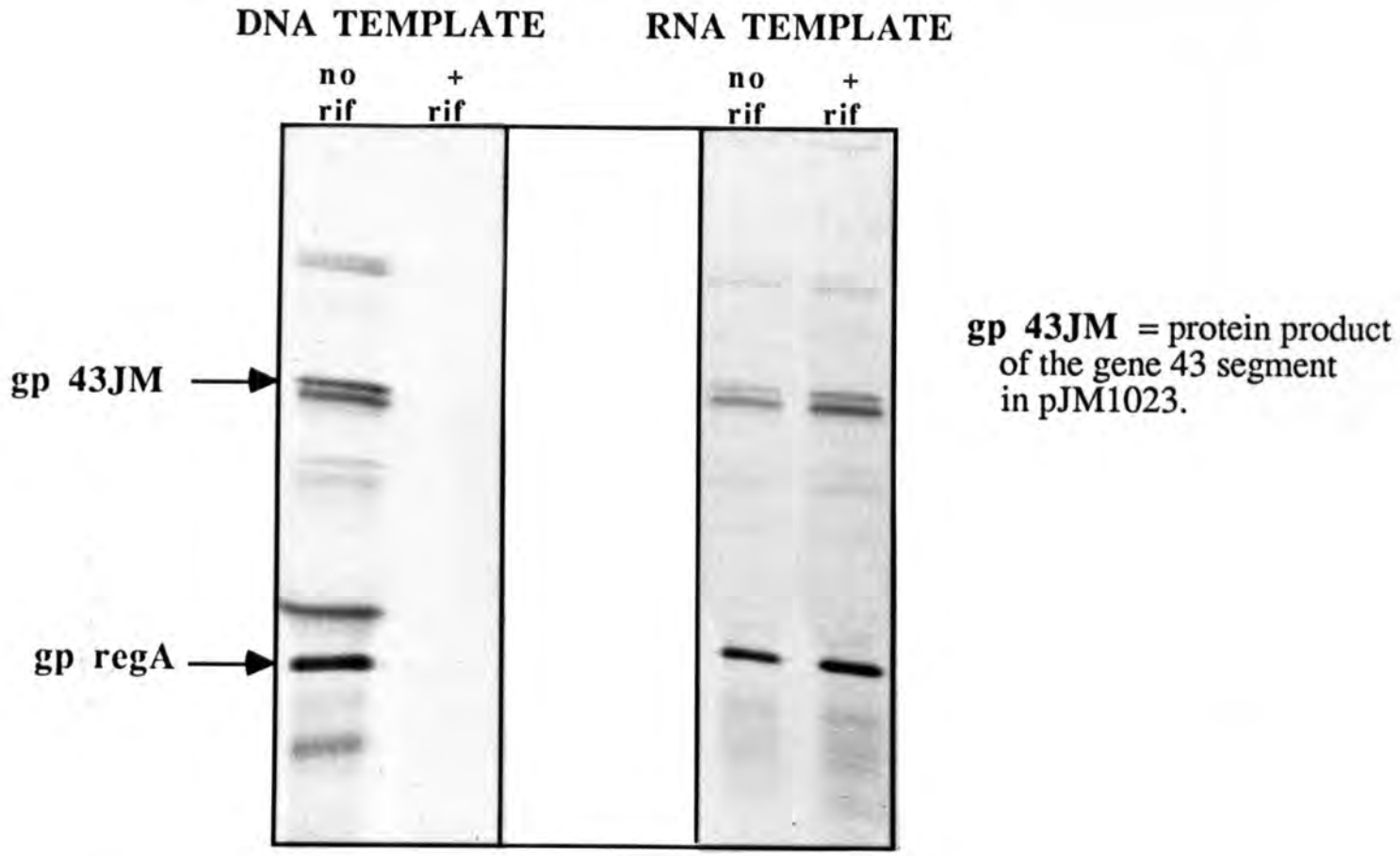


Figure 18: Translation *in vitro* is insensitive to rifampicin

RNA- and DNA-directed syntheses were tested for sensitivity to rifampicin as described in MATERIALS and METHODS. The + or - denotes the presence or absence of rifampicin in the assay. The plasmid DNA and source of RNA were pJM1023, which harbors a wild-type gpregA, in contrast to pEM104. The positions of the $gp43_{JM}$ and gp regA are noted on the left of the gel.

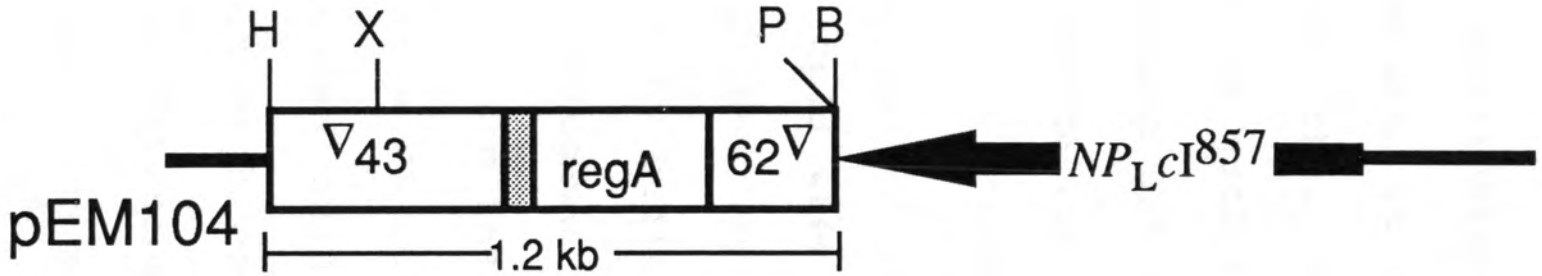
Expression of proteins in this RNA directed S30 system was resistant to the presence of rifampicin, whereas, synthesis in a DNA-directed assay was completely halted by rifampicin (Figure 18). I concluded that this method of RNA preparation contained no DNA contamination which could induce protein expression, and that the cellular extracts were adequate for assaying regulation of translation *in vitro*.

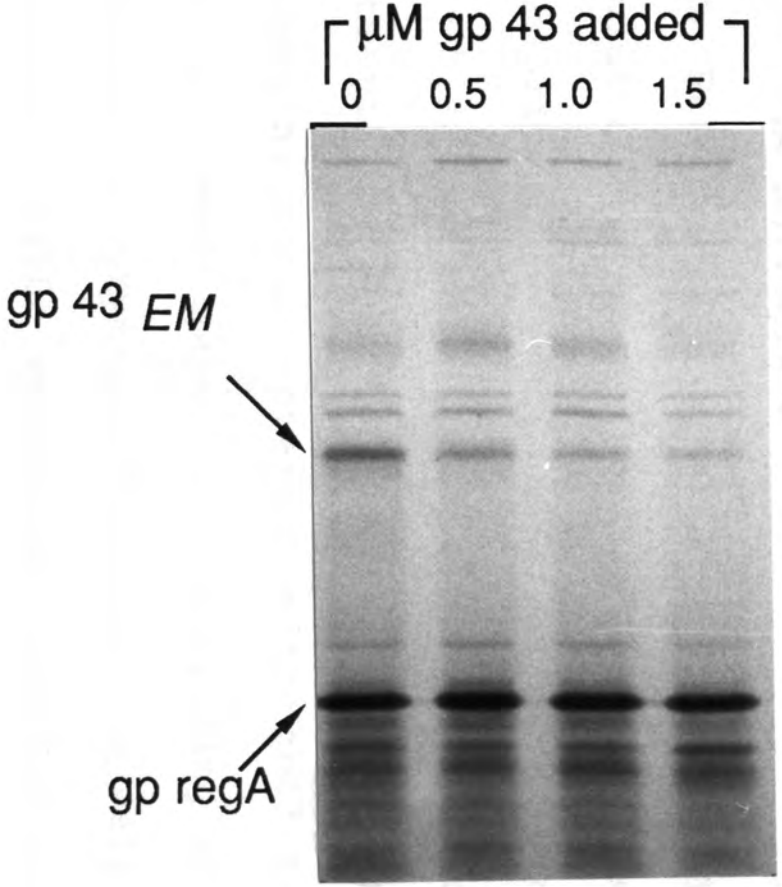
d) Gp43 translation is specifically inhibited by purified T4 DNA polymerase, in vitro.

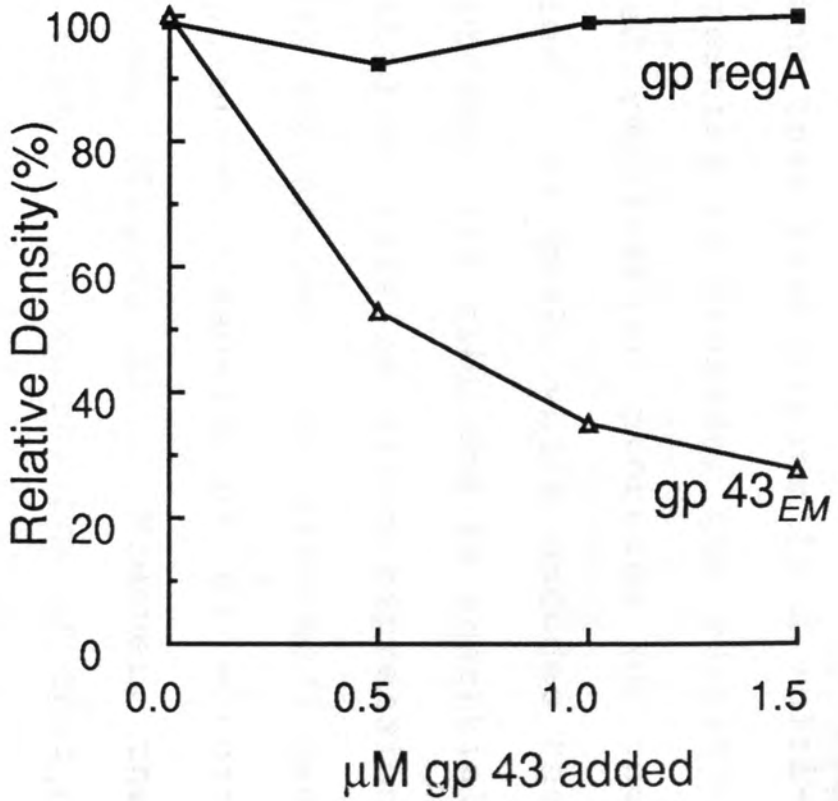
The purified RNA was subsequently used in cell-free translation assays that measured the effects of purified wild-type T4 DNA polymerase addition on the production of the two pEM104-encoded T4 proteins. The added enzyme inhibited $gp43_{EM}$ synthesis in a dose-dependent fashion, yet did not affect *gpregA* synthesis from the plasmid-generated polycistronic mRNA (Figure 19). These results demonstrate that T4 DNA polymerase is a repressor of its own translation and that it binds to an mRNA site which is located in a region distal to the *regA* cistron since *regA* expression is not affected by T4 DNA polymerase either *in vivo* (Karam and Bowles, 1974; Karam, *et al.*, 1977) or *in vitro* (Figure 19).

Figure 19: Translational repression by T4 DNA polymerase *in vitro*.

Cell-free extracts from *E. coli* MRE600 were used to translate an RNA mixture enriched for polycistronic mRNA from heat-induced pEM104 (diagrammed at top). The pEM104-generated mRNA encodes *RegA* protein (gp*regA* and an amino-terminal gene *43* protein fragment (gp43_{EM}). Translations were carried out in the absence and in the presence of 0.5, 1.0, and 1.5μM purified wild-type T4 DNA polymerase and ^{35}S-Methionine, and were subsequently analyzed by SDS-gel electrophoresis and autoradiography (see MATERIALS AND METHODS). Densitometric scans of the autoradiogram shown on the left were used to prepare the plots on the right. Note the inhibition of gp43_{EM} synthesis relative to gp*regA* synthesis. At concentrations greater than 1.5μM, T4 DNA polymerase exhibited inhibitory effects on the translation of all *in vitro* products.


H
X
P
B
∇43
regA
62∇
NPLcI857
pEM104
1.2 kb


μM gp 43 added
0
0.5
1.0
1.5
gp 43 EM
gp regA


100
80
60
40
20
0
Relative Density(%)
gp regA
gp 43EM
0.0
0.5
1.0
1.5
μM gp 43 added

e) Role of other replication proteins in repression of gp43 translation

Given the need to coordinately regulate the dosage of different replication proteins that participate in a multi-enzyme complex, it was interesting to consider the possible effects of other T4 encoded replication proteins on the regulation of gp43 translation. T4 gp32, which encodes the single-stranded DNA binding protein (T4 *ssb*) and is essential for DNA replication, was tested in this *in vitro* expression system for its effects on gp43 expression. Purified gp32 had no effect on the transcription or translation of either gpregA or gp43 when added alone (Figure 20). However, the addition of gp32 appeared to enhance the repression of $gp43_{JM}$ translation by purified T4 DNA polymerase. Protein bands on the autoradiogram from Figure 20 were quantitated by densitometry and the results are shown in the graph of Figure 21. The role of other replication accessory proteins could not be tested because of the unavailability of purified proteins in high concentration. It is possible that other replication accessory proteins could interact with gp43 to enhance repression at the gp43 mRNA operator; the binding and cooperation of these proteins in enzymatic functions is well-characterized (Jarvis, *et al.*, 1990; Venkatesan and Nossal, 1982); binding and cooperation for regulatory roles could also be feasible, especially in light of the need for coordinate regulation of gene expression.

Figure 20 and **21**: Effects of gp32 on the autogenous repression of T4 DNA polymerase

DNA and RNA directed in vitro expression from pJM1023 was assayed with the S30 system, proteins were extracted and analyzed by SDS-PAGE, and the dried gels subjected to autoradiography and quantitation by densitometry. The positions of relevant protein bands are marked on the left, and the additions of protein noted below each lane. The relative amounts of proteins synthesized are graphed in Figure 21.

Figure 20

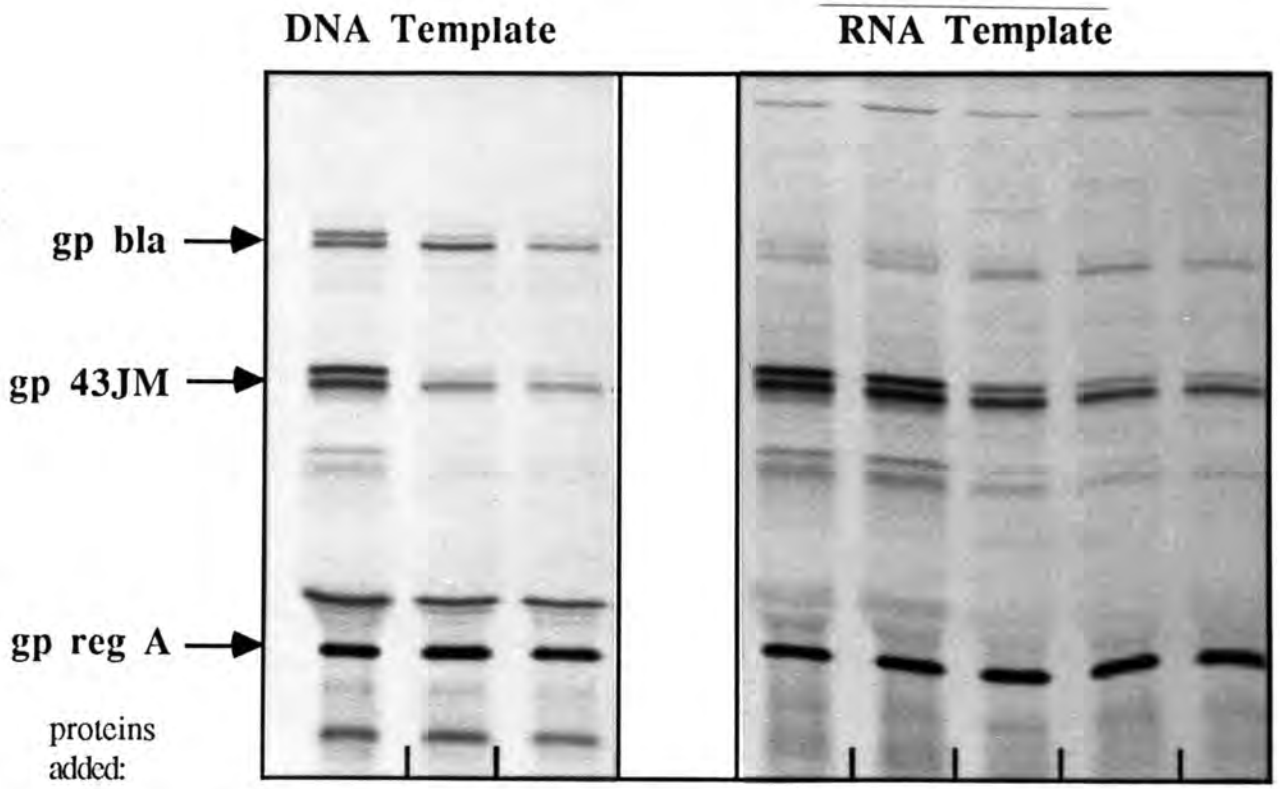
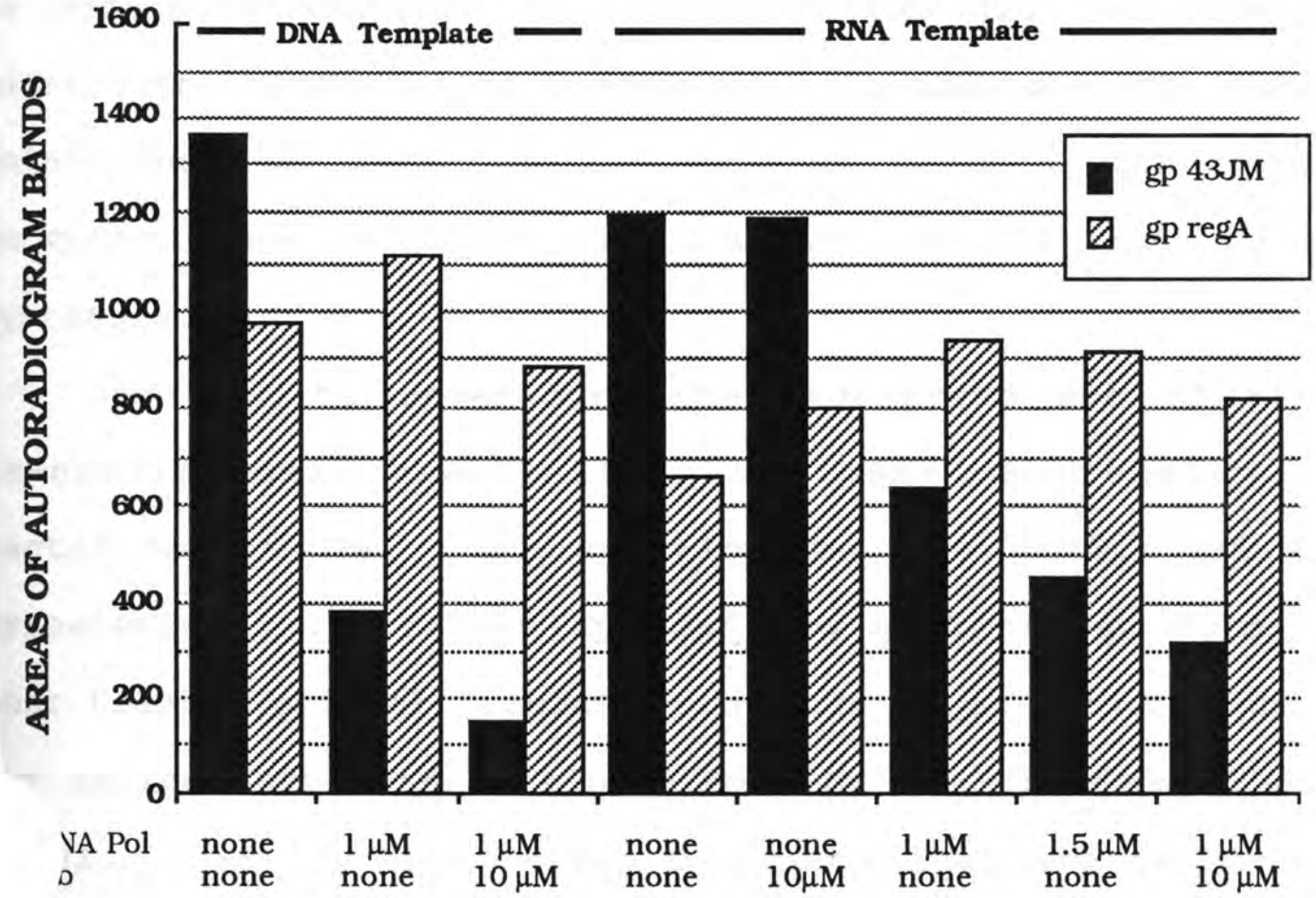


Figure 21



6) DISCUSSION

I have shown that autogenous repression of gp43 biosynthesis occurs at the level of translation. De-repression of protein synthesis in structural gene 43 mutants is not due to increased transcription or stability of the mRNA. Purified DNA polymerase specifically inhibits translation of gene *43* mRNAs *in vitro*.

Early experiments were cited as indicating a transcriptional mechanism for this feedback repression.(Krisch, *et al.*, 1977; Miller, *et al.*, 1981). In fact, these experiments only demonstrated that gene *43* mRNA is generally unstable in infection, and did not directly address the mechanism of repression. In addition, the mutant infections that were compared were not matched with respect to other phage mutations that involve transcription and DNA synthesis.

Our results underscore the widespread use of post-transcriptional regulation of gene expression by bacteriophage T4. In addition to the action of the translational repressors gp regA and gp32, other forms of post transcriptional regulation include 1) an intron that is not spliced but rather skipped over by the ribosome (Huang, *et al.*, 1988),2) three group I introns that are capable of self-splicing activity, (Shub, 1988) and, 3) a T4 induced endonuclease that cuts mRNAs at a specific site near the

Shine-Dalgarno complementarity in the ribosome binding site of several genes (Ruckman, *et al.*, 1989).

It is interesting to consider the possible roles of translational regulation in T4 infection. While regulation of transcription is primarily responsible for the temporal program of infection, the most likely purpose of translational repression is to fine tune the final dosage of certain gene products. Some might consider this dosage regulation superfluous; however, it is difficult to ascertain the importance of translational regulation when phage production is examined under optimal growth conditions in the lab, in a host we are not sure is the natural target for T4 infection. "Non-essential" genes and various modes of regulation may play a more essential role in other hosts, more adverse growth conditions, or in the presence of other genetic elements that may exclude T4 infection. The classic example of this latter idea is the exclusion of T4 *rII* mutants (a "non-essential" gene) from infection in *E. coli* strains which also harbor a lysogen of phage lambda. The presence of other genetic elements, such as insertion sequences, transposons, etc. may confer a selective advantage to the phage which regulates the dosage of replication components. In any case, several of these translational regulatory mechanisms have been maintained in a variety of prokaryotic organisms; their conservation is indicative of the importance of this mode of regulation.

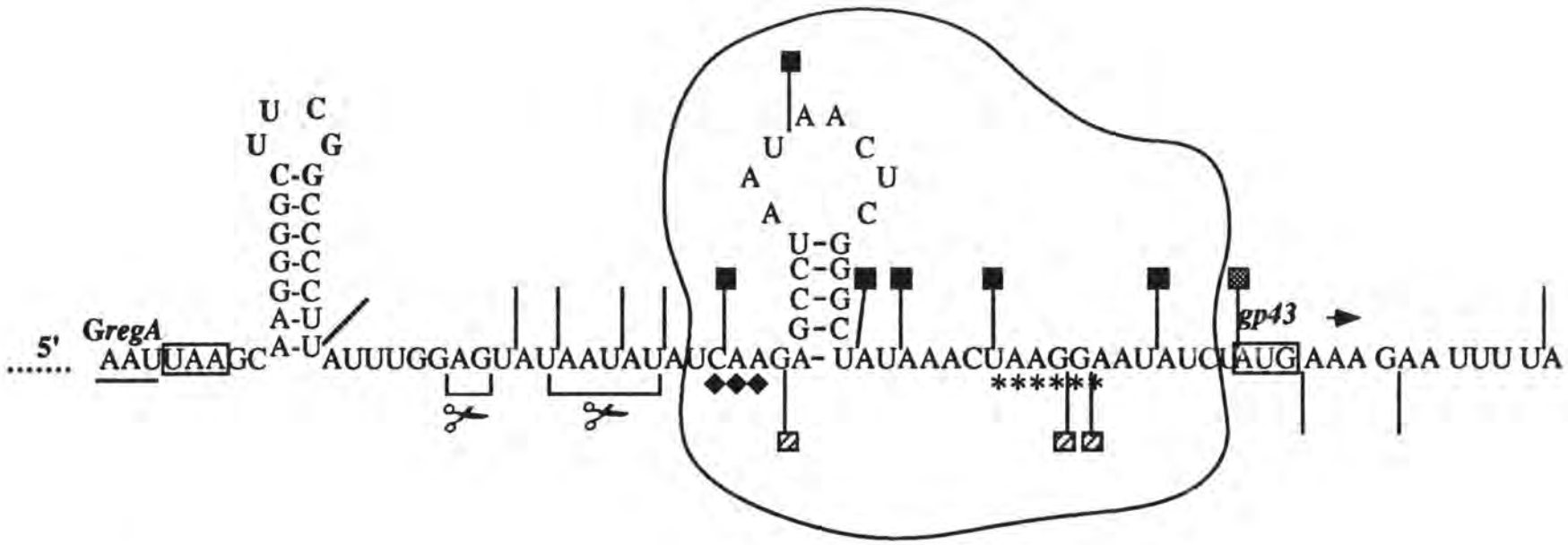
Another analogy relates this dose regulation to the lysis/lysogeny decision made by other lysogenic bacteriophage. Perhaps limiting the amount of DNA replication and burst size contributes some advantage to phage propagation and survival.

The advantage of autogenous regulation at the level of translation remains unknown. Perhaps a separation of the protein domains essential to its regulatory function (*i.e.* RNA binding domains) from those domains involved in the enzymatic function of DNA replication confers some advantage. T4 DNA polymerase, as a multi-functional protein must possess active sites responsible for protein, substrate, DNA, and RNA binding aspects of its function. However, it is yet to be determined whether these protein domains are separate and discrete or overlapping. Perhaps crystallographic investigations will soon clarify these structure-function questions as they relate to regulation.

The most common form of translational repression involves the binding of a repressor protein to the leader sequence of its target mRNA, blocking the RBS and inhibiting the initiation of translation. Such binding was clearly demonstrated by our collaborators using nuclease protection assays (Andrake, *et al.*, 1988). Figure 22 summarizes the mRNA residues that are protected from RNase digestion by gp43 binding.

Figure 22: RNA binding site for T4 DNA polymerase

The mRNA sequence of the intercistronic region between genes *regA* and *43* is presented, including a putative stem-loop hairpin. Symbols are explained on the left. Places of RNase digestion are denoted as vertical slashes (RNase A above the sequence, and RNase T1 below). The termination and initiation codons are boxed, and symbols noting the 5′ ends of various gene 43 mRNAs are shown. All mRNAs contain the entire region where the protein binds, shown diagrammatically by the shaded area.



RNase protection symbols

- ■ indicates protection from RNase A digestion (vertical slashes above the sequence) by gp43
- ▩ indicates partial protection
- ▨ indicates protection from RNase T1 digestion (vertical slashes below the sequence) by gp43

transcription/translation symbols

- ◆ indicates 5' ends of mRNAs that are transcribed from the phage dependant promoter in this region
- ✂ indicates 5' ends of mRNAs that are processed from polycistronic messages.
- *** marks the Shine/Dalgarno complementarity

With a known region where the protein contacts its mRNA, termed the *mRNA operator*, several questions now arise regarding the molecular determinants necessary for protein-operator binding and repression of translation: 1) Given the fact that repressor proteins must recognize and bind mRNAs which possess a unique three-dimensional RNA structure (including primary, secondary and higher order structure), what are the structural requirements of the gp43 mRNA operator? 2) Are particular residues essential for proper interaction of the mRNA with the repressor gp43? 3) What are the rules that govern RNA protein binding and in this case the physiological result of translational repression? These questions are the motivation for the studies described in the second part of this dissertation.

B) PART II. MUTATIONAL ANALYSIS OF THE mRNA OPERATOR FOR T4 DNA POLYMERASE

1) RATIONALE AND SIGNIFICANCE

Knowledge of a defined region, where T4 DNA polymerase contacts its mRNA leader, made a mutational analysis of the gp43 RNA operator feasible. The approach is similar to those studies done on DNA operator regions; however, since mRNA folding plays an important role in translation, a mutational analyses involves the added complexity of interpreting results in relation to RNA structure. In the case of transcription factors that bind DNA promoters, there is the benefit of comparing several binding sites to determine a primary sequence consensus. This is not the case for gp43 with a single known RNA binding site, and we are forced to look at more than primary structure in elucidating the determinants of gp43 binding; we must determine a "structural (or conformational) consensus" through our mutational analysis. This problem is exactly analogous to the discussion of protein "structure-function" relationships, and how primary sequence determines structure and ultimately function.

All three known T4-induced translational repressors contact their mRNA substrates within translation initiation domains and probably inhibit specific translation by blocking

ribosomes at these sites (Gold, 1988). On the basis of studies that mapped the mRNA-binding sites of T4 gp32 (McPheeters, *et al.* 1988), T4 gp43 (Andrake, *et al.* 1988; Tuerk, *et al.* 1990) and T4 gpregA (Karam, *et al.* 1981; Unnithan, *et al.* 1990; Webster, Adari and Spicer 1989; Winter, *et al.* 1987), it is clear that each of these proteins recognizes its own set of features in RNA sequence or structure. The precise characteristics of these features remain largely unknown, especially within the context of the three-dimensional structures of folded RNA and protein molecules. T4 gp32 utilizes RNA folding ("pseudoknot" formation), single-strandedness, and sequence specificity (Fulford and Model, 1984; McPheeters, *et al.* 1988) whereas RegA protein seems to utilize mainly nucleotide sequence elements around an initiator AUG (Unnithan, *et al.* 1990). In the case of T4 DNA polymerase, the mRNA binding site includes about 40 nucleotides to the immediate upstream of the initiator AUG for gene *43* and harbors a putative hairpin structure consisting of a 5-base-pair stem and an 8-base loop (Andrake, *et al.* 1988; Tuerk, *et al.* 1990). Evidence for this putative stem-pairing is derived from RNase protection assays (Andrake, 1988 - *i.e.* paired G residues are protected from RNase T1 digestion of RNA alone) and supported by computer assisted analysis of RNA folding (see Discussion and Conclusions Zuker, 1989, 1981). In addition, the experiments presented below confirm the existence and importance of this hairpin structure in the gp43 operator. The studies that

follow are a fine-structure mutational analysis aimed at identifying features and residues of the hairpin that are important for specificity of the gp43-mRNA interaction.

To begin these studies we designed a strategy for the construction of mutations in the operator region. This was done by *in vitro* mutagenesis using synthetic DNA duplexes, which were cloned in a T7 vector expression system (described in section 3 below). To examine the effects of these operator mutations on gene expression in vivo, we added the remaining portion of the structural gene 43; the strategy for this cloning is described in the next section. Subsequently, plasmid constructs were used 1) to synthesize labeled mRNA *in vitro*, for protein-RNA binding studies, and 2) for expression studies by inducing plasmid expression in *E. coli* strains harboring the T7 RNA polymerase gene.

2) Cloning of an autogenously regulated T4 gene *43* in a T7 expression vector:

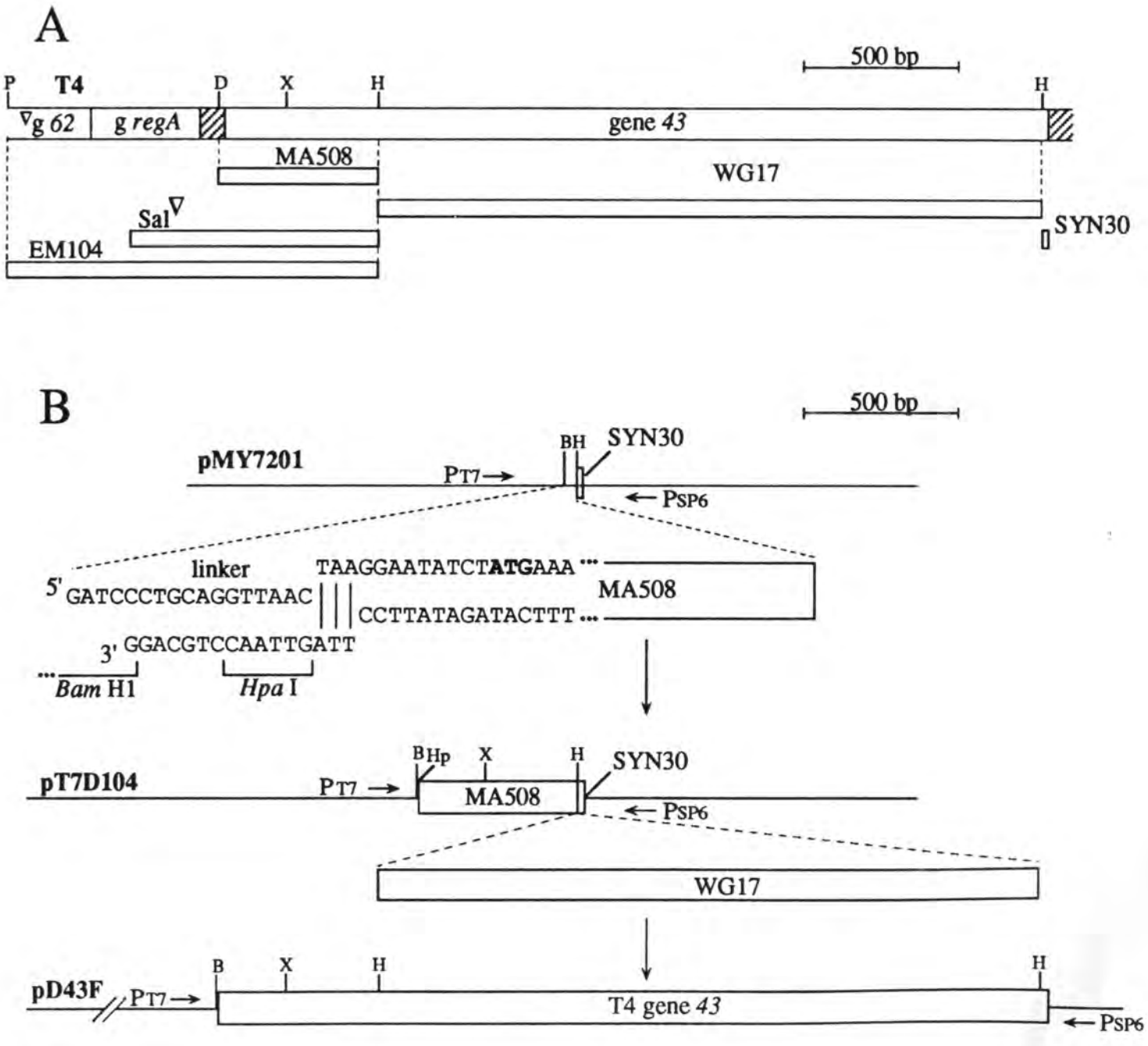
Figure 23 depicts the approach I used to clone the wild-type T4 gene *43* in the T7/SP6 expression plasmid pSP72, under control of wild-type and mutant translational operators. Previous work had shown that all but the 3'-terminal 30 base-pairs of the protein encoding portion of this gene could be retrieved from lambdoid or plasmid T4 DNA libraries (e.g. Hsu, *et al.* 1987; Hughes, *et al.* 1987; Spicer, *et al.* 1988; Trojanowska, *et al.* 1984). A synthetic DNA duplex (SYN30, Figure 23) served as a source of the 3'-terminal region of

the gene in the design of pMY7201, a cloning vector that was specifically designed for reconstructing the complete gene *43* from DNA fragments (Figure 23; M. Dawson and J. Karam, *in preparation*).

In Figure 23, the MA508 DNA used for the construction of pT7D104 originated from EM104 DNA that had been digested with endonucleases *Dde*I and *Hin*dIII; *Dde*I cleaves EM104 DNA at a site just upstream of the Shine-Dalgarno sequence for gene *43*. MA508 DNA was subsequently inserted into the *Bam*HI - *Hin*dIII interval of pMY7201, together with a synthetic DNA linker containing a *Bam*HI sticky end and a complete *Hpa*I site. This yielded pT7D104 (Figure 23), which served as the cloning vector for synthetic DNAs that specified wild-type and mutant translational operator sequences (see Figure 24 and below). The ligation of T4 DNA fragment WG17 into *Hin*dIII-digested pT7D104, or its translational-operator-bearing derivatives (see Figure 24), yielded clones in which WG17 was inserted in both possible orientations. The pD43F clone diagrammed in Figure 23 has a truncated operator (deletion of nucleotides -16 through -40 relative to the AUG) and a WG17 fragment that inserted in forward orientation (F designation), *i.e.* it is capable of encoding wild-type T4 DNA polymerase. Clones in which WG17 was inserted in the alternate orientation (reversed, R designation) specified NH_2-terminal gp43 fragments and served as controls in expression experiments that utilized *E. coli* BL21(DE3) as host for the plasmids (see Figure 28 below).

Figure 23: Cloning T4 gene *43* under the control of T7 RNA polymerase

Panel A depicts the T4 genetic region under study and shows relevant restriction sites. Open boxes denote structural genes and striped boxes denote intercistronic regions. The DNA fragments that were used for cloning gene *43* are also shown. Panel B diagrams the strategy used for cloning the gene. Experimental details are given in the text. Abbreviations: P_{T7} = T7ϕ10 promoter, P_{SP6} = SP6 promoter (horizontal arrows indicate the direction of transcription), P = *Pst* I, D = *Dde* I, X = *Xho* I, H = *Hin*d III, B = *Bam* HI.

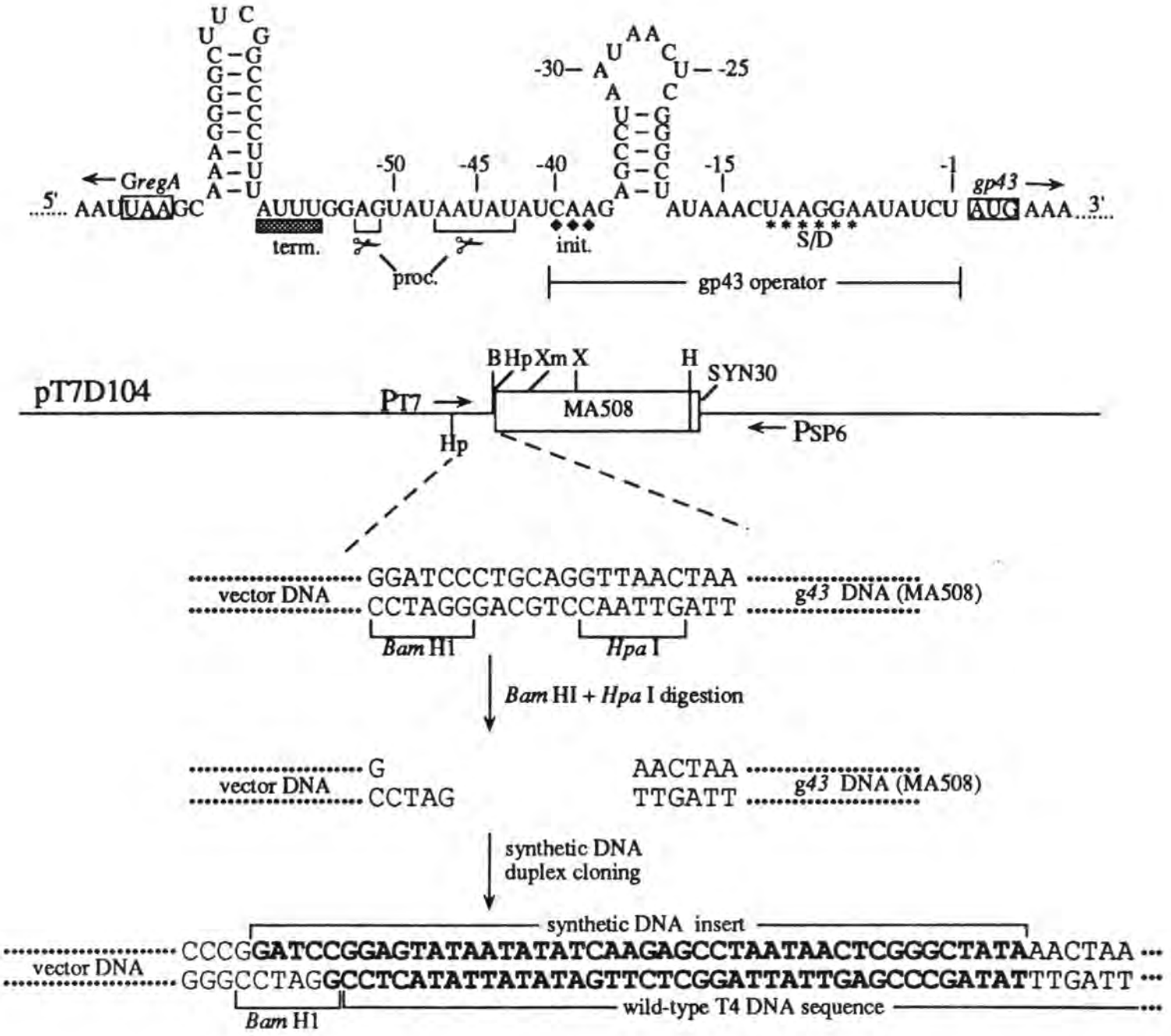

A
500 bp
P
T4
D
X
H
H
∇g 62
g regA
gene 43
MA508
WG17
Sal∇
SYN30
EM104
B
500 bp
pMY7201
BH
SYN30
PT7
PSP6
linker
TAAGGAATATCTATGAAA
5'
GATCCCTGCAGGTTAAC
MA508
CCTTATAGATACTTT
3'
GGACGTCCAATTGATT
Bam H1
Hpa I
pT7D104
BHp
X
H
SYN30
PT7
MA508
PSP6
WG17
pD43F
B
X
H
H
PT7
T4 gene 43
PSP6

3) Summary of mutant translational operators constructed

Figure 24 diagrams RNA primary and putative secondary structure in the 78 - nucleotide region between T4 genes *43* and *regA*, as derived from a number of studies that included RNase protection assays and nucleotide sequence determinations (Andrake, *et al.* 1988; Hsu and Karam 1990; Tuerk, *et al.* 1988). In T4-infected cells, gene *43* is normally transcribed by several overlapping modes, some of which initiate far to the upstream of the *regA-43* intercistronic region (Hsu and Karam 1990). Processing of readthrough mRNA and transcription initiation within this region yield a population of mRNA species with overlapping 5' ends that map at nucleotide positions -38 to -52 relative to the initiator AUG for gp43 biosynthesis (Figure 24). Thus, in such infections all gp43-encoding mRNAs contain the binding site for autogenous translational repression, *i.e.* residues -40 to approximately -1 relative to the AUG (see also Andrake, *et al.* 1988; Tuerk, *et al.* 1990). I designed a cloning strategy that allowed us to generate and test the effects of mutations in the -40 to -16 region of this translational operator. The -15 to -1 segment, which contains the Shine-Dalgarno sequence, was left intact in order to avoid major additional effects on intrinsic ribosome binding efficiency at this translational initiation domain. In our cloning scheme (Figure 24 - bottom), pT7D104 DNA was used to insert synthetic DNA duplexes containing the desired

Figure 24: Cloning synthetic operators for gp43 translation.

The upper portion of the figure shows the mRNA sequence of the *regA-43* intercistronic region, including putative RNA hairpin structures. The Shine-Dalgarno complementary sequence is marked with asterisks, and termination and initiation codons are boxed. Nucleotides protected from RNase digestion by gp43 are bracketed and define the gp43 operator (Andrake *et.al*, 1988). Diamonds denote 5' ends of transcripts that arise from initiation at the MotA-dependent promoter, scissors denote 5' ends of transcripts that arise from RNA processing, and the shaded bar marks bases at which transcription from upstream cistrons is terminated (Hsu and Karam, 1990). Nucleotides are numbered relative to the gene *43* initiation codon. The lower portion of the figure diagrams the strategy used for inserting synthetic DNA duplexes that specify gp43 translational operators. The pT7D104 recombinant plasmid (see also Figure 14) was digested partially with *Hpa* I and completely with *Bam* HI. Synthetic DNA duplexes, each possessing a *Bam* HI overhang, were then directionally inserted into the double-digested pT7D104. In the example shown (*i.e.* the wild-type T4 operator sequence), bases originating from the synthetic DNA are bolded. Abbreviations: Hp = *Hpa* I, Xm = *Xmn* I, B = *Bam* HI, X = *Xho* I, H = *Hin*d III.

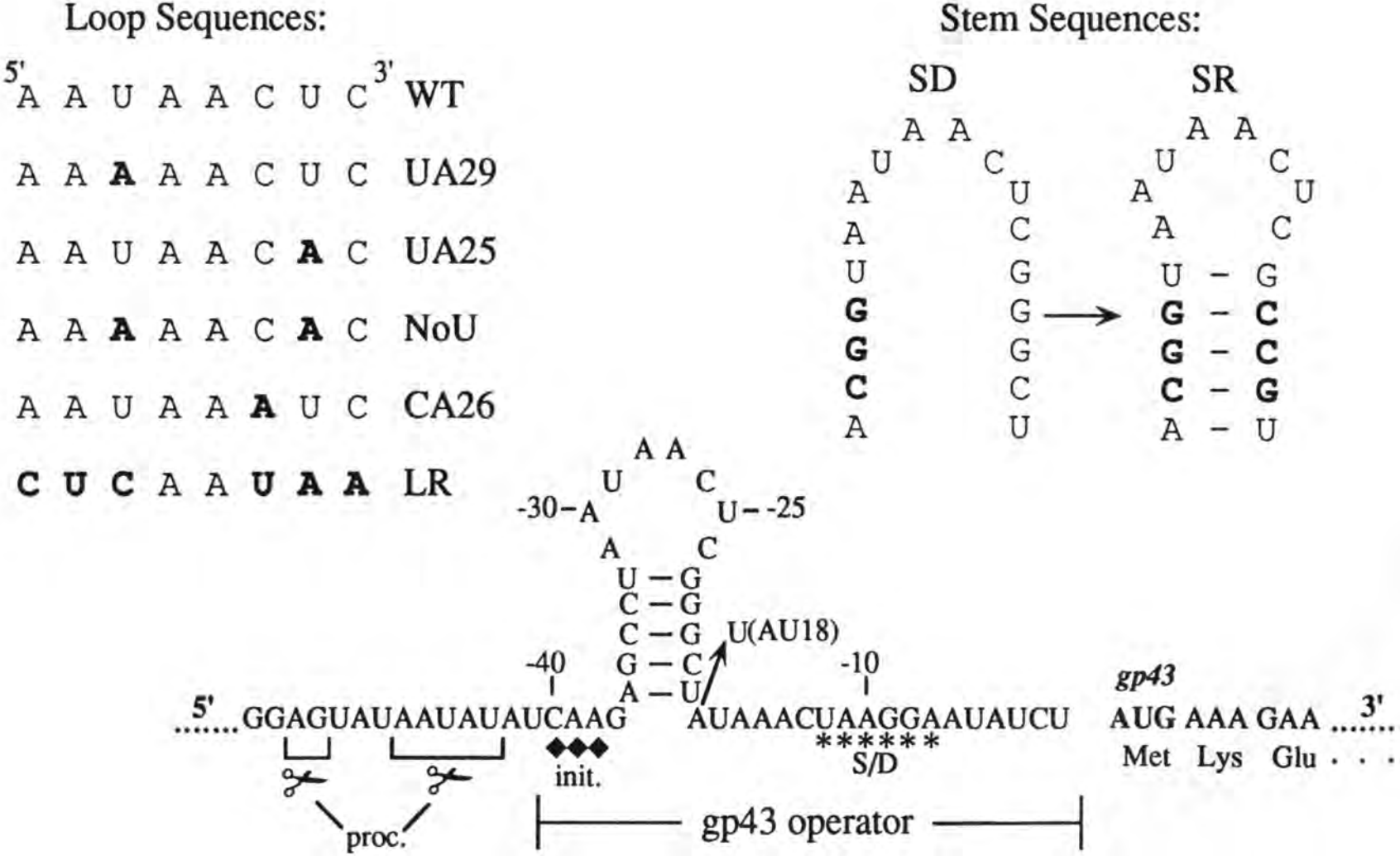

GregA
gp43
-50
-45
-40
-30
-25
-15
-1
5'
AAUUAAGC
AUUUGGAGUAUAAUAUAUCAAG
AUAAACUAAGGAAUAUCU
AUG
AAA
3'
term.
proc.
init.
S/D
gp43 operator
pT7D104
PT7
BHpXmX
H
SYN30
MA508
PSP6
Hp
GGATCCCTGCAGGTTAACTAA
CCTAGGGACGTCCAATTGATT
vector DNA
g43 DNA (MA508)
Bam H1
Hpa I
Bam HI + Hpa I digestion
G
CCTAG
AACTAA
TTGATT
synthetic DNA
duplex cloning
synthetic DNA insert
CCCGGATCCGGAGTATAATATATCAAGAGCCTAATAACTCGGGCTATAAACTAA
GGGCCTAGGCCTCATATTATATAGTTCTCGGATTATTGAGCCCGATATTTGATT
wild-type T4 DNA sequence

differences from wild-type operator sequence between the *Bam*HI and *Hpa*I sites of the plasmid; *Hpa*I cleaves to the 5' side of the residue at position -15 from the AUG (Figure 24). Thus, mutant (or wild-type control) sequences were introduced into the -16 to -38 segment of the RNA, while the other regions (-38 to -54 and downstream of the residue at -16) contained wild-type T4 sequences. Figure 25 summarizes the gp43 translational operators I constructed for this study.

To examine the importance of the stem pairing of the operator hairpin, I constructed two operator mutants, SD (stem destroyed) and SR (stem restored) (Figure 25). SD changed the 5′ nucleotide of each G-C pair of the stem. SR maintained these same base substitutions while also making compensatory substitutions on the 3′ half of the stem. Several other base substitutions were aimed at assessing the importance of certain loop residues, and their position: 1) UA29 and UA25 mutants changed U residues to A in the positions -29 and -25 respectively, 2) CA26 changed the C residue at -26 to an A, 3) NoU substituted A for both U residues of the loop , and 4) the LR mutation reversed (or "flipped") the loop sequence. This latter mutation attempted to assess the importance of residue position while maintaining loop composition. The former operator mutants were constructed to test the individual contribution of those residues to gp43 affinity and binding computer assisted analysis predicted that none of the loop mutations altered secondary structures of the operator.

Figure 25: A summary of translational operator mutants constructed in this study

The upper right portion of the figure highlights the altered residues of mutants in the operator stem pairing, and the upper left lists mutants of the loop segment. Changed residues are bolded. Designations and nomenclature are the same as in Figure 2. The ΔG's for the entire *stem and loop* region are shown for the WT and SR operators (calculated by GCG fold). Abbreviations: WT = wild-type sequence, UA29 = U changed to A at position -29, UA25 = U changed to A at position -25, NoU = both U's changed to A, CA26 = C changed to A at position -26, LR = entire loop sequence reversed, AU18 = A changed to U at position -18, SD = stem destroyed, and SR = stem restored.


Loop Sequences:
5' A A U A A C U C 3' WT
A A A A A C U C UA29
A A U A A C A C UA25
A A A A A C A C NoU
A A U A A A U C CA26
C U C A A U A A LR
Stem Sequences:
SD
SR
-30
-25
-40
-10
U(AU18)
5' ..... GGAGUAUAAUAUAUCAAG AUAAACUAAGGAAUAUCU AUG AAA GAA ..... 3'
gp43
Met Lys Glu . . .
init.
S/D
proc.
gp43 operator

Finally a mutation (AU18) at the base of the operator hairpin changed an A residue to U. Additional plasmid clones (named Sal I deletions - Sal^V - Figure 23) were made which isolated gene 43 from upstream genes yet encoded 5′ mRNA leaders of longer lengths. These DNA fragments (analogous to MA508 - Figure 23) were also ligated into the pMY7201 T7 expression vector, for expression studies *in vivo*.

4) *IN VITRO* BINDING OF GP43 TO THE TRANSLATIONAL OPERATOR

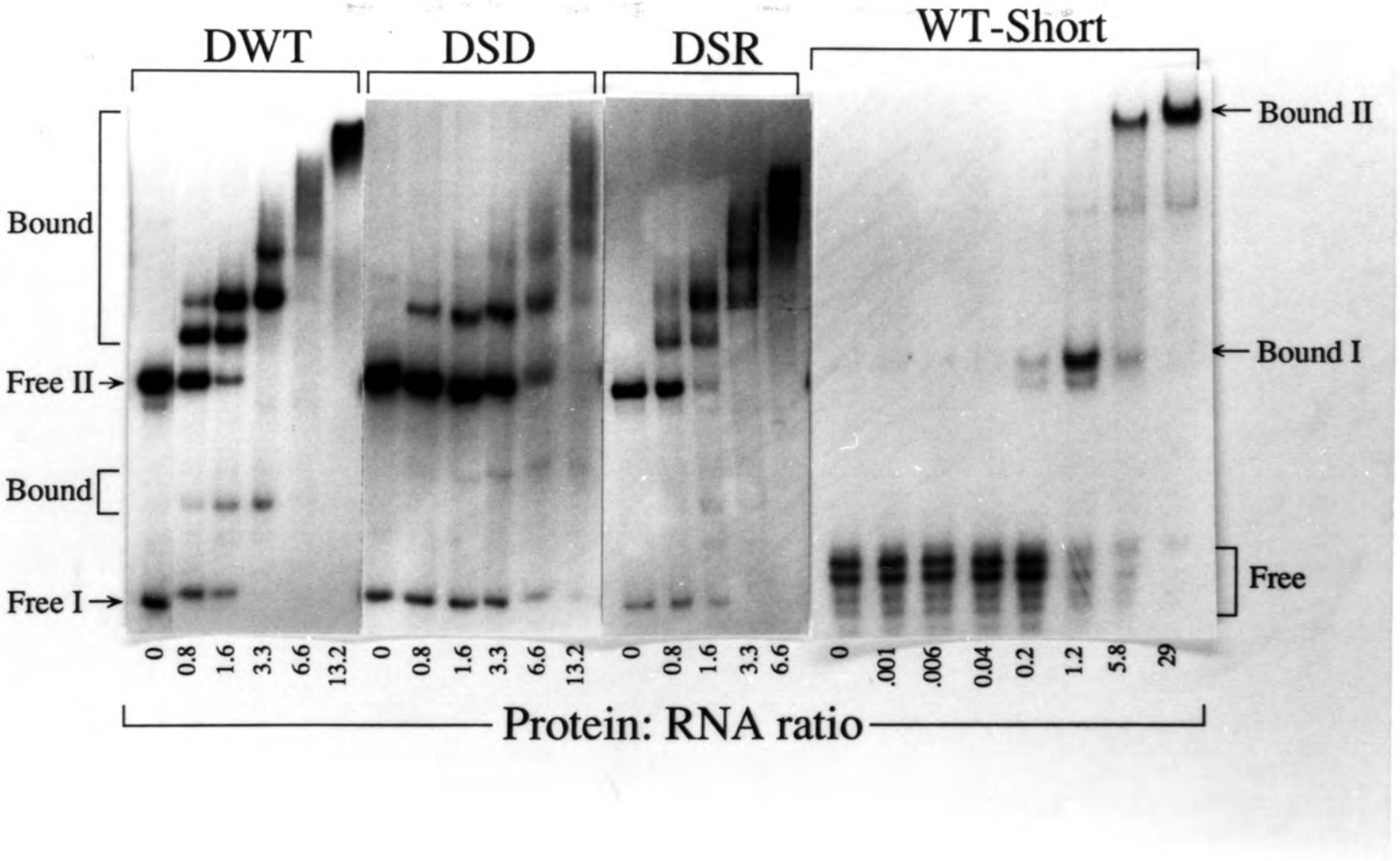
Operator constructs were transcribed *in vitro* with T7 RNA polymerase and the ^{32}P-labeled RNA products were assayed for binding to purified T4 DNA polymerase by using the RNA gel-retardation assay described in MATERIALS AND METHODS. Results in which the wild-type (WT), stem-destroyed (SD), and stem-restored (SR) operators were compared are shown in Figure 26. In the native gel system used for these assays, each of the 207-nucleotide long transcripts resolved into two widely separate bands (marked Free I and Free II, Figure 26) which probably represent alternate conformations of the same RNA since identical samples yielded single species on denaturing urea gels (not shown). Also, analysis of a much shorter (74-nucleotide) RNA harboring the wild-type operator sequence exhibited a simpler electrophoretic pattern on an 8% acrylamide gel (WT-short, Figure 26).

Figure 26: *In vitro* binding of gp43 to operator RNA

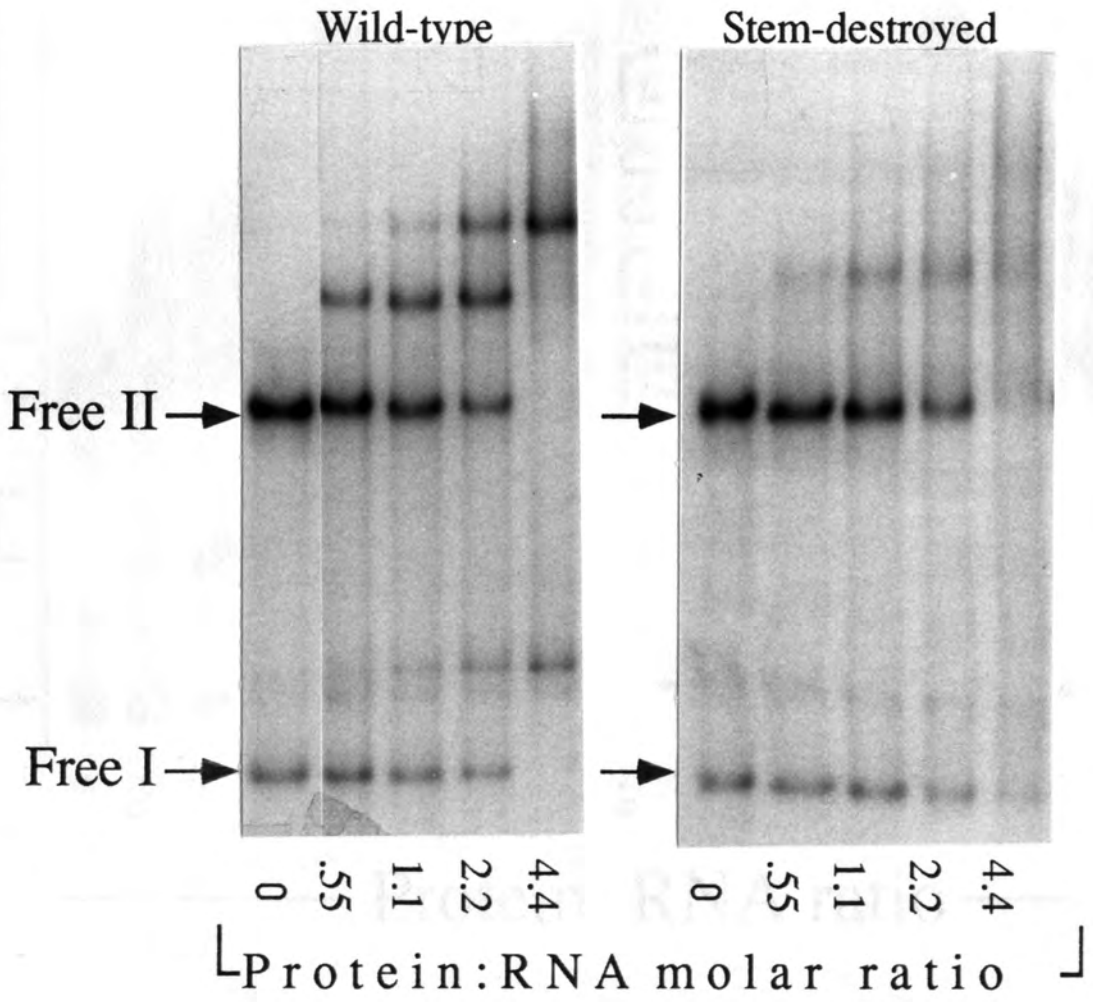
Five pages, panels A through E.

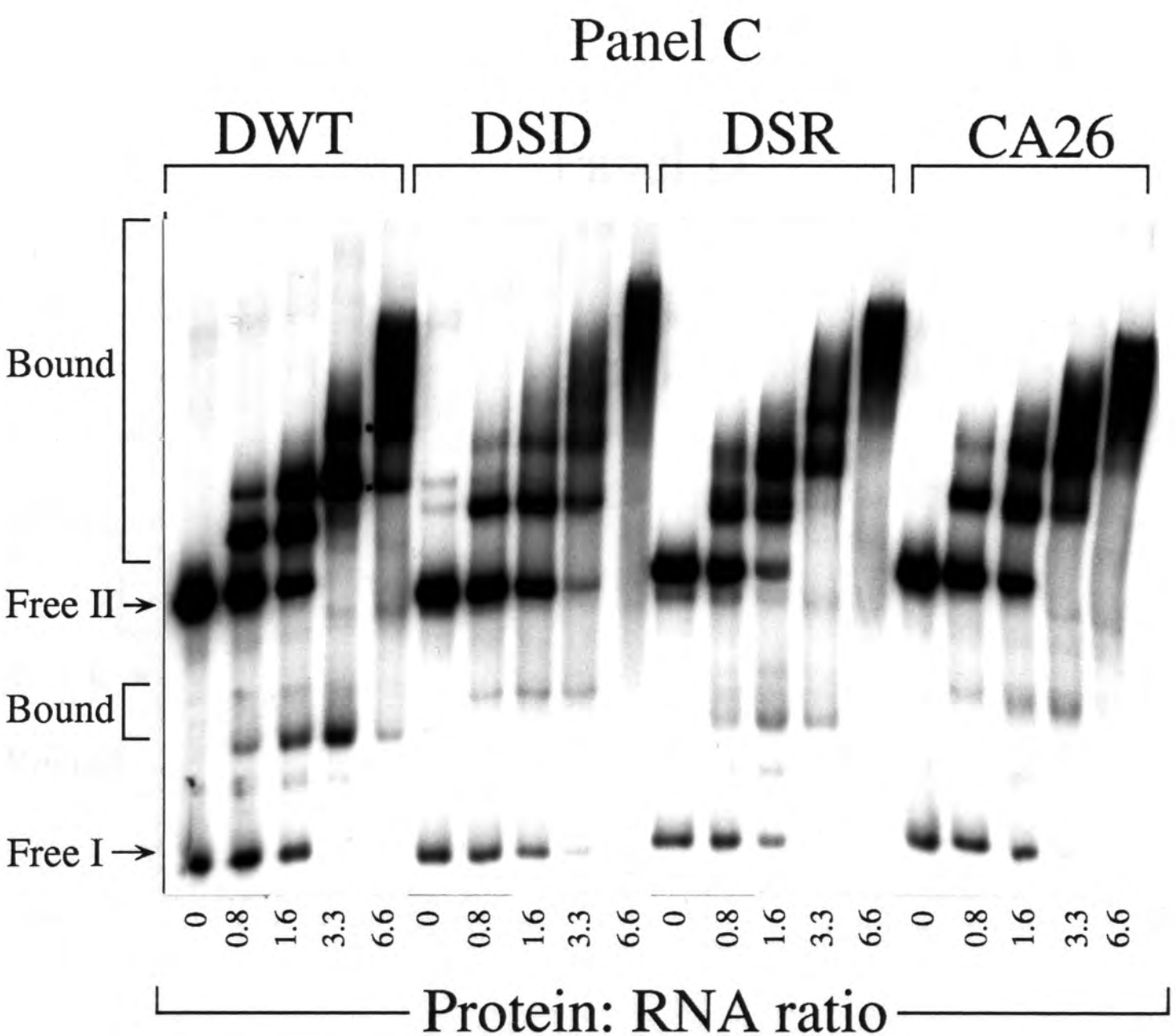
Operator RNA- gp43 binding was assayed by using a gel mobility shift technique (MATERIALS AND METHODS). Arrows point to the bound and unbound (free) RNA fractions resolved. The gp43:RNA molar ratios used are indicated below each lane, and the RNA used for each set of assays is indicated above the lanes. The RNAs used in the WT, SD, SR sets were 207-base run-off RNAs that were transcribed from *Xmn* I digested operator derivatives of pT7D104 (see Figure 24). The "WT-short" panel shows results of an assay with a 74-base RNA containing a wild-type operator sequence (resolved on an 8% acrylamide gel). Several other gel shift assays are shown in panels B through E to show data for all constructs, but are not directly refered to in the text. The operator RNA used in each assay is labeled above each autoradiogram.

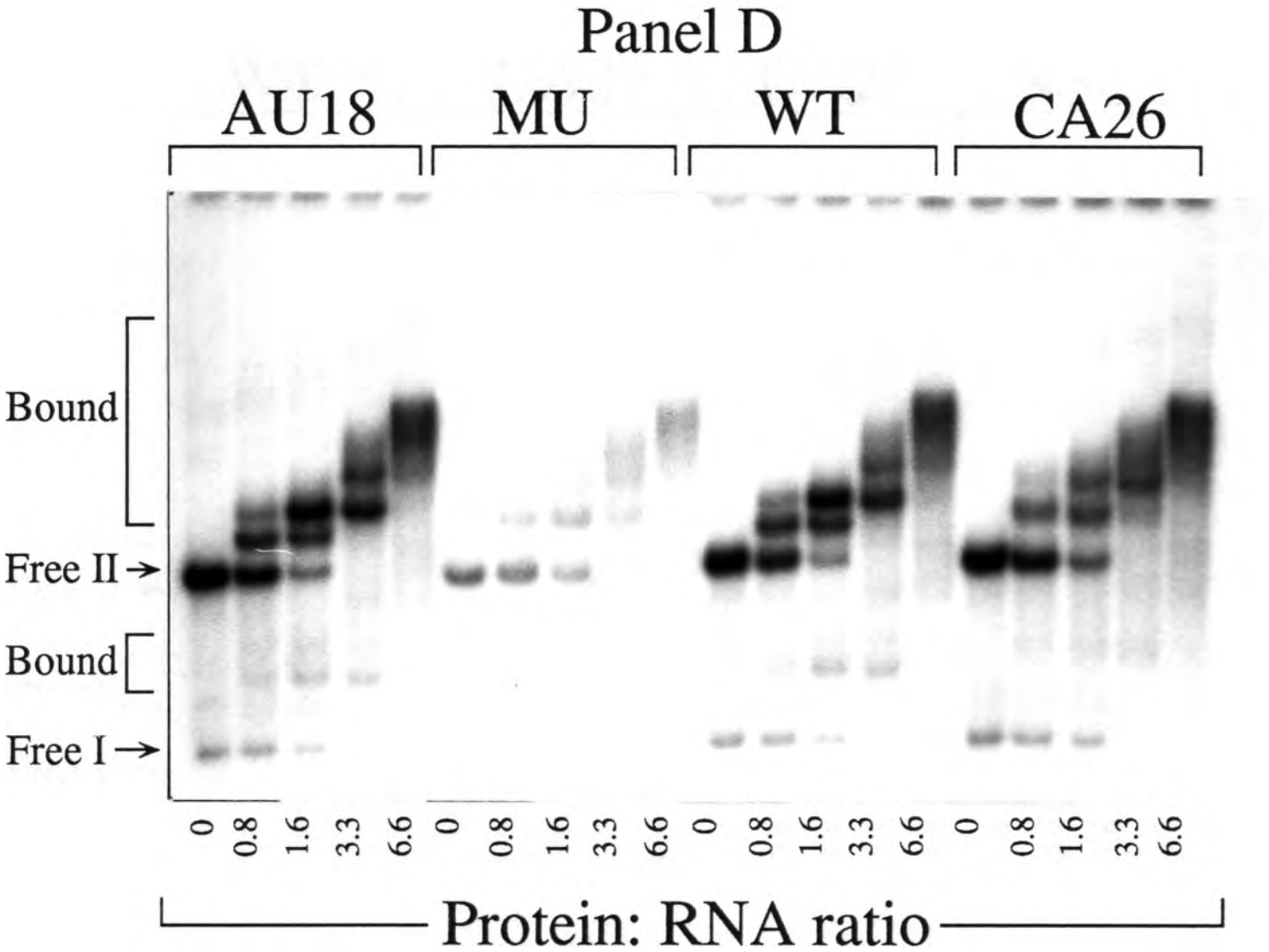
Panel A

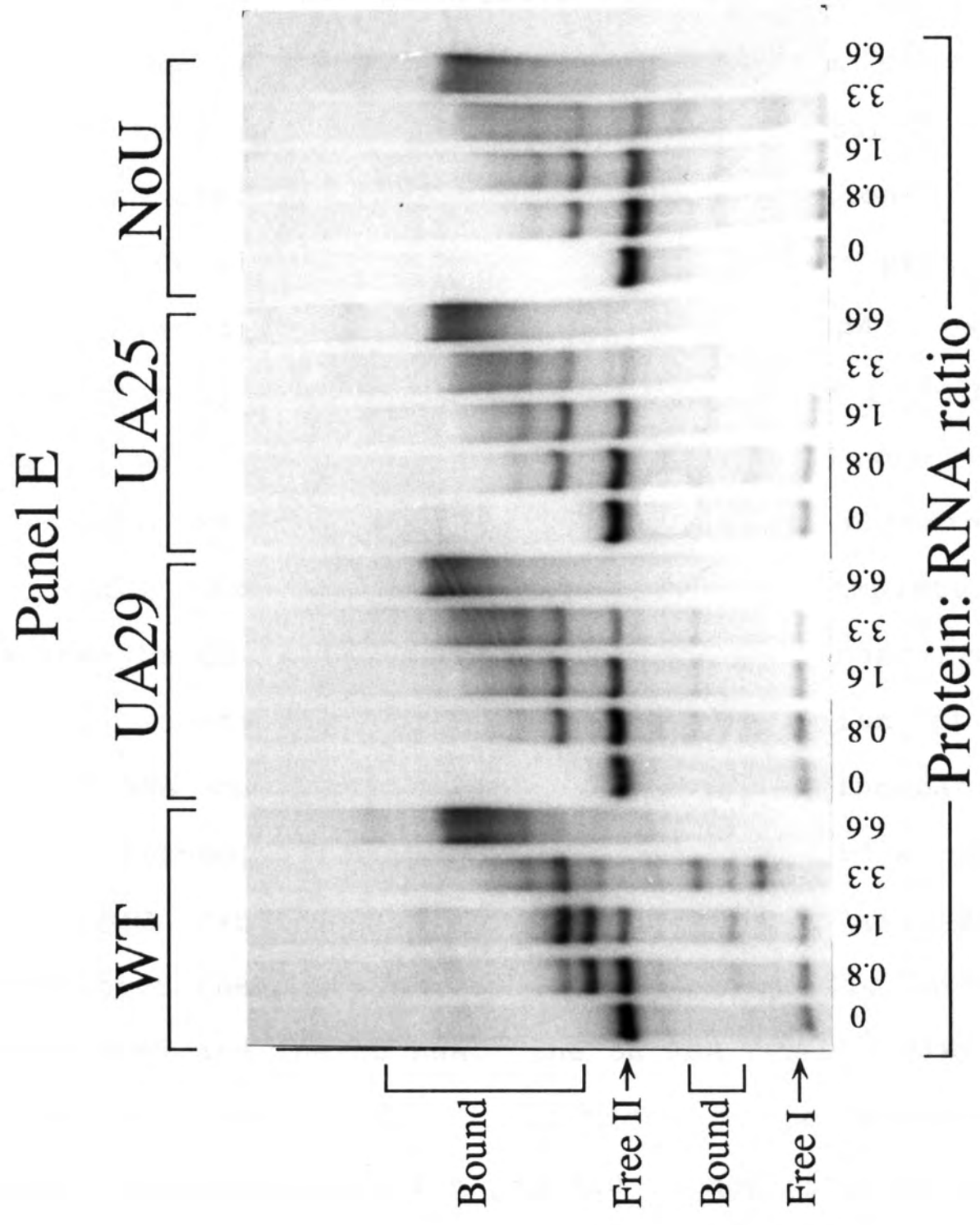

DWT
DSD
DSR
WT-Short
Bound
Free II
Bound
Free I
Bound II
Bound I
Free
0
0.8
1.6
3.3
6.6
13.2
0
0.8
1.6
3.3
6.6
13.2
0
0.8
1.6
3.3
6.6
0
.001
.006
0.04
0.2
1.2
5.8
29
Protein: RNA ratio

Panel B


Wild-type
Stem-destroyed
Free II
Free I
0
.55
1.1
2.2
4.4
0
.55
1.1
2.2
4.4
Protein:RNA molar ratio


Panel C
DWT
DSD
DSR
CA26
Bound
Free II
Bound
Free I
0
0.8
1.6
3.3
6.6
Protein: RNA ratio


Panel D
AU18
MU
WT
CA26
Bound
Free II
Bound
Free I
0
0.8
1.6
3.3
6.6
Protein: RNA ratio


Panel E
WT
UA29
UA25
NoU
0
0.8
1.6
3.3
6.6
Protein: RNA ratio
Free I →
Bound
Free II →
Bound

This RNA yielded a doublet, possibly due to the template-independent addition by T7 RNA polymerase of a nucleotide to the end of the transcript (Milligan, *et al.* 1987). Alternatively, T7 RNA polymerase may terminate transcription at both the beginning and end of the Bam H1 overhang; although this explanation is less likely since only the template strand is required for transcription elongation (Milligan, et al., 1987). WT samples that were pre-incubated with increasing amounts of purified gp43 exhibited correspondingly decreasing intensities of the Free I and II bands as well as the appearance of new, slower migrating bands (labeled "Bound" in Figure 26). The protein:WT RNA molar ratio that was required to effect a complete "shift" from free bands to discrete slower migrating bands was about 3:1. In contrast, electrophoretic migration of the SD operator RNA was much less sensitive to pre-incubation with protein, although this RNA did ultimately yield a smear when protein:RNA ratios in excess of 10:1 were analyzed. We presume that the smearing is due to non-specific interactions between gp43 and the SD RNA. The SR RNA resembled WT closely in its response to gp43 addition, which suggests that operator recognition by the protein is not strongly dependent on specific nucleotide sequence of the hairpin's stem segment in this defined *in vitro* system.

Densitometric scans of "Free" and "Bound" RNA bands from gel-shift assays were used to calculate K_d values for the

interaction of gp43 with WT and variant operators (MATERIALS AND METHODS). Examples of such analyses are shown in Figure 27 and a summary of calculated K_d values on all operator constructs analyzed is presented in Table 5. Both lengths of WT RNA examined yielded K_d values of about 10^{-8}M, which demonstrates that the experimental approach used here to determine Kds for protein-RNA interaction is applicable to a wide range of RNA sizes. The use of higher concentrations of protein with WT RNAs resulted in a second shift to even slower migrating bands, which, on the basis of relative density, we suspect to be formed by the incorporation of additional protein molecules per mole of RNA, *i.e.* protein aggregation, rather than by recruitment of more RNA per mole of protein.

The two most dramatic effects on operator-gp43 binding were observed with the SD and UA29 operator variants. Each exhibited about a 7-fold decrease in gp43 binding relative to WT (Table 5). Restoration of stem-complementarity with a non-wild-type sequence restored *in vitro* binding to wild-type levels (SR construct, Figure 26 and Table 5).

Figure 27: Kinetics of gp43 binding with the WT, SD, SR, and UA29 operators

Densitometric scans of autoradiograms from gel-shift experiments, such as those depicted in figure 17, were used to plot the fraction of RNA bound against increasing gp43 concentrations. The squares and triangles mark the actual data points for the WT and SD operators, respectively, with standard deviations shown for comparison. The curves drawn through the data points are theoretical and were generated from the calculated apparent Kds listed in Table 1 for each operator (see MATERIALS AND METHODS).

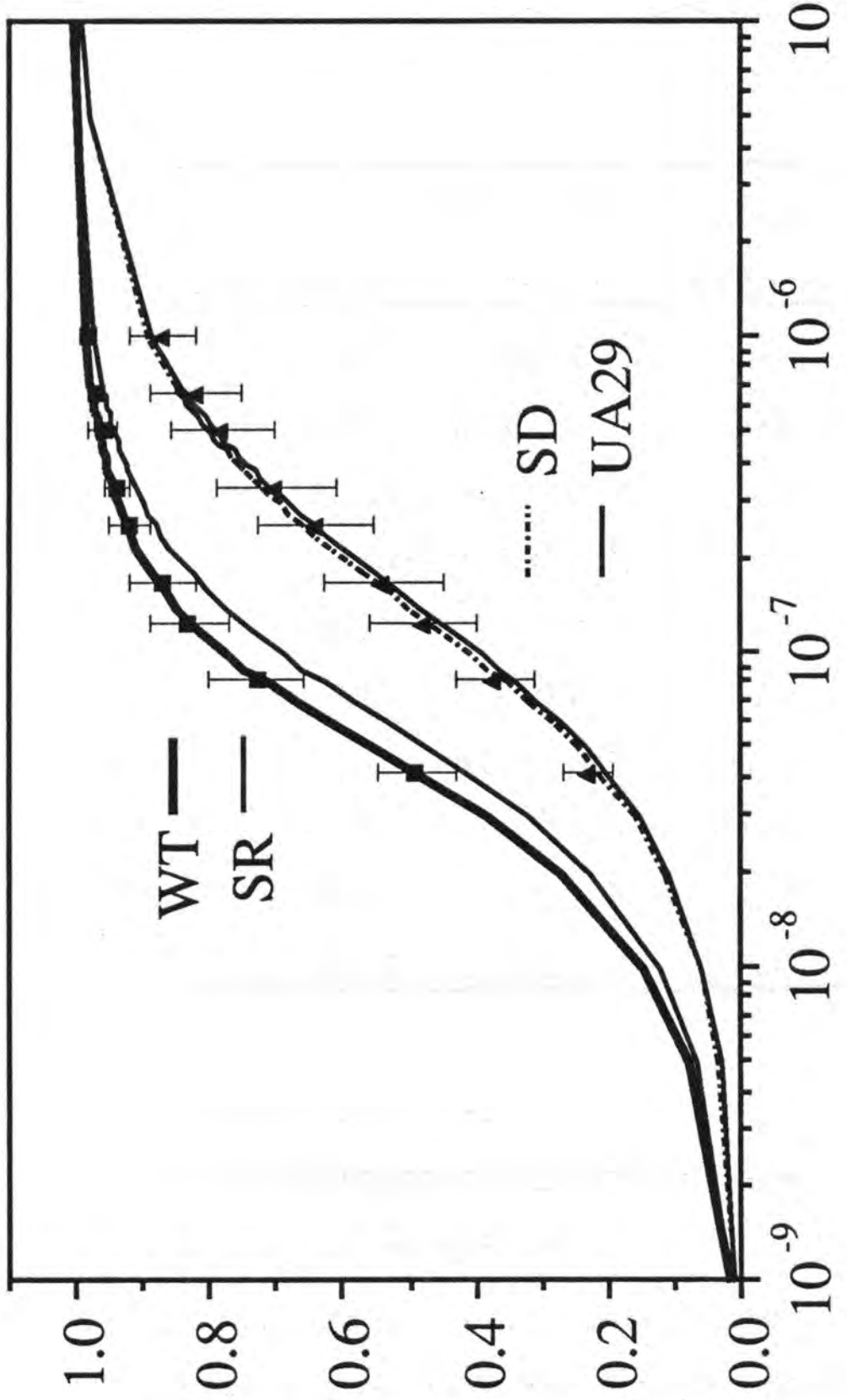

Fraction of RNA bound
1.0
0.8
0.6
0.4
0.2
0.0
WT
SR
SD
UA29
10^{-9}
10^{-8}
10^{-7}
10^{-6}
10^{-5}
[gp43], M

Table 5: Dissociation constants for the interaction of gp43 with wild-type and mutant operators

Operator[a]	Apparent K_d[b]	Fold effect
DWT	1.9×10^{-8}[c]	--
DSD	1.2×10^{-7}	6.5
DSR	1.9×10^{-8}	1.0
LR	9.6×10^{-8}	5.0
UA29	1.3×10^{-7}	6.9
UA25	9.0×10^{-8}	4.7
NoU	8.9×10^{-8}	4.6
MU	8.9×10^{-8}	4.6
CA26	3.3×10^{-8}	1.7
AU18	1.9×10^{-8}	1.0

[a] See Figure 25.

[b] See MATERIALS AND METHODS for method of determination.

[c] This Kd value is approximately 4-fold higher than one determined by Tuerk and Gold (1990) See the Discussion for details.

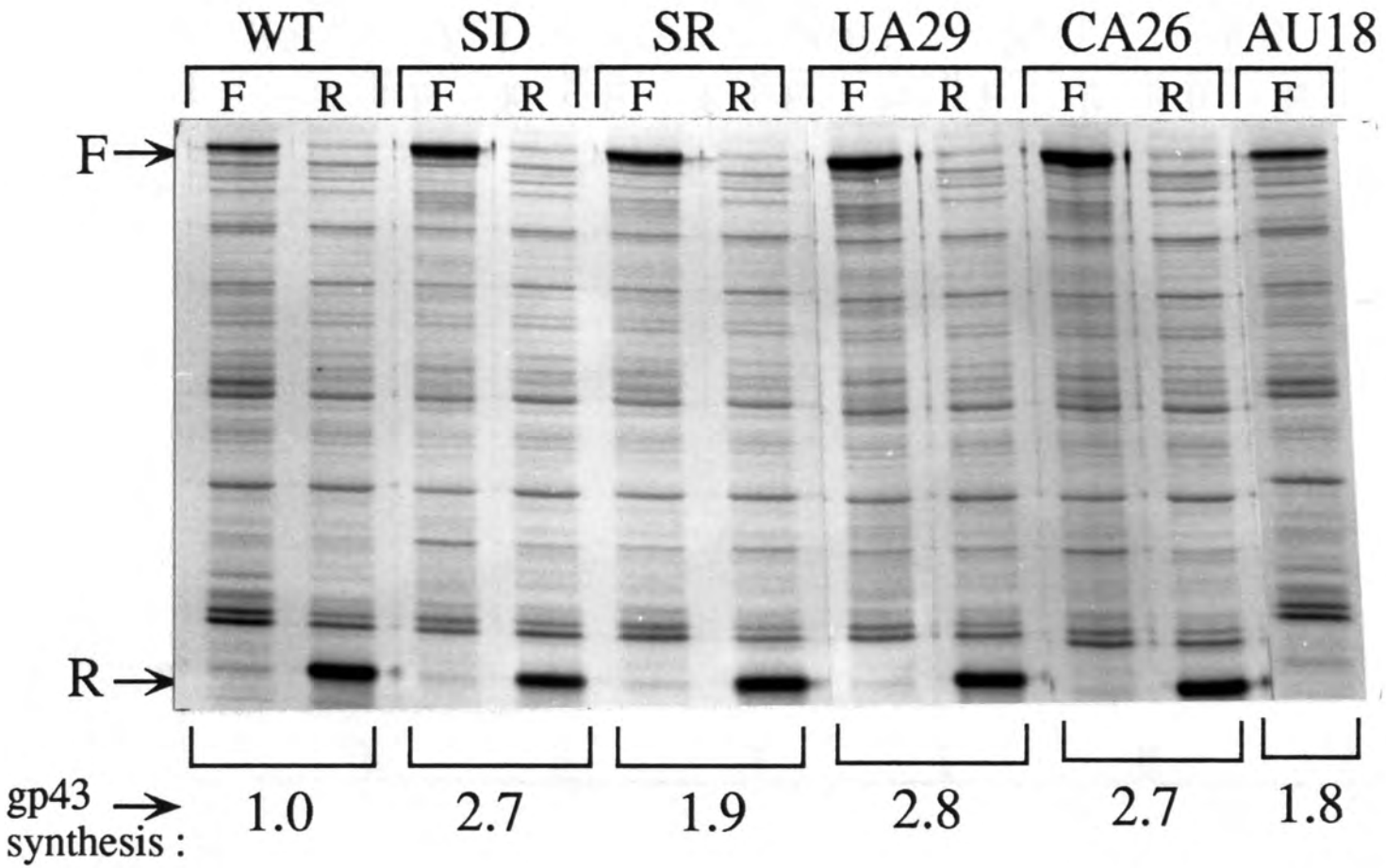
However, this change in operator stem sequence still exhibited strong biological effects, *in vivo*.(see next 2 sections). Similarly, a C to A change in the loop at position -26 (CA26) had only a small effect on protein binding *in vitro* (Table 5) while exhibiting a strong effect on plasmid-directed gp43 biosynthesis and gp43 synthesis in infections (see below). A base substitution near the base of the hairpin (AU18) had only small effects on the gp43-operator interaction both *in vivo* (Figure 28 - below) and *in vitro* (Table 5)

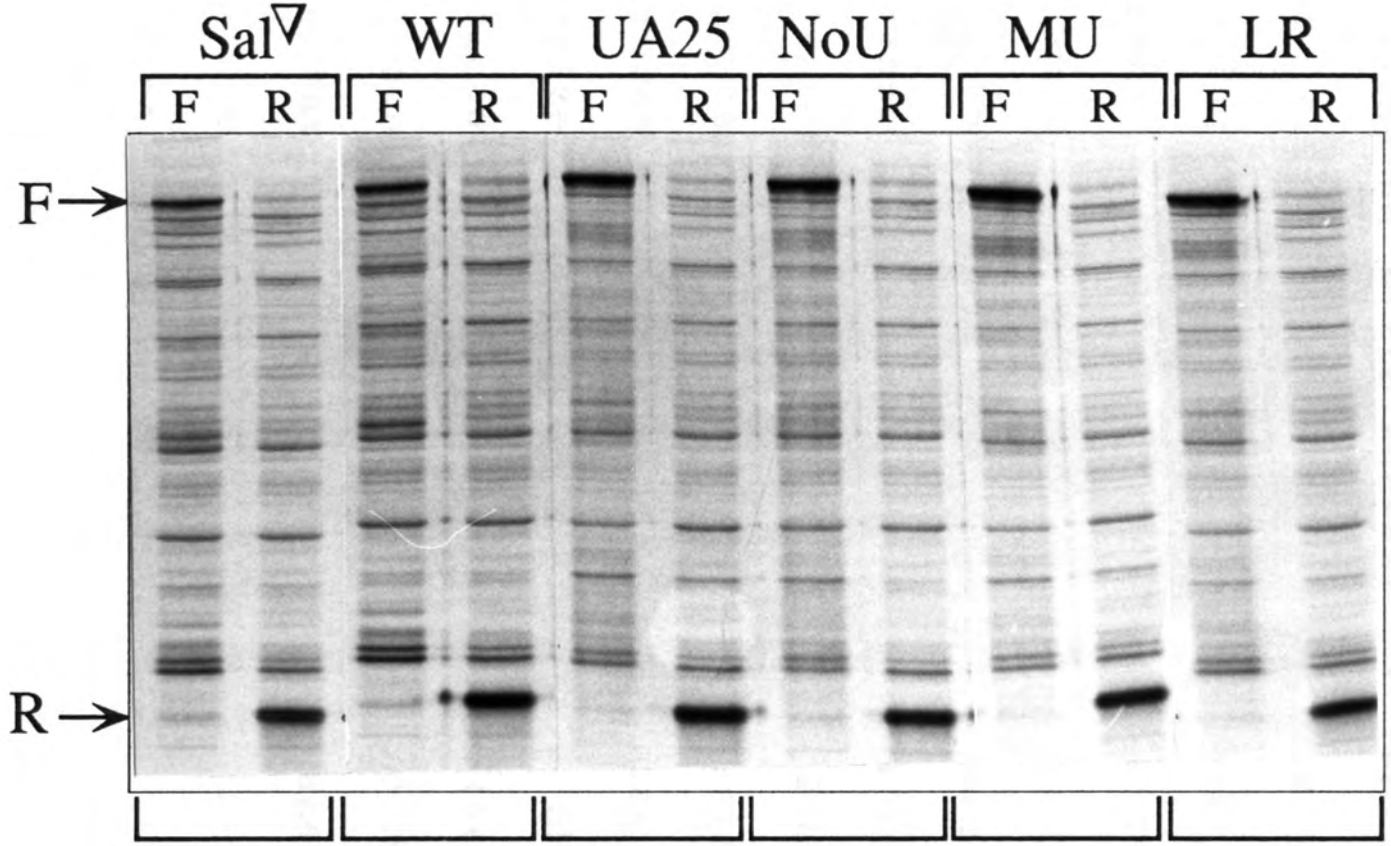
5) Plasmid mediated expression of T4 gene *43* with mutant translational operators

My analysis of these RNA operator mutations *in vivo*, was done in two stages: 1) assaying expression of gp43 from plasmid constructs (a setting which includes cell factors involved in gene regulation), and 2) expression of gp43 in T4 infections (a setting which also adds phage factors that may participate in translational regulation - see next section). Figure 28 shows examples of plasmid-mediated expression of cloned gene *43* sequences in the presence of some the operator constructs diagrammed in Figure 25. In clones expressing the wild-type structural gene (clones assigned the F designation), it was easy to distinguish between derepressed and autogenously repressed constructs.

Figure 28: Effects of translational operator mutations on plasmid-directed gp43 biosynthesis

Recombinant-plasmid-bearing *E. coli* BL21(DE3) cells were grown, induced with IPTG, labeled for proteins, and analyzed by SDS gel-electrophoresis and autoradiography as described in MATERIALS AND METHODS. F denotes plasmid clones in which internal WG17 fragment was inserted in a normal (forward) orientation (see Figure 1), R denotes reverse orientation for WG17. The full length 103.5 kD gp43, produced by F clones is marked by an F->, and the truncated N-terminal protein produced by R clones is marked by R-> on the autoradiogram. The mutant operators analyzed are described in Figure 3.


WT
SD
SR
UA29
CA26
AU18
F
R
F
R
F
R
F
R
F
R
F
F
R
gp43
synthesis :
1.0
2.7
1.9
2.8
2.7
1.8


Sal∇
WT
UA25
NoU
MU
LR
F
R
F
R

Differential effects by the various operator mutations were also evident in this *in vivo* assay. All the clones encoding a complete gp43 polypeptide (F constructs) supported growth (100% plating efficiencies) of the gene *43* double amber mutant *43amE4301-E4322* under un-induced conditions, suggesting that derepression was not caused by inadvertent mutations in the cloned structural gene (data not shown). The R constructs did not support the growth of amber mutant phage. These reverse orientation clones (R), which exhibited derepression, served as controls for possible alterations in translation initiation efficiency. I detected no significant differences in the level of synthesis among these reverse constructs (Figure 28). Note that although SR is not synthesized at as high a level as SD, nonetheless, it is significantly derepressed, *in vivo*.(see Discussion for Part II) I concluded that minor sequence changes in the operator region have definite effects on gene expression *in vivo*. Next, I examined the effects of these mutations on gp43 expression in infections.

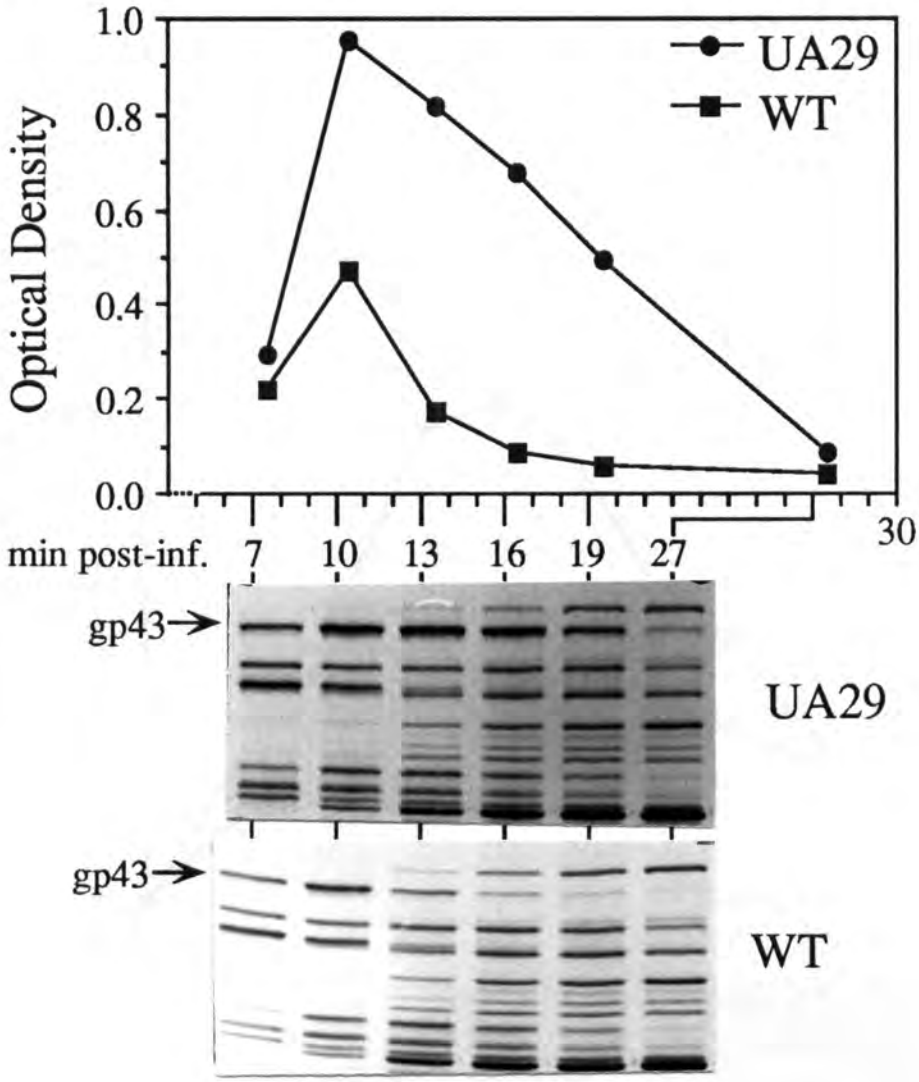
6) Effects of gp43 operator mutations in T4 infections

I introduced several operator mutations into the T4 phage genome by phage-plasmid recombination (*i.e.*, marker rescue, MATERIALS AND METHODS) and examined their effects on gp43 biosynthesis, DNA synthesis, and phage production in standard phage infections.

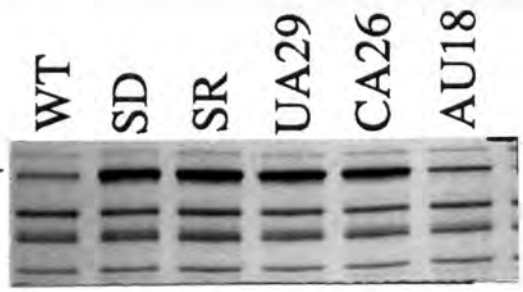
Figures 29 and **30**: Effects of translational operator mutations on gp43 synthesis in T4 infections

Panel A shows the level of gp43 synthesized during infections of *E. coli* Nap IV with phage strains carrying WT or the UA29 operator. Experimental details were as described previously (Andrake *et.al.*, 1988; Hughes *et al.*, 1987) and in MATERIALS AND METHODS. Infected cultures were pulse-labeled with ^{35}S precursor for 5 min at overlapping time intervals post-infection, and subsequently analyzed by SDS-polyacrylamide electrophoresis and autoradiography. The position of gp43 is marked by arrows and dots, and the operator genotype for each infection is indicated on the graph and autoradiograms. Levels of gp43 synthesis were calculated from densitometric scans of autoradiograms, and the data were plotted at the midpoints of the pulses on the time-scale shown. Panel B compares the relative levels of gp43 synthesis at 13 min post-infection (11-16 min pulse) with several operator phage mutants. The position of gp43 in the autoradiograms and operator mutant names are shown. In addition, listed below each gel lane are values for the levels of DNA synthesis (^{3}H-thymidine incorporation) and burst size for each mutant phage relative to WT. The WT infection in these comparisons yielded a burst size of 504 (measured at 2 hours post-infection) and a ^{3}H-thymidine incorporation rate of 4290 dpms/min (during the interval between 20 and 60 minutes post-infection).

A

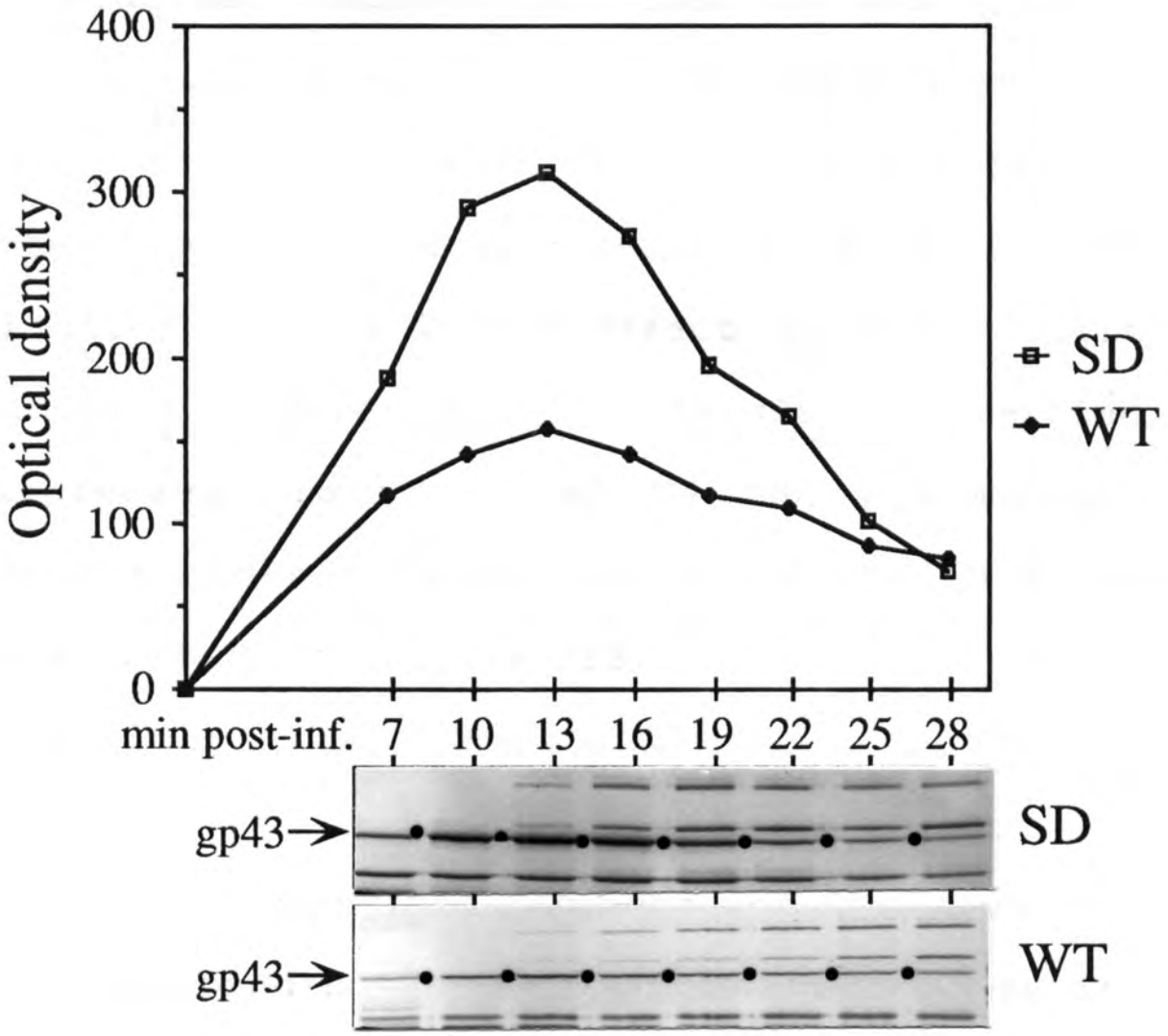


B



Relative values of

	WT	SD	SR	UA29	CA26	AU18
gp43 synthesis	1.0	7.3	8.3	8.2	7.2	1.9
DNA synthesis	1.0	0.9	1.0	1.0	1.0	1.0
Burst size	1.0	0.9	1.1	1.0	0.9	1.0


400
300
200
100
0
Optical density
min post-inf. 7 10 13 16 19 22 25 28
SD
WT
gp43
SD
gp43
WT

Examples of results are presented in Figures 29 and 30. Operator constitutivity was reflected by a derepression of gp43 biosynthesis during infection. As was observed with plasmid-mediated expression, the SR and CA26 mutations exhibited strong effects on gene *43* expression, while their effects *in vitro* were weak (Figures 26 and Table 5). The AU18 mutant, which bound gp43 nearly normally *in vitro* (Table 5), exhibited about a 2-fold effect on gp43 biosynthesis in T4 infections (Figure 29B). It is also interesting to note that increased production of T4 DNA polymerase in such infections neither enhanced nor diminished DNA replication and phage production (Figure 29B).

7) Discussion

Operator Sequence and Structure

The mRNA binding site for T4 DNA polymerase is at least 36-40 nucleotides long and includes a putative hairpin structure that consists of a 5 base-pair stem and an 8-base loop (Figs. 22 and 25 - Andrake, *et al.* 1988; Tuerk, *et al.* 1990). The studies reported here show that dimensions and stability of this structure constitute important determinants for specificity of the gp43-mRNA interaction. These results agree with the results of others examining different mutations in this mRNA operator (Tuerk, *et. al* 1990) Our mutational analyses further demonstrates that this specificity depends on the nucleotide sequence of the loop.

In particular, single base changes involving uracil residues in the hairpin's loop (positions -25 and -29 from the AUG, Figure 25) drastically reduced gp43-RNA binding *in vitro* as measured by a gel retardation assay (Figs. 26 and 27), and exhibited strong derepression of gp43 biosynthesis in vivo as measured by plasmid-mediated expression (Figure 28) and by expression in T4 infections (Figures 29 and 30). Base substitutions at other residues within the loop and near the base of the hairpin resulted in differential effects both *in vitro* and *in vivo*. Results with altered stem sequences, *i.e.* stem-destruction (SD) and stem-restoration (SR), suggest that both stability *and* sequence of the base-paired region are important criteria for protein binding and repression *in vivo*. Although the SR operator RNA bound gp43 as effectively as wild-type *in vitro* (Figure 26), it exhibited derepression of gp43 biosynthesis *in vivo* (Figs. 28 and 29). Similarly, another operator mutation, a C to A change in the loop (CA26), showed small effects *in vitro*, but exhibited strong effects *in vivo*. The contrasting results between *in vitro* and *in vivo* experiments may be due to a variety of factors that could influence the gp43-operator RNA interaction and RNA conformation *in vivo*, e.g. proteins that can interact with gp43 or RNA, or long-range intramolecular interactions between the hairpin and other segments of the gp43 mRNA (pseudoknot formation). In addition, a somewhat decreased stability of the SR hairpin relative to WT (ΔG of -4.8 vs. -5.4, respectively, for stem-and-loop stability), may be

magnified *in vivo* without the helix-stabilizing effect of the high salt buffer (Vorlickova, *et al.*, 1990) used in the *in vitro* binding assay (MATERIALS AND METHODS). In regard to loop mutations, I have no proof that changes in loop sequence do not alter the stability of operator structure. The stability of RNA hairpin structures is generally independent of loop base composition (Groebe and Uhlenbeck 1988), but in some cases it is not (Cheong, Varani and Tinoco 1990; Tuerk, *et al.* 1988). None of the loop mutants were predicted to change operator folding as assessed by computer assisted analysis (See Discussion and Conclusions).

Other RNA operators

Our results also suggest that the general principles underlying RNA-protein interaction at the gp43 translational operator are similar to those that govern several other prokaryotic translational repressors, including R17 coat protein (Uhlenbeck, Wu and Sampson 1987), *E. coli* ribosomal proteins (Nomura, Gourse and Baughman 1984), *E. coli* threonyl-tRNA synthetase (ThrS) (Springer, *et al.* 1988), and others. In all these cases, both RNA primary and higher-order structures play roles in determining specificity of RNA-repressor interactions. In the case of the *E. coli thrS* gene, operator mutations that were selected for derepression of translation *in vivo* either disrupted RNA stem formation (as confirmed by RNase mapping studies) or changed key hairpin loop residues (Springer, *et al.* 1988). Interestingly,

the same nucleotides that maintain stem-and-loop stability in the wild-type *thrS* operator are present in $tRNA^{Thr}$, demonstrating an interesting case of molecular mimicry between primary and secondary ligands of the ThrRSase protein. In the case of the RNA operator hairpin that binds R17 coat protein, specific protein binding was found to require three of the four loop residues, as well as a bulged stem residue (Uhlenbeck, *et al.* 1987). In addition, mutations that disrupted the stem pairing resulted in reduced binding of the protein, while base pair changes in the stem did not affect binding as long as hairpin stability was maintained (Romaniuk, *et al.* 1987). In another example, mutational analysis of the putative *tar* mRNA binding site of the HIV-1 Tat protein revealed that 5 of the 6 residues in the loop could not be substituted without loss of Tat-dependent *trans* activation (and presumably Tat protein-RNA binding). Disrupting the top base pair of the stem resulted in loss of *trans* -activation, whereas an alternate base pair in this position restored activity (Feng and Holland 1988). Thus, the theme for site-specific protein-RNA interaction in these examples portrays a structured operator in which specific single-stranded loop (or bulged stem) residues are placed in an appropriate three-dimensional context for contact and recognition by specific amino acid residues in the protein. There may be a tendency, based on the data available from these examples, to suspect that unpaired bases are more likely to interact with amino acid residues of RNA-

binding proteins. In principle, however, the grooves in A-form RNA helices may well provide RNA-protein contacts in a manner analogous to the recognition of sequence specific DNA-binding proteins (Schleif 1988). The interaction of T4 gp32 with its RNA operator offers yet another variation to the RNA sequence/structure requirement. Here, the protein seems to recognize structure (a pseudoknot) for entry into the RNA, but ultimately blocks translation by binding cooperatively further downstream from the structure at a specific sequence that overlaps the RBS (Fulford and Model 1984; McPheeters, *et al.* 1988). Recognition via an RNA pseudoknot has also been documented in the α operon of *E. coli* ribosome biogenesis (Tang and Draper 1989). One translational repressor, the T4 RegA protein, may not utilize RNA shape at all in finding its target. This protein is small (122 amino acids) and binds to short and diverse uridine-rich targets around the initiator AUGs of a subset of T4-induced early mRNAs (Unnithan, *et al.* 1990); however, RNA structure in RegA-mediated translational repression may play a role by sequestering unwanted targets or exposing functionally relevant ones.

Operator Loop Sequence and gp43 Interaction

Another indication that T4 DNA polymerase recognizes specific nucleotide residues in addition to structure comes from studies of the chromatographic behavior of the enzyme on RNA columns (unpublished results from this laboratory). In those studies, the protein exhibited vastly different binding

affinities to different homopolyribonucleotides. PolyU bound gp43 five times as strongly as polyA (or polyG) and twice as strongly as polyC. So, we suspect that pyrimidines (particularly uridines) constitute major contact points between gp43 and its mRNA target. The carbonyl groups in a pyrimidine ring (perhaps optimally in the number 4 position) may function as hydrogen acceptors while interacting with specific gp43 amino acid residues in the protein's binding site. In a three-dimensional context, we envisage a structured gp43 RNA target in which size of the stem-and-loop segment and a distribution of adenosines (possibly non-binding residues) allow for the placement of pyrimidines in the optimal positions for such contact with the protein. An extension of the mutational analyses described here as well as direct structural analyses on gp43, its mRNA-binding site, and RNA-gp43 complexes will ultimately elucidate the fine details of this interaction.

Physiology of gp43 Autogenous Regulation

Finally, I shall comment on the physiological aspects of T4 DNA polymerase autogenous regulation. This enzyme is remarkably versatile in that it harbors DNA synthesizing, 3'-DNA-exonucleolytic, nucleotide-binding, DNA-binding, and RNA-binding activities, all contained in a single polypeptide chain of fewer than 900 residues (Goulian, Lucas and Kornberg 1968; Spicer, *et al.* 1988). The protein also possesses the capacity to interact specifically with other protein

components of the T4 DNA replication complex (protein-protein binding activities) (Jarvis, *et al.* 1990). We are interested in how the RNA binding activity of gp43 relates physiologically to the other activities of the protein, particularly in regard to its participation in DNA replication. For example, can gp43 autogenous translational repression vs. its assembly into the DNA replication complex be likened to feedback translational regulation of ribosome biogenesis by ribosomal subunit proteins? The analogy may well be a pertinent one, and Tuerk *et al.* (1990) demonstrated that a reconstructed DNA replication apparatus could compete with a synthetic RNA target for gp43 binding *in vitro*. However, it should be noted that the DNA-binding and RNA-binding activities of T4 DNA polymerase may or may not be mediated by the same protein domain. This enzyme binds to DNA primer-template junctions in a sequence non-specific manner while its binding to RNA is characteristically specific to a sequence/structure. From a regulatory stand point, one should also consider the temporally restricted synthesis of gp43 mRNA and its high instability *in vivo* (Hsu and Karam 1990). These two factors ultimately determine the amount of mRNA available for gp43 biosynthesis and autogenous repression at various times during the phage growth cycle. Autogenous translational repression of gp43 biosynthesis may be an important mechanism by which the phage sets enzyme dosage during the pre-replicative stage of the phage growth cycle, and may become irrelevant at later times when the gene

43 mRNA concentration is too low (due to degradation and cessation of early transcription) to compete with the abundant binding sites on DNA in the replication pool. Furthermore, gp43 functions in concert with other proteins whose dosages are also important for determining replication levels. This is perhaps why a selective overproduction of wild-type T4 DNA polymerase (translational operator constitutive mutants) is not accompanied by increased DNA synthesis and phage yield (Figure 29 and Tuerk, *et al.* 1990). Also, the fact that such phage mutants are viable may also mean that the DNA region specifying the translational operator is not a site for a critical gp43-DNA interaction *in vivo*.

IV) CONCLUSIONS AND GENERAL DISCUSSION

A) CONCLUSIONS

T4 DNA polymerase is an autogenous translational repressor

- In mutant phage infections, overproduction of the protein is not accompanied by increased gene *43* transcription or increased mRNA stability.
- Purified T4 DNA polymerase specifically represses the translation of gene *43* mRNA. Other T4 encoded proteins may affect gp43 activity in repression.

When binding to its mRNA operator, T4 DNA polymerase recognizes specific nucleotides held in a defined secondary structure conformation.

- The stem pairing of an RNA hairpin structure is required for gp43 binding to its mRNA operator.
- Varying the sequence of the stem base pairs does not affect gp43 binding affinity *in vitro*, but does result in derepression *in vivo*.
- The presence and position of specific loop residues

are required for wild-type binding affinity *in vitro*, and repression *in vivo*.

- Since the effect of a particular operator mutation *in vivo* can differ from its effect on *in vitro* binding experiments, it is possible that elements other than base sequence may affect binding and repression (*e.g.*RNA higher order structure, phage or host factors, *etc*).

B) COMPUTER ASSISTED ANALYSIS OF RNA SECONDARY STRUCTURE

INTRODUCTION

Ultimately, the biological activity of any macromolecule is dependent on its three-dimensional conformation in solution. In the case of RNA, the tools currently available for determining 3-D conformation are limited. Solution mapping by RNase digestion (using a battery of single and double-stranded specific nucleases) and chemical modification studies are among the methods commonly in use. These methods largely assess the accessibility of nucleotide residues to the action of enzymes or modification by chemicals. The results of such analysis only reveal whether or not particular residues are paired; this information must still be interpreted to develop a higher order structure model approximates three-dimensional conformation. This is still far from the molecular resolution that crystalographic analysis can potentially provide.

In the midst of these methods, computer algorithms for predicting RNA secondary structure present first approximations of RNA structure and constitute an adequate basis for experimental design and for models that explore the relationship of structure to function in RNA molecules.

There are several limitations to these computer assisted analytical methods. The most obvious shortcoming is that these programs are based on assumptions that oversimplify reality. In the first place, the mathematical model for calculations is a discrete one; a nucleotide is either base paired or not. There is no provision for differing bond lengths, angles, etc. This simplification is necessary when applying a recursive algorithm; otherwise calculations would take inordinately long to perform. The algorithms compute a structure that minimizes the free energy of the folded RNA. The energy calculations are based on thermodynamic parameters which have been calculated from melting temperatures of short oligonucleotides in solution, and other theoretical considerations. The global energy of folding is then the sum of individual energies of base pairs and loops. Again this principle of additivity may or may not correspond to reality. Current algorithms can not predict higher order RNA structure (see examples below) nor can they account for the sequential(temporal) aspects of RNA folding (Martinez, 1984; Nussinov and Tinoco, 1981). In addition, RNA structure, *in vivo*, is not dictated solely by intramolecular interactions. The possible function of protein-mediated stabilization or

destabilization of structures *in vivo* may be an important determinant for RNA folding.

Nevertheless, when used in combination with experimental data, these programs can provide useful information, especially with respect to designing the next experiment. Confidence is increased when a predicted structure agrees with data from other analysis of the RNA structure under consideration, whether that be phylogenetic comparisons or the above mentioned solution mapping techniques. A prediction of local structure is considered "robust" (more reliable) when various criteria are met: 1) the same structure is predicted when differing lengths of upstream and downstream sequences are included in the analysis 2) the same structure is predicted with minor changes in the energy parameters used for the analysis ("energy perturbation" analysis), 3) the same structure is predicted in sub-optimal foldings that are within a certain percent (10-20%) of the global energy minimum. If certain local motifs are shared among predicted structures despite variation of these parameters then a prediction is considered well defined.

As with any research tool, the use of computer-aided analysis must be done in an informed fashion. The user should be aware of how different factors influence the calculations and final predictions. A careful analysis will 1) try alternate computer algorithms, 2) examine the effects of changing energy parameters on the predicted structures, and 3) try to incorporate experimental data.

RNA HIGHER ORDER STRUCTURE

The classical building block of RNA structure is the stem-and-loop hairpin. Other terminology and types of RNA structure are portrayed in Figure 31. Loops can be further classified as hairpin, bulge, bifurcation (or multi-loop), or interior loops. A pseudoknot structure is the higher order interaction between single-stranded residues of any of these loops (enclosed by a stem) with other complementary bases outside that region. Figures 32 and 33 show some of the ways these interactions might occur. A typical H-type pseudoknot can be defined by two stems (S1 and S2 helices) and two loops (L1 and L2). The minimum length of the connecting loops depends on the length of the helix being crossed. The stability of formation for a given RNA structure arises from both 1) the free energy of formation of hydrogen bonds between base pairs, and 2) the energy of formation when base pairs are stacked in a helix. The energetic favorability of this second parameter is exploited by pseudoknot structures which coaxially stack one helix on the second, forming a quasi-continuous helix (Figure 32). Other structures have been proposed (Abrahams, et al., 1990), differing in the placement of the loop segments with respect to stems (Figure 33). These structures have not been found in natural RNAs to date. Programs able to predict these higher order structures are just now becoming available (Abrahams, et al., 1990); however, these algorithms do not always produce the structure

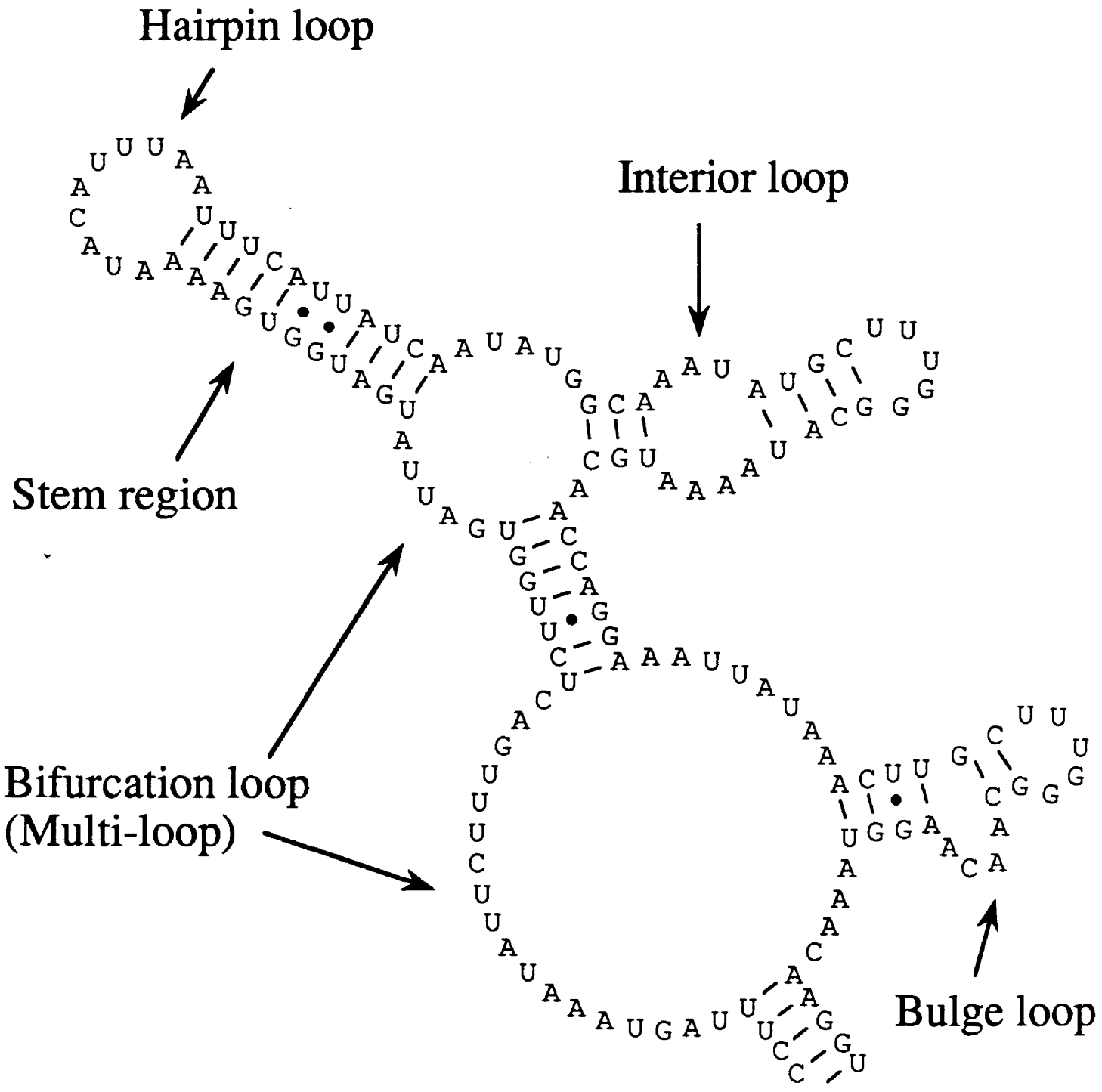


Figure 31: Elements of RNA secondary structure. See text for details.

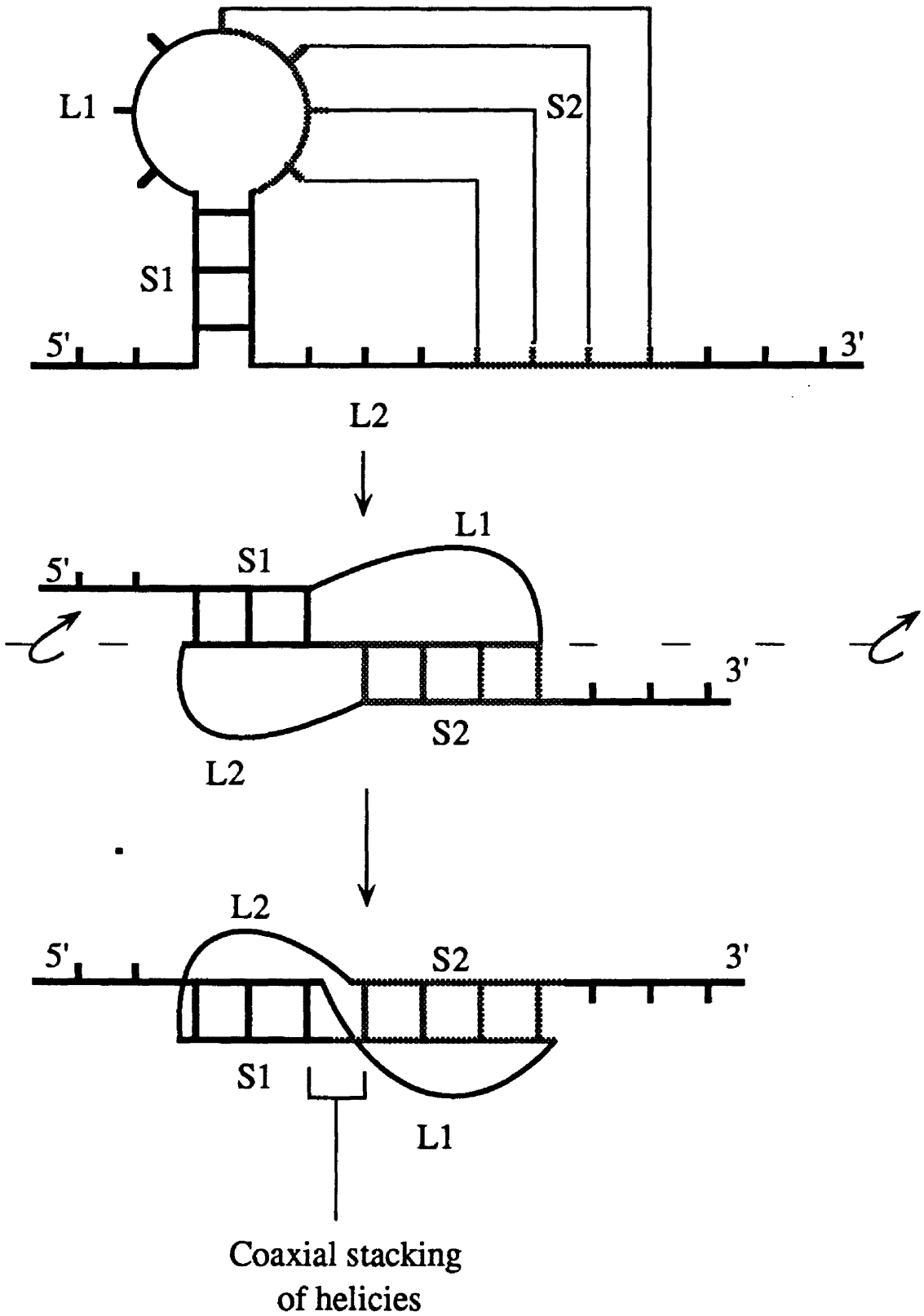
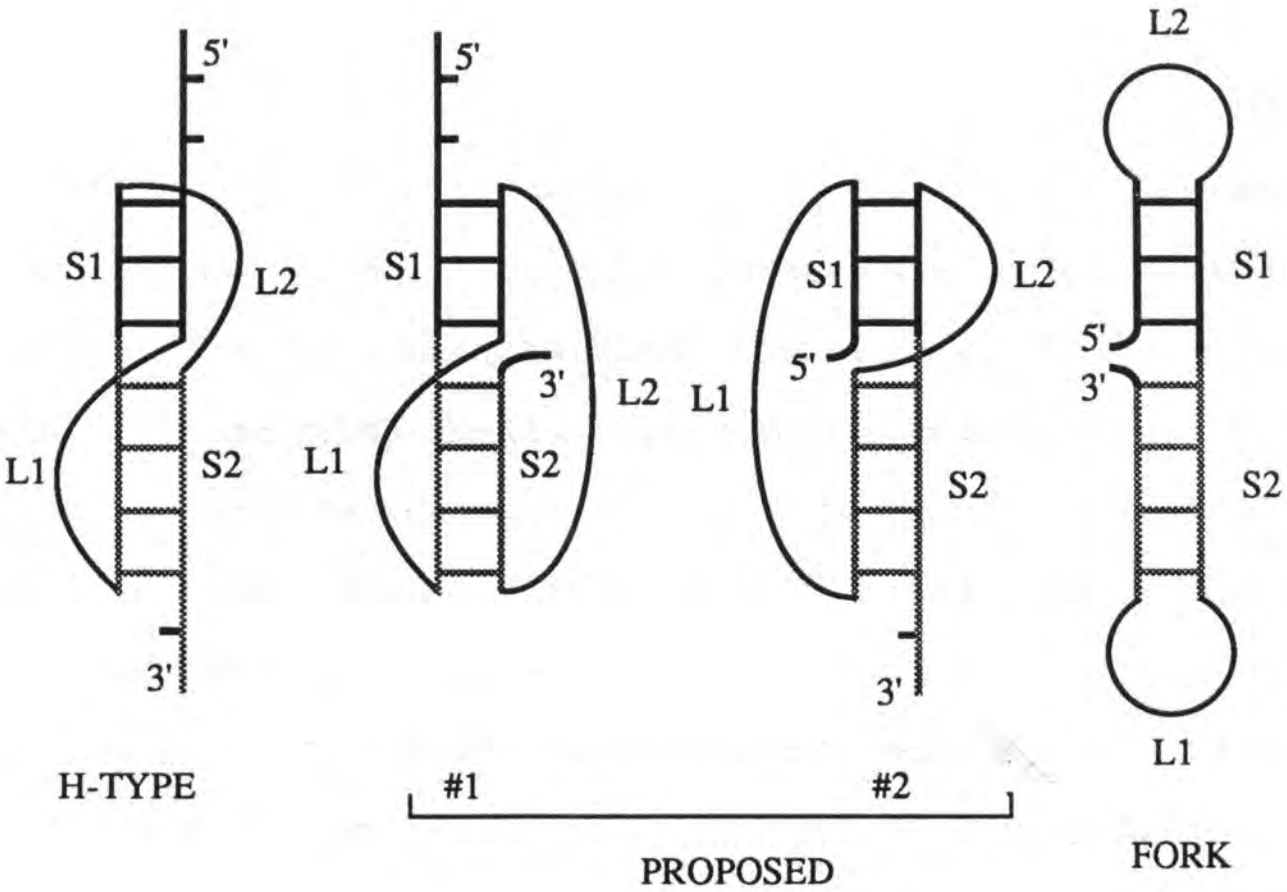


Figure 32: Folding of a common pseudoknot structure.
The H-type ("hairpin") pseudoknot folding is portrayed in three schematic representations to demonstrate the stacking of helicies. Single-stranded loops have a thinner line, and the stem S2 is lightly shaded for clarity. In this structure L1 crosses the shallow groove and L2 the deep groove of A-form RNA.



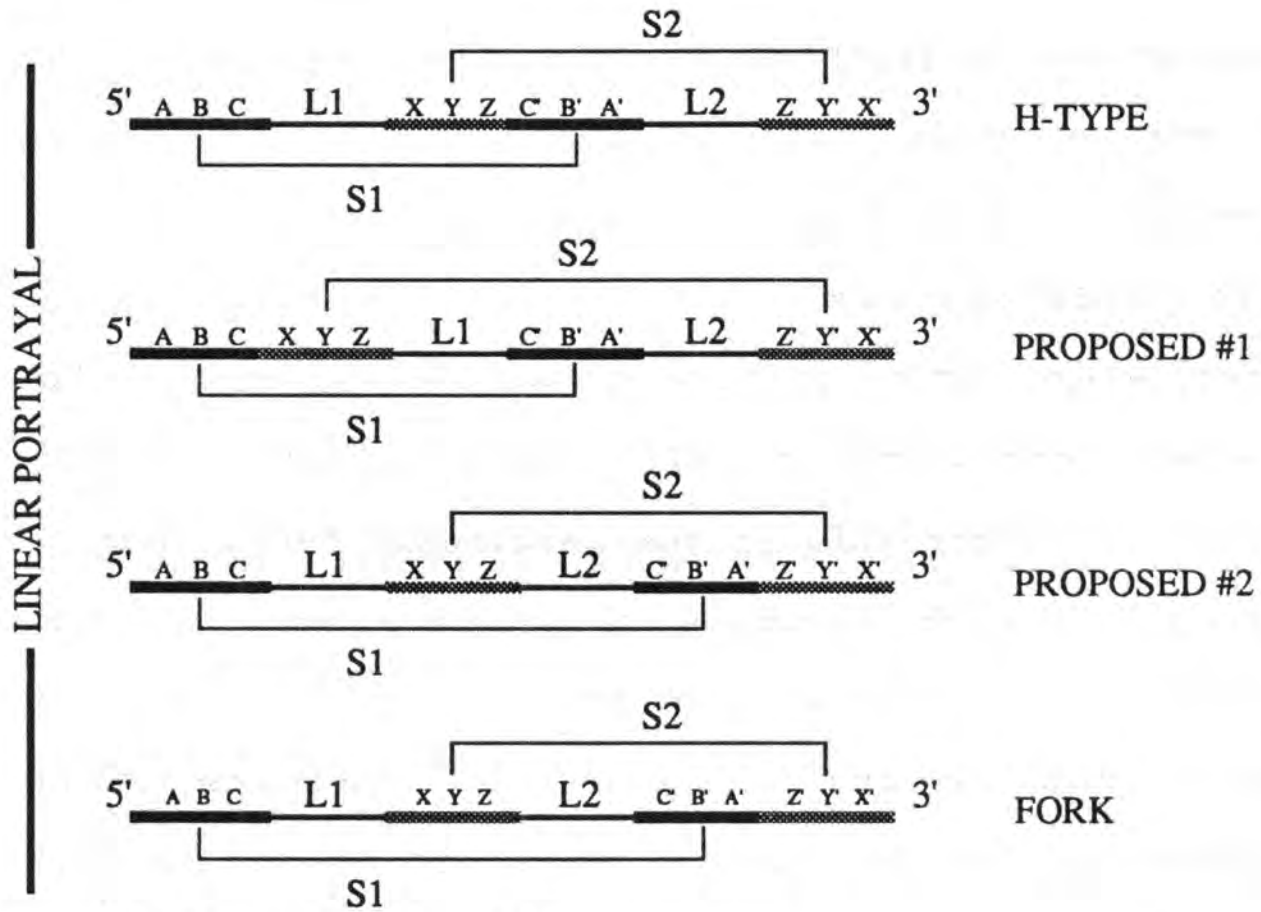


Figure 33: Other Higher order RNA interactions
The same concentions apply as in Figure 32. A' is a base complemetary to A. All of the helicies shown are of the antiparallel. A linear portrayal of each structure is shown below.

with the global minimum free energy, due to the additional computational difficulties introduced by tertiary interactions. The lack of accurate energy parameters for these minimally characterized structures (i.e. connecting loops and coaxial helix stacking energies) has been an obstacle to the development of these programs. These parameters are just now becoming available (Puglisi, et al., 1990; Puglisi, et al., 1990).

Pseudoknots were first described in the tRNA-like structures at the 3′ end of turnip yellow mosaic virus RNA (Rietveld, et al., 1982). Since then they have been described in several other RNAs, including 1) tRNAs - confirmed with X-ray crystalography, 2) rRNA - supported by extensive phylogenetic analysis, 3) at frame shift sites in viral RNAs (required for efficient shifting in some genes), 4) various functional roles in mRNA (protein binding sites, participation in group I intron self-splicing catalysis); for a review see (Pleij, 1990). It is likely that these and other higher order structural motifs play important roles in RNA function. By analogy, it would be absurd to consider protein function while strictly considering secondary structure interactions only; tertiary and subunit interactions are at the heart of several protein catalytic activities. Likewise, several types of pseudoknots (Figure 33) may only be the beginning of elucidating the three-dimensional structural code for RNA function.

Two aspects of nucleic acid conformation which are as yet unexplored in RNA structure are shown schematically in Figure 34. All of the helices and pseudoknot interactions described to date have involved anti-parallel strands in A-form helices. This is certainly the most common helix conformation; however, both of the following have been shown to exist 1) parallel DNA duplexes between two *Drosophila* sequences (Tchurikov, et al., 1989) and 2) intramolecular DNA triplexes in supercoiled plasmids (involving Hoogsteen base pairs where one of the three strands exists in parallel) (Hanvey, et al., 1988). These non-Watson-Crick interactions may seem to be a remote possibility presently; however, RNA may have even less constraint on flexibility (compared to DNA), permitting a variety of these interactions with less steric hindrance. An alternate RNA helix conformation (Z-RNA) has already been described (Davis, et al., 1990). The difficulty in determining higher order RNA structure in general, may account for the lack of their detection to date. Several tempting possibilities for parallel helices (detected by direct complementary repeats in the linear sequence) exist in the *regA-43* polycistronic mRNA (Figure 35). If these structures do exist in RNA, they would certainly present a unique three-dimensional shape that may be important for protein recognition and binding.

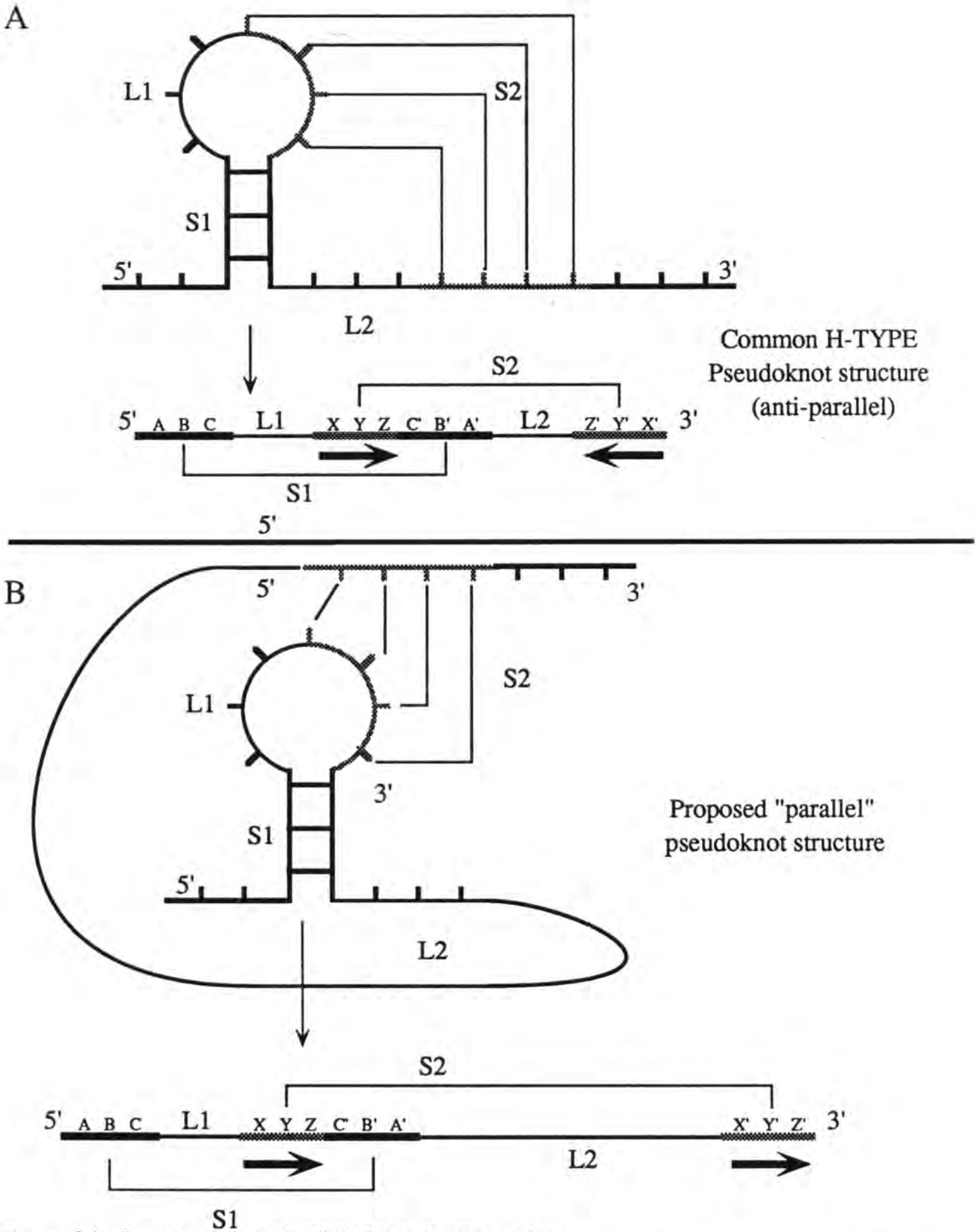
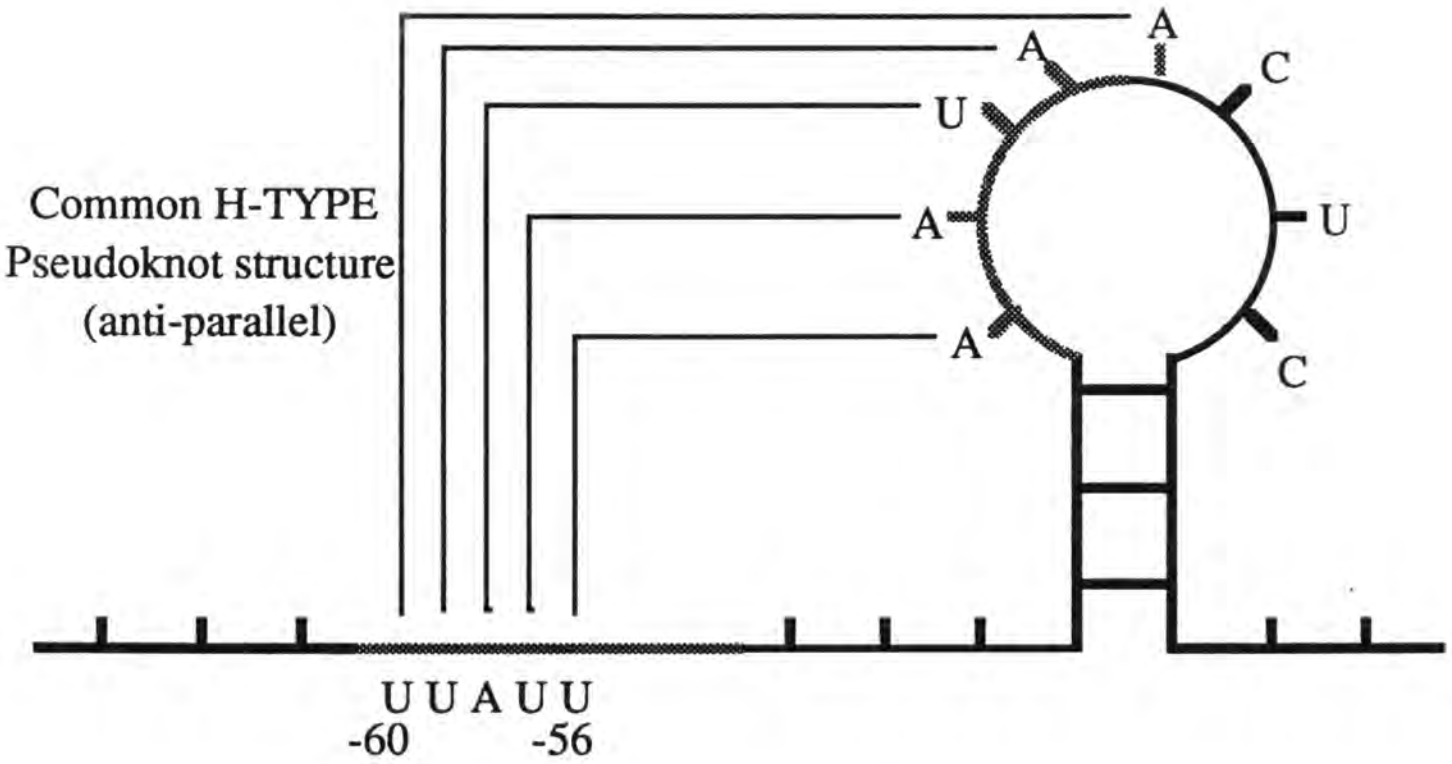


Figure 34: Proposed "parallel" RNA helix formation.
Conventions are the same as in previous figures. A' is a base complementary to A; note the 5' to 3' direction of arrows portraying the parallel or anti-parallel nature of helicies. A parallel helix interaction is seen as a "direct complement repeat" in the linear sequence.



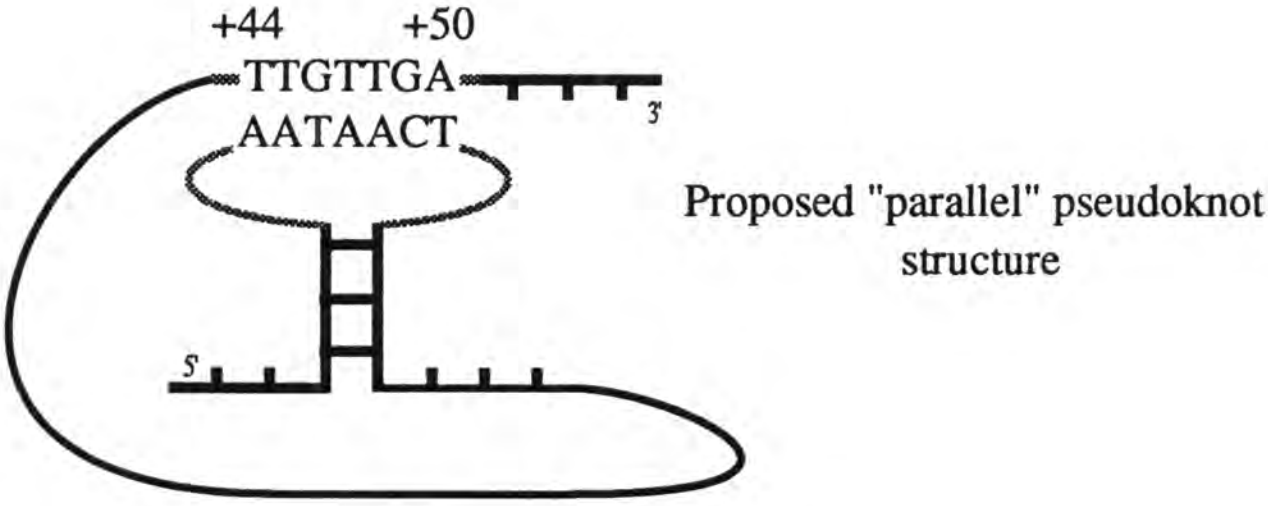


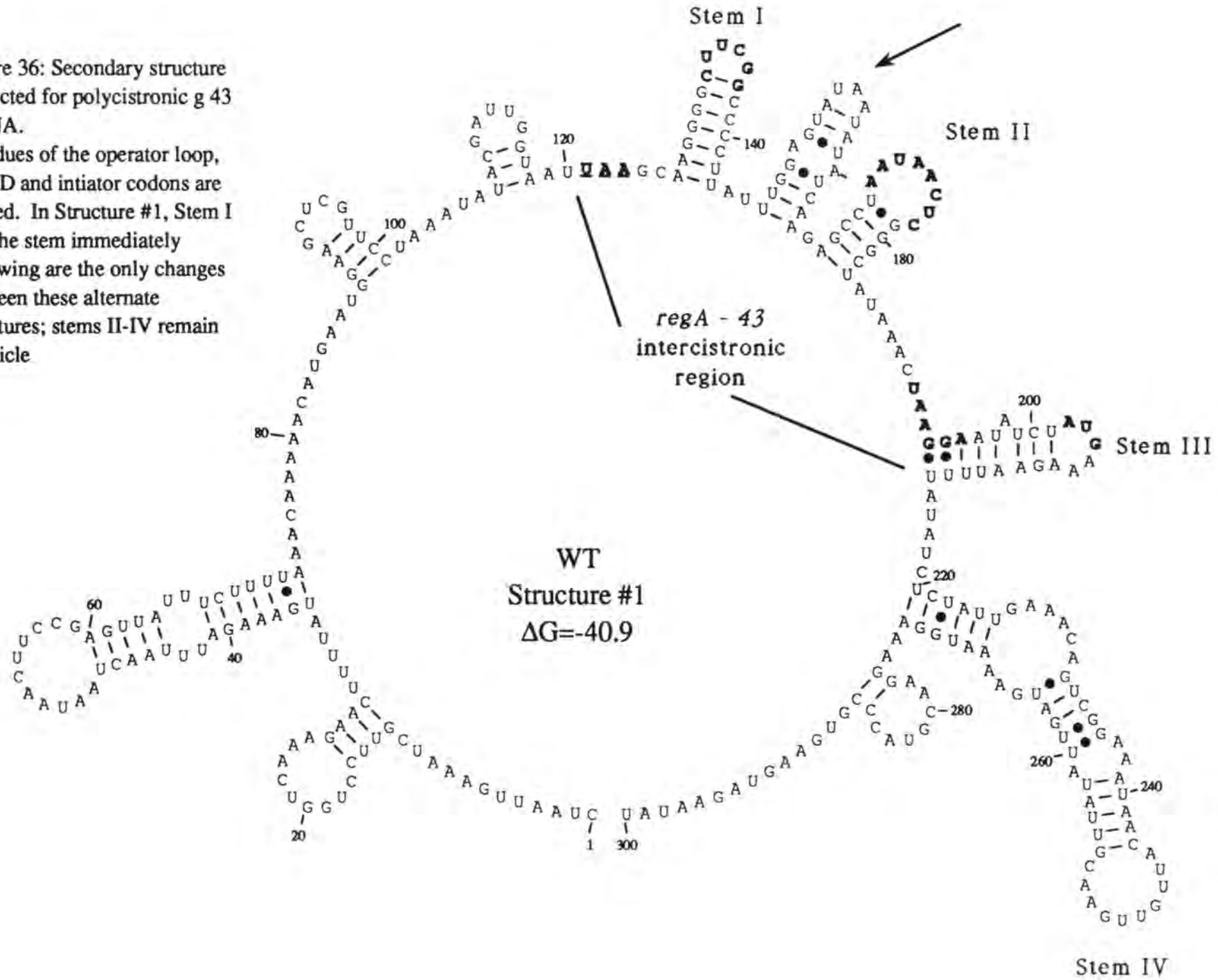
Figure 35: Possible higher order interactions in the gp43 operator
Two possible structures are given as examples, the top being a common H-type pseudoknot and the bottom a parallel helix structure. Interestingly, the +44 to +50 residues exist in the loop of an extended stem structure (see Figure 6).

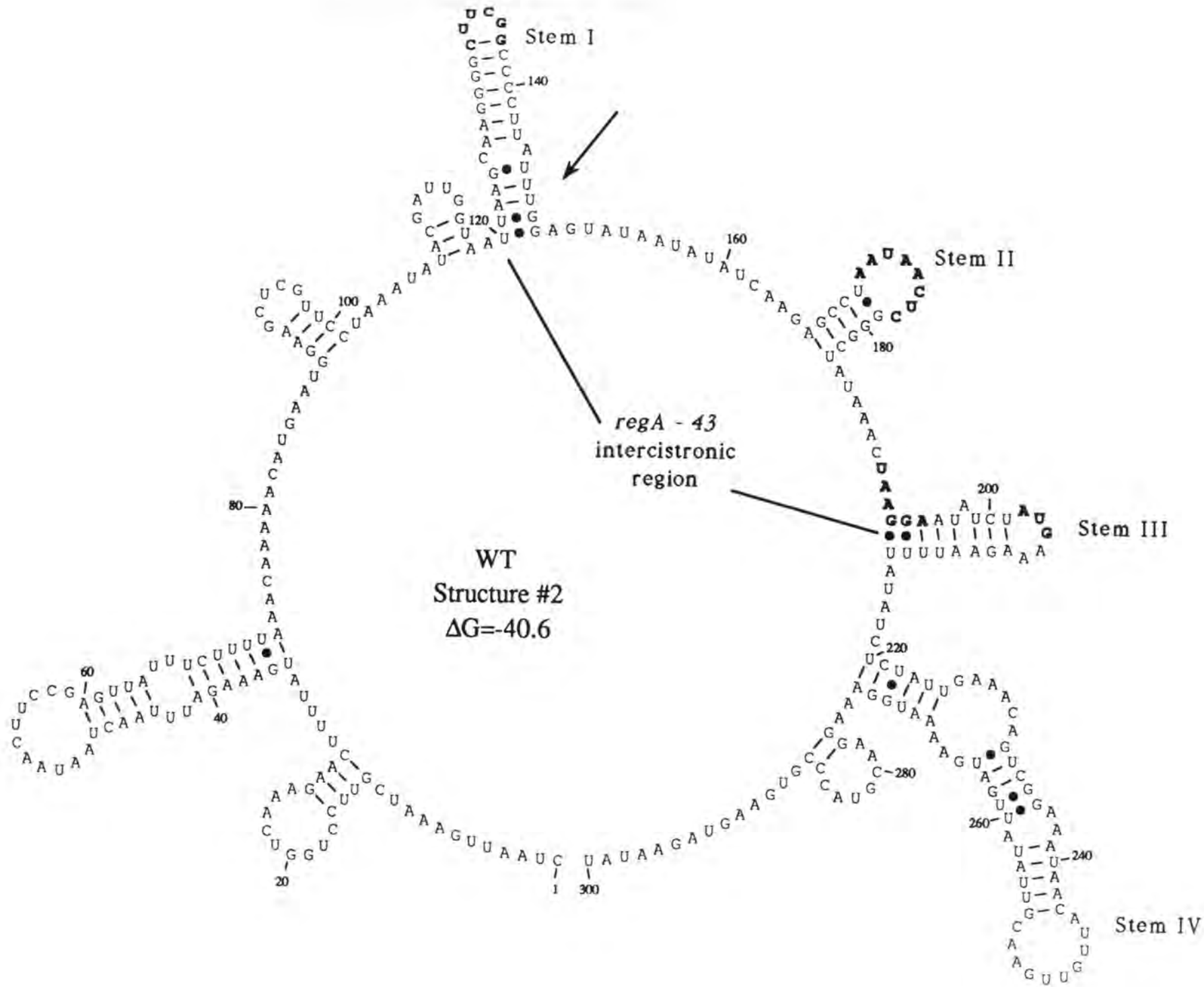
REGA-43 INTERCISTRONIC REGION

Figure 36 (two pages) depicts two possible RNA secondary structures predicted for the *regA-43* intercistronic region using the newest version of Mulfold, which also predicts sub-optimal folding structures (Jaeger, et al., 1990; Zuker, 1989). Both the upstream and operator hairpin loops are predicted with different programs and with varying lengths of upstream and downstream sequences included in the analysis. Therefore, the predictions of each of these local secondary structures are convincing. Also, these structures have been confirmed by RNase digestion experiments (Andrake, et al., 1988). Thirdly, the predictions are confirmed by the experimental data relating the function of this structure to gp43 binding.

It seems that the nucleotide sequence and resultant three-dimensional conformation of this operator region have been carefully selected and highly refined to perform its function. Perhaps the most dramatic detail is that several single base changes in the loop can greatly affect expression. None of the loop mutants constructed for this dissertation work changed the predicted stem pairing of the mRNA folds. Note that the SD operator can still form a hairpin in the operator region with most of the same loop residues maintained in a smaller loop (Figure 37). The essential factor is that this smaller hairpin is predicted to

Figure 36: Secondary structure predicted for polycistronic g 43 mRNA.
Residues of the operator loop, the SD and intiator codons are bolded. In Structure #1, Stem I and the stem immediately following are the only changes between these alternate structures; stems II-IV remain identicle




Stem I
Stem II
Stem III
Stem IV
regA - 43
intercistronic
region
WT
Structure #2
ΔG=-40.6

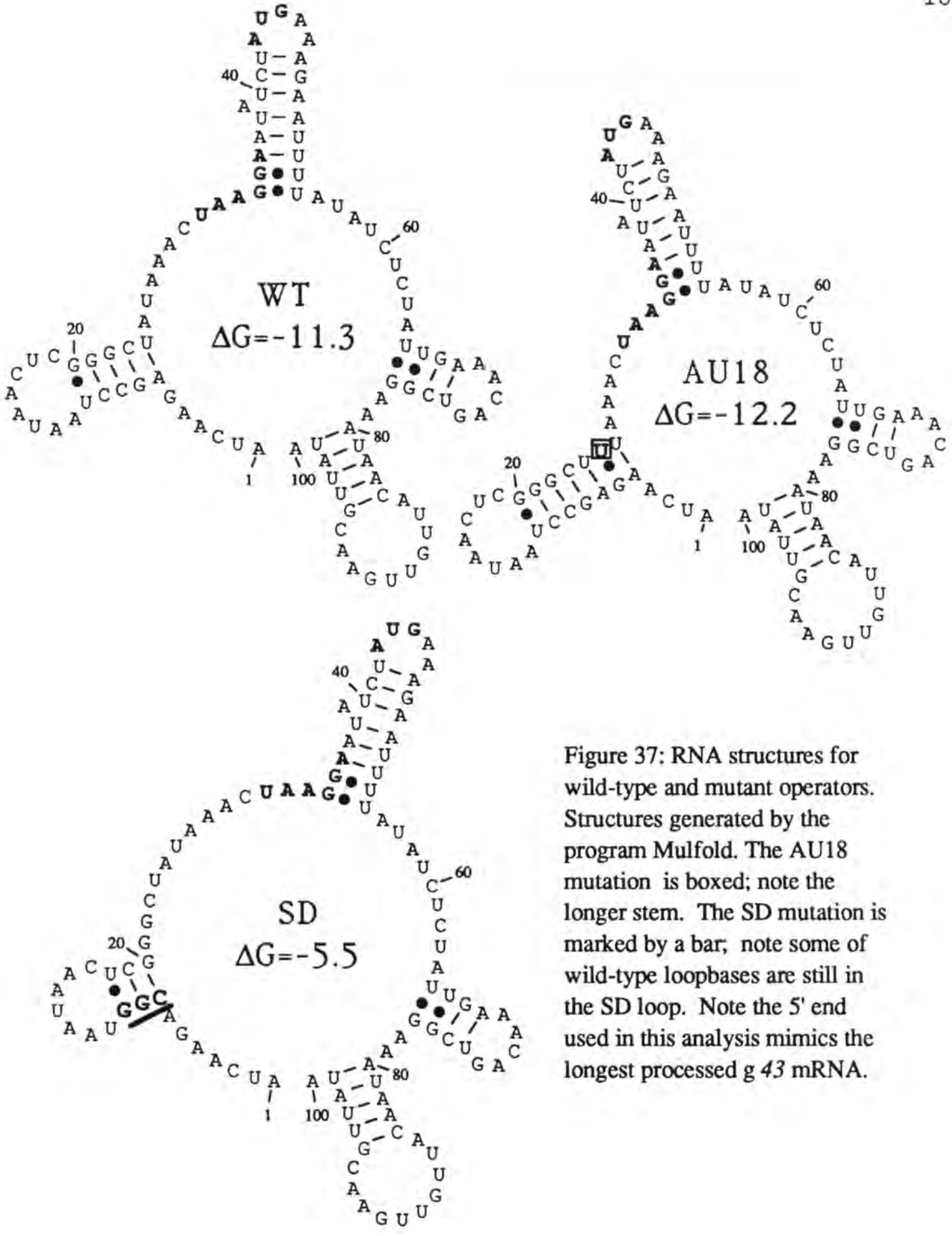


Figure 37: RNA structures for wild-type and mutant operators. Structures generated by the program Mulfold. The AU18 mutation is boxed; note the longer stem. The SD mutation is marked by a bar; note some of wild-type loopbases are still in the SD loop. Note the 5' end used in this analysis mimics the longest processed g *43* mRNA.

be considerably less stable. Given these thermodynamics, the SD RNA operator will spend statistically less time in the folded hairpin structure (rigid for presentation to the protein binding site) than the wild-type hairpin. As noted in the Introduction with respect to 30S subunit binding, this difference in free energy stability can be quite significant to the kinetics of binding. In contrast, the AU18 operator is predicted to have a longer stem with higher stability than WT; yet, as shown in Results this operator mutant while binding as well as wild-type *in vitro*, is still derepressed slightly *in vivo*.

There are two additional factors which confirm that the gp43 operator is a highly refined regulatory element: 1) Comparison of the analogous region in several other T-even phage indicates little variability in this region which does not encode amino acids. If these corresponding phage DNA polymerases also regulate their translation, utilizing similar protein RNA interactions, we might expect this conservation. 2) An *in vitro* selection procedure that enriches for loop sequences that bind gp43 with high affinity selected for only minor variations in loop residues (Tuerk and Gold, 1990), most remaining conserved through this selection process.

C) RNA STRUCTURE AND RECOGNITION OF RNA BY PROTEINS

Proteins seem to recognize a combination of structure

and specific nucleotide residues when binding to a specific RNA site. Interaction with single-stranded nucleotides may be more common, since unpaired regions present more functional groups for hydrogen bonding with amino acid side chains on the RNA binding protein. However, the three-dimensional structure of A-form RNA helices may also provide functional groups for hydrogen bonding with amino acids, as is found in double-stranded DNA. The stability of folding for various TIRs may also play an important role in protein binding. Some RNA-protein interactions may depend on a very stable RNA structure that statistically spends most of its time in a specific folded conformation (i.e. gp43, R17 coat protein, ribosomal proteins S4 and L1, gp32, ThrSase). In the gp43 operator, the calculated decrease in stability of the SR operator (relative to the WT operator) may account for the derepression of synthesis *in vivo*. More studies to explore the relationship of folding stability to gp43 binding affinity will be needed before definite conclusions can be made. In contrast, other RNA-protein interactions may require an open TIR where the specific nucleotides that bind protein are located in an unstructured region (T4 gpregA, f1 gpV). Here, the introduction of secondary structure may interfere with protein binding.

It is clear that a better understanding of these interactions will require more detailed knowledge of mRNA structure in solution. RNA structure modeling will play the same important role that has in studies of protein structure.

In addition, detailed investigation of the protein determinants of RNA binding will be essential to understanding the specificity protein-RNA interactions.

D) POTENTIAL RNA BINDING MOTIFS WITHIN GP43

EUKARYOTIC RNP CONSENSUS

In Eukaryotes, the ribonucleoprotein (RNP) binding domain consists of an approximately 90 amino acid region, shown to be conserved in several proteins that interact with RNA (Bandziulis, et al., 1989; Query, et al., 1989). Most are involved in mRNA biogenesis in eukaryotes (hnRNP A1, scRNP, polyA binding protein, U1-70K protein, nucleolin, and others). The motif was first described in polyA binding protein where the domain is repeated several times within the protein (reviewed in Dreyfuss, 1988). Within the large domain are two short areas of a higher degree of conservation that have been termed the RNP-1 and RNP-2 consensus. Conserved phenylalanines in these two regions have been shown to cross-link directly to RNA (Merrill, et al., 1988). Since many of these proteins may recognize RNA in a non-site-specific manner, the RNP domain may differ significantly from other protein motifs responsible for site-specific RNA binding. In fact, there is no perfect match for the entire RNP-1 octamer in gp43 (only half sites), while there are matches for the RNP-2 consensus. One RNP-2 is located up from an RNP-1 half site, yet separating the conserved aromatic residues (shown to cross-link to RNA) half the

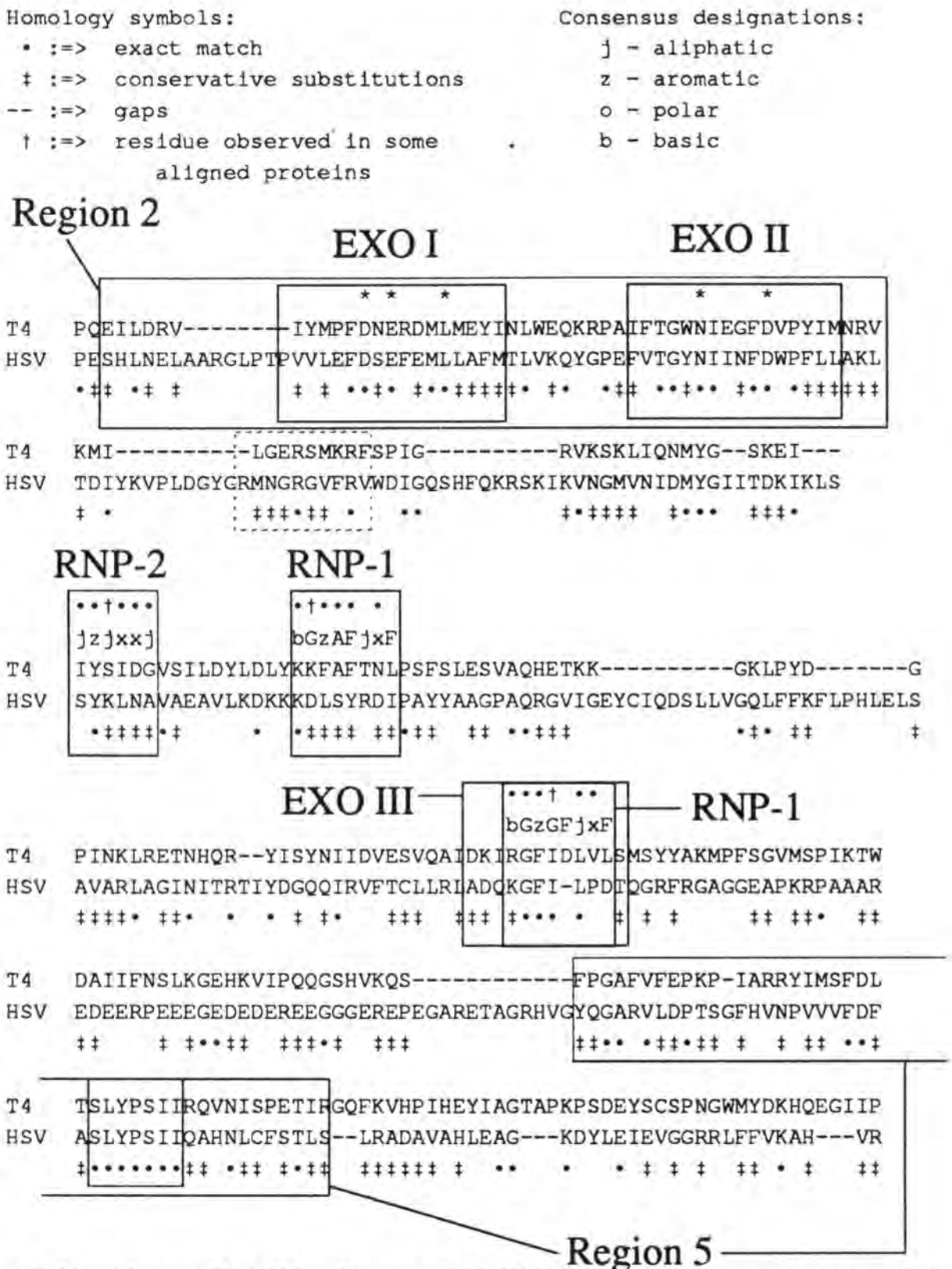


Figure 38: Homology of T4 DNA polymerase with HSV DNA polymerase and other protein motifs. A key is given at the top to explain homology and consensus symbols; symbols below the aligned sequences apply to T4-HSV homology and symbols above the consensus motifs indicate T4 homology to a given motif. Alignment of proteins was done with the Clustal program (Higgins and Sharp, 1988). Asterisks above the residues in EXO I and EXO II are highly conserved residues for the exonuclease function (Bernad, *et al.* 1989). Regions 2 and 5 are the previously noted homologies for many viral and eukaryotic DNA polymerases (Spicer *et al.*, 1988). The dashed box highlights an Arg-rich motif (half-site).

distance found in the RNP domain. This region occurs between the three regions shown to have homology to the exonuclease domain of *E. coli* DNA pol I (marked EXO I, II, and III -see Figure 37). Residues in these regions have been shown to be essential to various polI activities: 1) metal binding, which is essential for enzymatic activity (D and E residues in EXO I) 2) dTMP binding and 3) single-stranded DNA binding and binding to DNA 3′ ends. The best match for the RNP-1 consensus (6/8 residues match) is directly in the center of the EXO III domain of homology. The close approximation of these two conserved sequence motifs raises the question of the possible overlap of these functional domains (RNA binding and 3′-5′exonuclease) in the protein. Here again, the protein determinants for ssDNA (3′end binding - an aspect of exonuclease activity) and RNA binding may be similar. Alternatively, the preference for basic and aromatic residues may reflect exonuclease function alone, and not potential RNA binding motifs found in this region.

ARG-RICH AND OTHER PROTEIN MOTIFS

Another conserved region of several prokaryotic antiterminator proteins has been shown to be involved in RNA binding and has been termed the ARG-rich motif (Lazinski, et al., 1989). Interestingly, the motif is also found in some eukaryotic retrovirus regulator proteins known to interact with RNA. This motif contains a high proportion of basic

residues also, but in contrast to the RNP consensus not aromatic residues. Again, there is no perfect match for this motif in gp43, although there are several runs of basic residues within the gp43 sequence. A region of some homology is shown in Figure 38 - (see the dotted box).

The as yet unidentified RNA binding domain of T4 DNA polymerase may represent a motif of site specific RNA binding repressors, and not bear any resemblance to these known RNA binding motifs. It is worth noting that preliminary protein alignments of prokaryotic translational repressor proteins did not yield any regions of extensive homology. The T4 DNA polymerase sequence was not found to contain any of the ribosomal protein signature motifs which may be relevant to RNA binding, especially since most of these proteins have site specific interactions with rRNA. Yet none of the known 10 protein motifs had a match in gp43 sequence. gp43 does contain (at position 616) the short sequence YGTDT which is highly conserved in DNA polymerases (Argos, 1988), and a putative nuclear localization signal (Gomez-Marquez and Segade, 1988) at position 295.

E) REGULATION OF THE BIOSYNTHESIS OF DNA REPLICATION ENZYMES

Surprisingly, T4 DNA polymerase shares more homology with animal virus DNA polymerases (Herpes Simplex(HSV) and

Epstein-Barr (EBV) Viruses) than with *E. coli* DNA polymerase I and has little or no similarity to phage T7 polymerase (Spicer, et al., 1988). Alignment of the sequences of T4 and HSV DNA polymerases reveals several regions of significant homology, the most significant of which were highlighted above (Figure 39). Interestingly, preliminary evidence suggests that the HSV polymerase translation may be repressed during infection (Yager, et al., 1990). These results show that polymerase mRNA is inefficiently translated in infection, while equal amounts of RNA from before and after repression directed equivalent amounts of polymerase synthesis, *in vitro*. Whatever the final details of the mechanism of regulation, it is interesting to consider the similarities in how vastly different organisms regulate DNA replication. Apparently, early shut off of HSV polymerase translation is part of the regulatory mechanism guiding genome duplication in this animal virus. Yager *et al.* also postulate a role for polymerase translational regulation in the fine tuning the stoichiometries of replication proteins and/or finer temporal control of polymerase expression. The *E. coli* dnaG operon, which contains genes for ribosomal proteins, a DNA replication initiation protein, a sigma subunit of RNA polymerase is also reminiscent of the regulation of the T4 replication gene cluster (Hsu and Karam, 1990) studied in our lab. A variety of regulatory mechanisms are employed to control gene expression, including multiple promoter utilization (responsive to varying cellular

conditions), transcription termination, and RNA processing. The final dosage and temporal regulation of the gene products for these (and many other) operons are the result of the unique combination and coordination of diverse transcriptional and post-transcriptional regulatory mechanisms in the physiology of the cell.

V) LIST OF REFERENCES

Abrahams, J. P., M. van den Berg, E. van Batenburg and C. W. A. Pleij. "Prediction of RNA secondary structure, including pseudoknotting, by computer simulation."; Nucleic Acids Res. 18; 3035-3044, 1990.

Altuvia, S., G. H. Locker, S. Koby, N. O. Ben and A. B. Oppenheim. "RNase III stimulates the translation of the cIII gene of bacteriophage lambda."; Proc Natl Acad Sci U S A 84; 6511-5, 1987.

Andrake, M., N. Guild, T. Hsu, L. Gold, C. Tuerk and J. Karam. "DNA polymerase of bacteriophage T4 is an autogenous translational repressor."; Proc. Natl. Acad. Sci. 85; 7942-7946, 1988.

Argos, P. "A sequence motif in many polymerases."; Nucleic Acids Res. 16; 9909-9916, 1988.

Atkins, J. F., R. B. Weiss and R. F. Gesteland. "Ribosome gymnastics--degree of difficulty 9.5, style 10.0."; Cell 62; 413-23, 1990.

Bandziulis, R. J., M. S. Swanson and G. Dreyfuss. "RNA-binding proteins as developmental regulators."; Genes Dev 3; 431-7, 1989.

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

Bernad, A., L. Blanco, J. M. Lazaro, G. Martin and M. Salas. "A conserved 3'-5' exonuclease active site in prokaryotic and eukaryotic DNA polymerases."; Cell 59; 219-228, 1989.

Blumberg, D. D. "Creating a ribonuclease-free environment."; Methods Enzymol. 152; 20-24, 1987.

Breidt, F. and D. Dubnau. "Identification of cis-acting sequences required for translational autoregulation of the ermC methylase."; J Bacteriol 172; 3661-8, 1990.

Butler, J. S., M. Springer and M. Grunberg-Manago. "AUU-to-AUG mutation in the initiator codon of the translation initiation factor IF3 abolishes translational autocontrol of its own gene (*infC*) *in vivo*."; Proc. Natl. Acad. Sci. USA 84; 4022-4025, 1987.

Calogero, R. A., C. L. Pon, M. A. Canonaco and C. O. Gualerzi. "Selection of the mRNA translation initiation region by Escherichia coli ribosomes."; Proc Natl Acad Sci U S A 85; 6427-31, 1988.

Cammack, K. A. and H. E. Wade. "The sedimentation behavior of RNase -active and -inactive ribosomes from bacteria."; Biochemical J 96; 671-680, 1965.

Davis, P. W., R. W. Adamiak and I. J. Tinoco. "Z-RNA: the solution NMR structure of r(CGCGCG)."; Biopolymers 29; 109-22, 1990.

Devereux, J., P. Haeberli and O. Smithies. "A comprehensive sequence analysis package for VAX computers."; Nucleic Acids Res. 12; 387-395, 1984.

Dreyfus, M. "What constitutes the signal for the initiation of protein synthesis on Escherichia coli mRNAs?"; J Mol Biol 204; 79-94, 1988.

Fuchs, R. "MacPattern: protein pattern searching on the Apple Macintosh."; CABIOS in press; 1990.

Ganoza, M. C., P. Marliere, E. C. Kofoid and B. G. Louis. "Initiator tRNA may recognize more than the initiation codon in mRNA: a model for translational initiation."; Proc Natl Acad Sci U S A 82; 4587-91, 1985.

Gerald, W. L. and J. D. Karam. "Expression of a DNA Replication Gene Cluster in Bacteriophage T4: Genetic Linkeage and the Control of Gene Product Interactions."; Genetics 107; 537-549, 1984.

Gold, L. "Posttranscriptional regulatory mechanisms in Escherichia coli."; Annu Rev Biochem. 57; 199-233, 1988.

Gold, L., D. Pribnow, T. Schneider, S. Shinedling and B. Singer. "Translation initiation in prokaryotes"; Ann. Rev. Microbiol. 35; 365-403, 1981.

Gomez-Marquez, J. and F. Segade. "Prothymosin alpha is a nuclear protein."; FEBS Lett. 226; 217-219, 1988.

Guild, N., M. Gayle, R. Sweeney, T. Hollingsworth and T. Modeer. "Transcriptional Activation of Bacteriophage T4 Middle Promoters by the *motA* Protein."; J. Mol. Biol. 199; 241-258, 1988.

Hahn, J., G. Grandi, T. J. Gryczan and D. Dubnau. "Translational attenuation of *ermC*: a deletion analysis."; Mol. Gen. Genet. 186; 204-216, 1982.

Hanvey, J. C., M. Shimizu and R. D. Wells. "Intramolecular DNA triplexes in supercoiled plasmids."; Proc. Natl. Acad. Sci. USA 85; 6292-6296, 1988.

Higgins, D. G. and P. M. Sharp. "CLUSTAL: a package for performing multiple sequence alignment on a microcomputer."; Gene 73; 237-244, 1988.

Hsu, T. and J. D. Karam. "Transcriptional mapping of a DNA replication gene cluster in bacteriophage T4."; J. Biol. Chem. 265; 5303-5316, 1990.

Hsu, T., R. Wei, M. Dawson and J. D. Karam. "Identification of two new bacteriophage T4 genes that may have roles in transcription and DNA replication."; J. Virol. 61; 366-374, 1987.

Huang, W. M., S. Z. Ao, S. Casjens, R. Orlandi, R. Zeikus, R. Weiss, D. Winge and M. Fang. "A persistent untranslated sequence within bacteriophage T4 DNA topoisomerase gene 60."; Science 239; 1005-12, 1988.

Hui, A. and H. A. de Boer. "Specialized ribosome system: preferential translation of a single mRNA species by a subpopulation of mutated ribosomes in Escherichia coli."; Proc Natl Acad Sci U S A 84; 4762-6, 1987.

Inouye, M. "Antisense RNA: its functions and applications in gene regulation--a review."; Gene 72; 25-34, 1988.

Jacob, W. F., M. Santer and A. E. Dahlberg. "A single base change in the Shine-Dalgarno region of 16S rRNA of Escherichia coli affects translation of many proteins."; Proc Natl Acad Sci U S A 84; 4757-61, 1987.

Jaeger, J. A., D. H. Turner and M. Zuker. "Improved predictions of secondary structure for RNA."; Proc. Natl. Acad. Sci. USA 86; 7706-7710, 1989.

Jaeger, J. A., D. H. Turner and M. Zuker. "Predicting optimal and suboptimal secondary structure for RNA."; Methods Enzymol. 183; 281-330, 1990.

Jaeger, J. A., D. H. Turner and M. Zuker. "Predicting optimal and suboptimal secondary structure for RNA."; Methods Enzymol. 183; 281-330, 1990.

Jarvis, T. C., D. M. Ring, S. S. Daube and P. H. von Hippel. ""Macromolecular crowding": Thermodynamic consequences for protein-protein interactions within the T4 DNA replication complex."; J. Biol. Chem. 265; 15160-15167, 1990.

Karam, J. D. and M. G. Bowles. "Mutation to Overproduction of Bacteriophage T4 Gene Products."; J. Virol. 13; 428-438, 1974.

Karam, J. D., C. McCulley and M. Leach. "Genetic control of mRNA decay in T4 phage-infected Escherichia coli."; Virology 76; 685-700, 1977.

Kozak, M. "Comparison of initiation of protein synthesis in prokaryotes, eukaryotes, and organelles."; Microbiol. Rev. 47; 1-45, 1983.

Krieg, P. A. and D. A. Melton. "In vitro RNA synthesis with SP6 RNA polymerase."; Methods Enzymol 155; 397-415, 1987.

Krisch, H. M., G. Van Houwe, D. Belin, W. Gibbs and R. H. Epstein. "Regulation of the expression of bacteriophage T4 genes 32 and 43."; Virology 78; 87-98, 1977.

Lazinski, D., E. Grzadzielska and A. Das. "Sequence-Specific Recognition of RNA Hairpins by Bacteriophage Antiterminators Requires a Conserved Arginine-Rich Motif."; Cell 59; 207-218, 1989.

Liang, Y. M., R. X. Wei, T. Hsu, C. Alford, M. Dawson and J. Karam. "Autogenous regulation of the regA gene of bacteriophage T4: derepression of translation."; Genetics 119; 743-749, 1988.

Lindahl, L. and J. M. Zengel. "Ribosomal genes in Escherichia coli."; Annu Rev Genet 1986.

Mans, R. M. W., C. Guerrier-Takada, S. Altman and C. W. A. Pleij. "Interaction of RNase P from Escherichia coli with pseudoknotted structures in viral RNAs."; Nuc. Acids Res. 18; 3479-3487, 1990.

Martinez, H. M. "An RNA folding rule."; Nucleic Acids Res. 12; 323-334, 1984.

Mazur, B. J. and N. D. Zinder.Virology 68; 490-502, 1975.

McCarthy, J. E. and C. Gualerzi. "Translational control of prokaryotic gene expression."; Trends Genet. 6; 78-85, 1990.

McClure, W. R. "Mechanism and control of transcription initiation in prokaryotes."; Annu Rev Biochem 1985.

McPheeters, D. S., G. D. Stormo and L. Gold. "Autogenous regulatory site on the bacteriophage T4 gene 32 messenger RNA."; J. Mol. Biol. 201; 517-535, 1988.

Merrill, B. M., K. L. Stone, F. Cobianchi, S. H. Wilson and K. R. Williams. "Phenylalanines that are Conserved among several RNA-binding Proteins Form part of a Nucleic Acid-binding Pocket in the A1 Heterogeneous Nuclear Ribonucleoprotein."; Journal of Biological Chemistry 263; 3307-3313, 1988.

Michel, B. and N. D. Zinder. "In vitro binding of the bacteriophage f1 gene V protein to the gene II RNA-operator and its DNA analog."; Nucleic Acids Res 17; 7333-44, 1989.

Michel, B. and N. D. Zinder. "Translational repression in bacteriophage f1: characterization of the gene V protein target on the gene II mRNA."; Proc Natl Acad Sci U S A 86; 4002-6, 1989.

Miller, R. C., E. T. I. Young, R. H. Epstein, H. M. Krisch and A. Bolle. "Regulation of the synthesis of the T4 DNA polymerase (Gene 43)."; 110; 98-112, 1981.

Moazed, D. and H. F. Noller. "Transfer RNA shields specific nucleotides in 16S ribosomal RNA from attack by chemical probes."; Cell 47; 985-94, 1986.

Moazed, D., S. Stern and H. F. Noller. "Rapid Chemical Probing of Conformation in 16 S Ribosomal RNA and 30 S Ribosomal Subunits Using Primer Extension."; J. Mol. Biol. 187; 399-416, 1986.

Model, P., C. McGill, B. Mazur and W. Fulford. "Replication of bacteriophage f1: gene V protein regulates the synthesis of gene II protein."; Cell 29; 329-335, 1982.

Murgola, E. J., K. A. Hijazi, H. U. Goringer and A. E. Dahlberg. "Mutant 16S ribosomal RNA: a codon-specific translational suppressor."; Proc Natl Acad Sci U S A 85; 4162-5, 1988.

Nomura, M., R. Gourse and G. Baughman. "Regulation of the synthesis of ribosomes and ribosomal components."; Ann. Rev. Biochem. 53; 75-117, 1984.

Nussinov, R. and I. J. Tinoco. "Sequential Folding of a Messenger RNA Molecule."; J. Mol. Biol. 151; 519-533, 1981.

Olins, P. O. and S. H. Rangwala. "A novel sequence element derived from bacteriophage T7 mRNA acts as an enhancer of translation of the lacZ gene in Escherichia coli."; J Biol Chem 264; 16973-6, 1989.

Peabody, D. S. "Translational repression by bacteriophage MS2 coat protein does not require cysteine residues."; Nucleic Acids Res 17; 6017-27, 1989.

Petersen, C. "Long-range translational coupling in the rplJL-rpoBC operon of Escherichia coli."; J Mol Biol 206; 323-32, 1989.

Petersen, G. B., P. A. Stockwell and D. F. Hill. "Messenger RNA recognition in Escherichia coli: a possible second site of interaction with 16S ribosomal RNA."; Embo J 7; 3957-62, 1988.

Pleij, C. W. A. "Pseudoknots: a new motif in the RNA game."; Trends Biol. Sci. 15; 143-147, 1990.

Pratt, J. M. "Coupled transcription-translation in prokaryotic cell-free systems." In: Transcription and Translation: a Practical Approach, edited by B.D. Hames and S.J. Higgins. Washington, DC: IRL, 1984, p. 179-209

Puglisi, J. D., J. R. Wyatt and I. J. Tinoco. "Conformation of an RNA pseudoknot."; J Mol Biol 214; 437-53, 1990.

Puglisi, J. D., J. R. Wyatt and I. J. Tinoco. "Solution conformation of an RNA hairpin loop."; Biochemistry 29; 4215-26, 1990.

Query, C. C., R. C. Bentley and J. D. Keene. "A Common RNA Recognition Motif Identified within a Defined U1 RNA Binding Domain of the 70K U1 snRNP Protein."; Cell 57; 89-101, 1989.

Rietveld, K., R. van Poelgeest, C. W. A. Pleij, J. H. van Boom and L. Bosch. "The tRNA-like structure at the 3' terminus of turnip yellow mosaic virus RNA. Differences and similarities with cannonical tRNA."; Nucleic Acids Res. 10; 1929f, 1982.

Romaniuk, P. J. and O. C. Uhlenbeck. "Nucleoside and nucleotide inactivation of R17 coat protein: evidence for a transient covalent RNA-protein bond."; Biochemistry 24; 4239-44, 1985.

Ruckman, J., D. Parma, C. Tuerk, D. H. Hall and L. Gold. "Identification of T4 gene required for Bacteriophage mRNA processing."; 1; 54-65, 1989.

Russel, M. "Control of Bacteriophage T4 DNA Polymerase Synthesis."; J. Mol. Biol. 79; 83-94, 1973.

Schauder, B. and J. E. McCarthy. "The role of bases upstream of the Shine-Dalgarno region and in the coding sequence in the control of gene expression in Escherichia coli: translation and stability of mRNAs in vivo."; Gene 78; 59-72, 1989.

Schleif, R. "DNA binding by proteins."; Science 241; 1182-7, 1988.

Schneider, T. D., G. D. Stormo, L. Gold and A. Ehrenfeucht. "Information content of binding sites on nucleotide sequences."; J Mol Biol 188; 415-31, 1986.

Shine, J. and L. Dalgarno. "The 3' terminal sequence of Escherichia coli 16S ribosomal RNA: Complementarity to nonsense triplets and ribosome binding sites."; Proc. Natl. Acad. Sci. USA 71; 1342-1346, 1974.

Spicer, E. K., J. Rush, C. Fung, L. J. Reha-Krantz, J. D. Karam and W. H. Konigsberg. "Primary structure of T4 DNA polymerase."; Journal of Biological Chemistry 263; 7478-7486, 1988.

Springer, M., M. Graffe, J. Dondon and M. M. Grunberg. "tRNA-like structures and gene regulation at the translational level: a case of molecular mimicry in Escherichia coli."; Embo J 8; 2417-24, 1989.

Springer, M., M. Graffe, J. Dondon, M. M. Grunberg, P. Romby, B. Ehresmann, C. Ehresmann and J. P. Ebel. "Translational control in E. coli: the case of threonyl-tRNA synthetase."; Biosci Rep 8; 619-632, 1988.

Springer, M., M. Graffe, J. S. Butler and M. Grunberg-Manago. "Genetic Definition of the Translational Operator of the Threonine-tRNA Ligase Gene in Escherichia coli."; Proc. Natl. Acad. Sci. 83; 4384-4388, 1986.

Steitz, J. A. In: Ribosomes, edited by G. Chambliss, G.R. Craven, J. Davies, K. Davis, L. Kahan and M. Nomura. Baltimore: University Park Press, 1979, p.

Tang, C. K. and D. E. Draper. "Unusual mRNA pseudoknot structure is recognized by a protein translational repressor."; Cell 57; 531-6, 1989.

Tapprich, W. E., D. J. Goss and A. E. Dahlberg. "Mutation at position 791 in Escherichia coli 16S ribosomal RNA affects processes involved in the initiation of protein synthesis."; Proc Natl Acad Sci U S A 86; 4927-31, 1989.

Tchurikov, N. A., B. K. Chernov, Y. B. Golova and Y. D. Nechipurenko. "Parallel DNA: generation of a duplex between two Drosophila sequences in vitro."; Febs Lett 257; 415-8, 1989.

Thanaraj, T. A. and M. W. Pandit. "An additional ribosome-binding site on mRNA of highly expressed genes and a bifunctional site on the colicin fragment of 16S rRNA from Escherichia coli: important determinants of the efficiency of translation-initiation."; Nucleic Acids Res 17; 2973-85, 1989.

Trifonov, E. N. "Translation framing code and frame-monitoring mechanism as suggested by the analysis of mRNA and 16 S rRNA nucleotide sequences."; J Mol Biol 194; 643-52, 1987.

Trojanowska, M., E. S. Miller, J. Karam, G. Stormo and L. Gold. "The bacteriophage T4 regA gene: primary sequence of a translational repressor."; Nuc. Acids Res. 12; 5979-5993, 1984.

Tuerk, C. and L. Gold. "Systematic evolution of ligands by exponential enrichment: RNA ligands to bacteriophage T4 DNA polymerase."; Science 249; 505-510, 1990.

Unnithan, S., L. Morrissey, J. Rinkley, J. Karam and L. Gold. "Binding of the bacteriophage T4 regA protein to mRNA targets: An initiator AUG is required."; Nucleic Acids Res. 8; 7083-7092, 1990.

Uzan, M., R. Favre and E. Brody. "A nuclease that cuts specifically in the ribosome binding site of some T4 mRNAs."; Proc Natl Acad Sci U S A 85; 8895-9, 1988.

Venkatesan, M. and N. G. Nossal. "Bacteriophage T4 gene 44/62 and gene 45 polymerase accessory proteins stimulate hydrolysis of duplex DNA by T4 DNA polymerase."; J. Biol. Chem. 257; 12435-12443, 1982.

Vorlickova, M., J. Kypr, T. M. Jovin and M. Planck. "CD of the synthetic RNA duplexes poly[r(A-T)] and poly[r(A-U)] in salt and ethanolic solutions."; Biopolymers 29; 385-92, 1990.

Webster, K. R., H. Y. Adari and E. K. Spicer. "Bacteriophage T4 regA protein binds to the Shine-Dalgarno region of gene 44 mRNA."; Nucleic Acids Res 17; 10047-68, 1989.

Weiss, R. B., D. M. Dunn, A. E. Dahlberg, J. F. Atkins and R. F. Gesteland. "Reading frame switch caused by base-pair formation between the 3' end of 16S rRNA and the mRNA during elongation of protein synthesis in Escherichia coli."; Embo J 7; 1503-7, 1988.

Winter, R., L. Morrissey, P. Gauss, L. Gold, T. Hsu and J. Karam. "Bacteriophage T4 regA protein binds to mRNAs and prevents translation initiation."; Proc Natl Acad Sci U S A. 84; 7822-7826, 1987.

Wulczyn, F. G., M. Bolker and R. Kahmann. "Translation of the bacteriophage Mu mom gene is positively regulated by the phage com gene product."; Cell 57; 1201-10, 1989.

Yager, D. R., A. I. Marcy and D. M. Coen. "Translational regulation of herpes simplex virus DNA polymerase."; J Virol 64; 2217-25, 1990.

Yen, B. T. and R. E. Webster. "Translational control of bacteriophage f1 gene II and gene X proteins by gene V protein."; Cell 29; 337-345, 1982.

Zuker, M. and P. Stiegler. "Optimal computer folding of large RNA sequences using thermodynamics and auxiliary information."; Nucleic Acids Res. 9; 133-148, 1981.

Zuker, M. "On Finding All Suboptimal Foldings of an RNA Molecule."; Science 244; 48-52, 1989.