Proc. Natl. Acad. Sci. USA
Vol. 78, No. 8, pp. 4669–4673, August 1981
Biochemistry

Translational regulation: Identification of the site on bacteriophage T4 *r*IIB mRNA recognized by the *regA* gene function

(translational repression/RNA sequence recognition)

JIM KARAM*, LARRY GOLD†, BRITTA SWEBILIUS SINGER†, AND MYRA DAWSON*

*Department of Biochemistry, Medical University of South Carolina, Charleston, South Carolina 29403; and †Department of Molecular, Cellular and Developmental Biology, University of Colorado, Boulder, Colorado 80309

Communicated by William B. Wood, December 15, 1980

ABSTRACT **The bacteriophage T4 gene *regA* encodes a protein that diminishes the expression of many unlinked early T4 genes. Previous work demonstrated that regA-mediated repression occurs after transcription. We report here on the identification of the target site on one regA-sensitive mRNA, the message encoding the phage T4 *r*IIB protein. The target for regA-mediated action overlaps the translational initiation domain of the *r*IIB messenger. The *regA* protein may be a repressor that operates translationally on a significant and interesting set of early phage T4 mRNAs.**

Evidence has accumulated over the last several years for control of prokaryotic gene expression beyond the level of transcription. Several factors may influence the translational yield from specific transcripts: discrimination by ribosomes in selecting one transcript over others, regulation of translation of particular mRNAs by repressors or activators, and destruction of transcripts at nonuniform rates. Direct evidence exists (in different systems) for all but one of these control mechanisms (1). The intrinsic strength of a ribosome binding site (refs. 1–7) determines the constitutive‡ level of specific translation, whereas regulation might encompass a reversible and specific repression or activation of translation through the intervention of a regulatory molecule. Alteration of transcripts by an RNase may be specific and, therefore, an example of irreversible regulation (8, 9). Reports of regulation of translation are rare: (*i*) the *Escherichia coli* RNA phages utilize translational regulation (10–14), (*ii*) the bacteriophage T4 controls the expression of its major DNA binding protein through autogenous translational regulation (15–18), and (*iii*) translational regulation is responsible for the careful titration of the quantities of several ribosomal proteins in *E. coli* (19–22). Compared with the abundant descriptions of transcriptional regulation, the literature on prokaryotic translational control is not vast.

Recently, a phage T4 function was discovered that plays a role in control of the translation of a number of phage transcripts. Mutations in the *regA* gene of T4 lead to overproduction of a small number of the proteins that are synthesized during the early stages of phage development (23–27). The effects of these mutations seem to be mediated by loss of a diffusible substance, probably a protein (26). Some observations suggest that the *regA* protein functions at a posttranscriptional level. T4 *regA* mutations do not affect RNA synthesis, although the protein overproduction of T4 *regA*$^-$ phage infections is accompanied by a prolonged functional lifetime for the corresponding mRNA (and by decreased breakdown of some phage RNA) (23, 24, 27, 28).

It is not known how the phage T4 *regA* protein represses translation of several early T4 mRNAs. Possibly, this protein initiates mRNA breakdown directly—e.g., by modifying or cleaving certain messages. Alternatively, translation of some early T4 mRNAs could be directly repressed by the *regA* protein, and mRNA decay (physical or merely functional) may ensue as a consequence of that repression. Either type of mechanism for T4 *regA* protein action would implicate the participation of specific sequence(s) in mRNA because the biological activities of most T4-induced early transcripts are not affected by *regA* mutations.

The T4 *r*IIB gene product is one of the phage proteins that are hyperproduced in T4 *regA*$^-$ phage infections. We tested several precisely mapped missense, nonsense, frameshift, and deletion mutations of the *r*IIB gene (7, 29–31) for effects that mimic the effects of *regA* mutations on *r*IIB protein synthesis—i.e., for inactivation of a presumed regA target site. We found several *r*IIB lesions that produced *r*IIB protein almost equally in *regA*$^+$ and *regA*$^-$ genetic backgrounds. These mutations all mapped within a short region of known nucleotide sequence (31, 32) that comprises the first three codons (including the initiation codon) of *r*IIB mRNA. From the size and location of the sequence, we suggest that the *regA* protein, a polypeptide of about 80 amino acids (26), exerts its primary effects at the level of initiation of translation. It is not clear, however, if the sequence is a target site for direct interaction between the *regA* protein and *r*IIB mRNA.

MATERIALS AND METHODS

Phage and Bacteria. Many of the phage T4 *r*II mutants came from S. Champe and S. Brenner. Some were used by us in related work (7). The T4 *regAR9* mutation has been described (23). Other phage mutants came from W. B. Wood and R. S. Edgar. Multiple phage mutants were constructed by genetic crosses and identified by standard complementation and recombination tests (33). *E. coli* CR63, which permits growth of T4 amber (*am*) mutants, was used as host for growing all phage stocks. *E. coli* CR63 (λ) restricts growth of *r*II mutants and was used for their identification. *E. coli* B^E (su^0), which restricts growth of phage T4 that carry nonsense mutations in essential genes, was used in all experiments involving the radioactive labeling of phage-encoded proteins for gel electrophoretic analyses.

Nucleotide Sequence in the Phage T4 *r*II Region. The nucleotide sequence for the region encompassing the T4 *r*IIA–*r*IIB gene junction was determined by Belin *et al.* (32) and by Pribnow *et al.* (31).

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

‡ The term "constitutive" is used to denote gene expression no longer under the control of a regulatory molecule. We know of no restraints against the use of this term in a translational context.

Measurement of Phage T4 rIIB Protein Synthesis. The methods for labeling phage T4-induced proteins with radioactive amino acids and for analysis of these proteins by polyacrylamide slab gel electrophoresis have been described (7, 24). The effects of T4 *regA*$^-$ mutations are most easily measured under conditions of inhibited late phage gene expression (24, 25, 28). To achieve such conditions for studying rIIB protein synthesis, phage strains were constructed that carried lesions in DNA replication functions [genes *32* (*32amA453*), *62* (*62amE1140*), and *44* (*44amN82*)] and in one or both of genes *33* (*33amN134*) and *55* (*33amBL292*), which are required for initiation of T4 late transcription (34, 35). Protein synthesis was measured at 30°C in M9 medium. Initially, a large number of T4 rIIB and rIIA mutations were screened in the *regA*$^+$ genetic background for effects on rII expression. In these screening studies, phage-infected cells were pulse labeled for 5 min with ^{14}C-labeled amino acids at a late time after infection (e.g., 30 min), and extracts were analyzed by $NaDodSO_4$/gel electrophoresis and autoradiography. Those rII mutants that exhibited higher rates of rII protein synthesis than was seen in control infections with rII$^+$ phage were examined more closely in *regA*$^-$ and *regA*$^+$ genetic backgrounds.

RESULTS

A Summary of the Effects of Phage T4 *regA*$^-$ Mutations on Phage-Induced Protein Synthesis. In normal infections of *E. coli* with phage T4, the synthesis of most early phage proteins ceases within a few minutes of the onset of DNA replication and late gene transcription (34, 36, 37). The shutoff of early expression and the transition to late gene expression are related to each other, although they are not strictly interdependent. When T4 late transcription is prevented (e.g., by mutations in phage genes 33 and 55 or by eliminating T4 DNA replication), overall early gene expression continues, albeit at a decreasing rate (34). Phenotypically, T4 *regA*$^-$ mutations result in virtual elimination of the shutoff of expression of certain early genes without detectably affecting the expression of other early genes; these regA-mediated effects are more conspicuous when late gene expression is restricted than when it is permitted (23, 27). Hyperproduction of many early proteins, mediated by a *regA* mutation, is accompanied by an increase in the functional lifetime of mRNA for the hyperproduced proteins (24, 27); however, these results are complex (see below). Early transcription, when it has been measured directly, seems unaltered by *regA* mutations (refs. 23, 27, and 28; unpublished data).

Fig. 1 shows examples of the specificity in regA-mediated hyperproduction of early proteins. Genes regulated by the *regA* function are interspersed with genes that are indifferent to this function; genetic linkage is not a factor in *regA* protein-mediated regulation (23–28, 38, 39–43). Because the *regA* protein appears to function after transcription, we think that this protein causes inhibition of translation (directly or indirectly) of only those T4 early transcripts that carry a certain RNA sequence or structure. Our results give credence to this notion.

Genetic Localization of a *regA*-Related Control Site on a Phage T4 mRNA. We used the phage T4 rIIA and rIIB genes to search for the presumed site on mRNA that responds to inhibition by *regA*$^+$ gene function. These two T4 genes are ideal for such studies because (*i*) they are well-characterized genetically (29–31), (*ii*) the nucleotide sequence surrounding and including the translation initiation region for the rIIB cistron is known (31, 32), and (*iii*) the rIIA and rIIB proteins are among those early proteins that are hyperproduced in phage T4 *regA*$^-$ infections (Fig. 1; refs. 23–27). We searched for rII mutations that rendered synthesis of one or the other of the rII proteins insensitive to *regA*$^+$ gene control. Such mutations might be expected to produce the relevant rII protein at the same rate

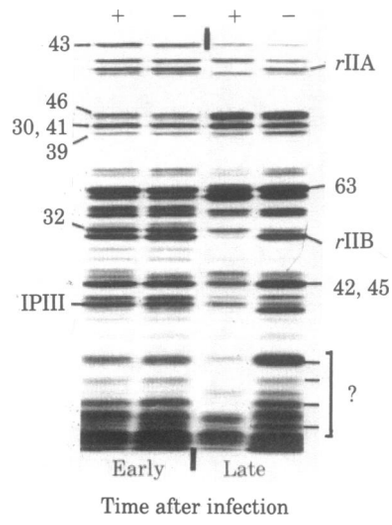

FIG. 1. Patterns of phage T4-induced protein synthesis after infections with *regA*$^+$ and *regAR9* T4 phage. The phage strains used carried nonsense mutations in genes *33*, *62*, *44*, and *55*. Samples of the infected *E. coli* B^E cultures were incubated with ^{14}C-labeled amino acids at an early time period (6–11 min) and at a late time period (31–36 min) after infection; extracts were prepared and analyzed by $NaDodSO_4$/gel electrophoresis and autoradiography (7, 24). The autoradiogram demonstrates the selective hyperproduction of several T4 early proteins during late time periods after infections with *regA*$^-$ phage. The identities of many of these gene proteins have been determined (24) and are indicated on the right-hand side of the autoradiogram; those gene proteins that are not hyper-produced in phage T4 *regA*$^-$ infections are indicated on the left-hand side.

in phage T4 *regA*$^+$ or *regA*$^-$ infections, that is, to be constitutive.‡

In preliminary screening studies, we used the types of assays in Fig. 1 to examine several rIIA and rIIB deletion mutations for constitutive phenotypes. The analyses were limited to those mutants that yielded detectable peptides, yet we were able to test directly many segments of the rIIA and rIIB cistrons. The genetic limits of the deletions we studied are diagrammed in Fig. 2, which also provides a qualitative summary of our observations. This summary and genetic map (Fig. 2) are presented now to facilitate description of the results.

Our initial studies implicated the domain defined by deletion *326* in *regA*-related control of rIIB protein synthesis. No other tested segment (as defined by deletions) within either rII cistron contains a site that, when missing, causes constitutive rII expression in a phage T4 *regA*$^+$ infection. Examples of the types of results that focused our attention to the rII segment deleted in *326* are shown in Fig. 3. Deletion *326* removes the normal initiator AUG for rIIB protein synthesis and allows initiation, with slightly reduced efficiency, at an internal and out-of-phase AUG codon (ref. 7; Fig. 2). The double mutant 326+FC6 initiates rIIB protein synthesis at this internal initiator, but the reading frame is corrected by the *FC6* deletion (7). The double mutant gives synthesis, at lowered levels, of an abbreviated but biologically functional rIIB peptide (7). Synthesis of the truncated protein generated by 326+FC6 is insensitive to inhibition by the *regA*$^+$ allele (Fig. 3). This effect is not a consequence of deleting information with *FC6* because the double mutant FC6+FC31 shows normal responses of rIIB expression in the *regA*$^+$ and *regA*$^-$ backgrounds (Fig. 3). Furthermore, as will be shown below, the rIIB mutant 326+a_2+FC9 responds like 326+FC6.

Fine-Structure Mapping of the Site for *regA*-Mediated Control on rIIB-mRNA. In view of the results with the double mutant 326+FC6, we focused our analyses on the region removed by *326*. We measured rIIB expression in phage T4 in-

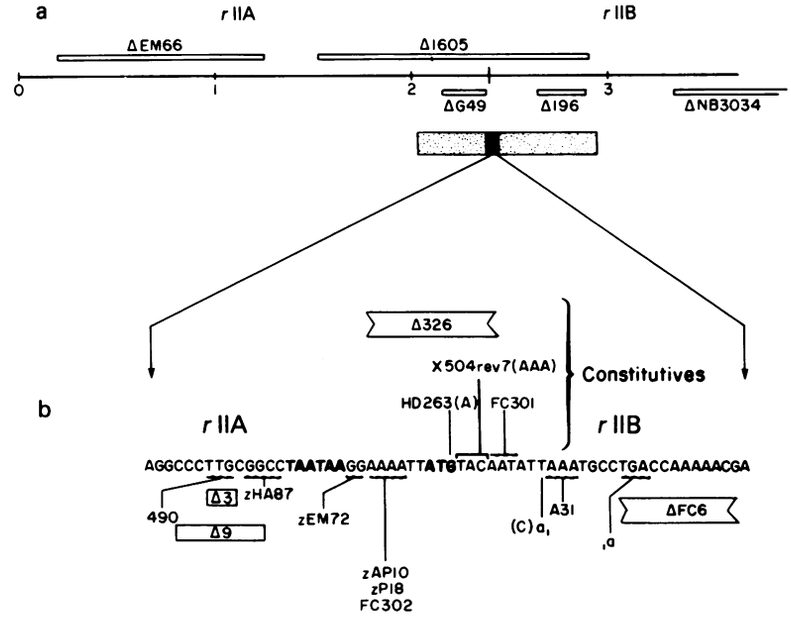

FIG. 2. Mapping the regulatory site for *regA* protein-mediated repression on *r*IIB mRNA. (*a*) Gross mapping. The five deletions are either precisely known in the *r*II sequence [G49 and 196 (31)] or have been placed approximately (31). Deletions 196 and EM66 remove 3n base pairs of DNA, and thus allow straightforward quantitation of gene products large enough to be visualized in the autoradiograms of $NaDodSO_4$/acrylamide gels. Deletion 1605 yields only *r*IIA expression; the *r*IIA fragment made can be seen on gels. None of the five deletions was defective in regA-mediated repression. The region stippled under the map of *a* has been sequenced (31). The digits 0–3 shown with vertical lines on the genetic map indicate the nucleotides, in kilobase pairs, of the *r*II cistrons (31). The unnumbered vertical line shows the boundary between *r*IIA and *r*IIB. All the *r*IIB translational defectives (7) and regA-insensitive mutations are found within a small domain at the *r*IIA/*r*IIB interface. (*b*) Fine mapping. The domain shown as a solid area in the sequenced region *a* of *r*II is expanded to highlight 60 base pairs, including the *r*II intercistronic domain. Within this region are shown three *r*IIA mutations (*490*, *Δ3*, and *Δ9*) and five *r*IIB mutations (*zHA87*, *zEM72*, *zAP10*, *zP18*, and *FC302*) that lie 5′ to the *r*IIB translational initiation codon. All but two of these mutations (490 and Δ3) reduce *r*IIB translation (7). All eight respond to the *regA* gene function (Table 1; ref. 7); overproduction of the *r*IIB protein at late times in a phage T4 *regA*$^+$ infection is not observed. Starting at the 12th nucleotide encoding *r*IIB are shown four mutations that are 3′ to the initiation codon (a_1, *A31*, $_1a$, and *FC6*) and that do not alter translation or regA regulation of the *r*IIB mRNA (Table 1; refs. 7 and 31). Between these two clusters (shown below the sequence) lie four mutations that are constitutive for *r*IIB translation. The mutant carrying deletion *326* initiates *r*IIB translation at the AUG just 3′ to *A31* (7). The mutation *HD263* has been described (32, 44); of the four constitutive mutations, only *HD263* greatly reduces translational initiation. The superposition of all the sites shown onto the wild-type sequence has been described (7, 31). Mutations *FC301*, *FC302*, *zP18*, and *A31* are most likely insertions or deletions of one or a few base pairs; *zHA87*, *zEM72*, *zAP10*, a_1, and $_1a$ are almost certainly transitions (7, 45). We have used wavy lines to indicate the approximate positions of mutations that are not precisely known. The sequences of mutations *HD263*, a_1, and *X504rev7* are shown in parentheses.

fections that utilized a number of small additions, deletions, and substitutions that mapped close to and under *326*. The precise locations of the mutations used have been determined (7). The rates of *r*IIB expression that were obtained in the analyses and the deletions screened in the first analyses are compared in Table 1. Only lesions within a small (less than 15 nucleotides) segment caused an alteration in the response to the *regA*$^+$ function. The frameshift mutation *FC301*, which caused a strong constitutive response, defines one end (3′ to the initiation codon) of the segment. The other boundary of the domain is set, for now, by the guanosine of the initiation codon [which is an adenosine in HD263 (Fig. 2)]. The mutations that map within the Shine–Dalgarno sequence (4, 5, 7) did not affect the response to *regA*$^+$ protein (see table 3 of ref. 7). In summary, the genetic data define an *r*IIB target site for *regA* protein-mediated control as a 7- to 15-nucleotide sequence that overlaps the 3′ boundary of the ribosomal binding region on *r*IIB mRNA (7, 32).

DISCUSSION

In these experiments and in unpublished data, a number of mutations flanking or abolishing the *r*IIB initiation codon have been tested for *regA* protein-mediated regulation (Fig. 2). All mutations 5′ to the *r*IIB AUG respond normally to the *regA* function. Similarly, the mutations 3′ to *FC301* respond normally to control by the *regA* gene. There are 15 nucleotides between *zAP10* and a_1, including the *r*IIB AUG (Fig. 2). The deletion *326* removes about 11 of these nucleotides, permits translational initiation at a downstream AUG (7), and is unresponsive to regA regulation. Because both mutants 326+FC6 and 326+a_2+FC9 are regA insensitive (Table 1), the regA target must lie within the region deleted by *326* itself. Three separate mutations that lie under deletion *326* (*HD263*, *X504rev7*, and *FC301*) render *r*IIB expression unresponsive to regA regulation. For these mutants, *r*IIB initiation occurs at the same codon [or position, in HD263 (44, 46)] used for wild-type translational initiation. X504rev7 is a double mutant. It was selected as a revertant of an ochre mutation at the second codon that alters the isoelectric point of the *r*IIB protein. The wild-type codon UAC (Tyr) is changed to AAA (Lys) in X504rev7 (7).

Thus, of the 15 nucleotides that could comprise the regA target, regA-insensitive mutations have been observed altering at least three. We cannot be sure that both of the nucleotides changed by *X504rev7* affect the *regA* protein site. The mapping of *FC301* strongly suggests that it lies just 3′ to the second codon in *r*IIB mRNA (7, 45). Until *FC301* is sequenced, however, we cannot rigorously extend the regA target further 3′ than the second codon.

The *regA* protein contains about 80 amino acids (26). Other small proteins recognize specific nucleic acid sequences. For example, the amino terminal "head" pieces of the *lac* and λ phage repressor proteins (47–50) are but 59 and 92 amino acids, respectively; the phage λ *cro* protein contains 66 amino acids (51, 52). Each protein recognizes operator DNA specifically (47, 51, 52). Analogously, the *regA* protein could recognize a site of six nucleotides or so. A site of this size is compatible with the

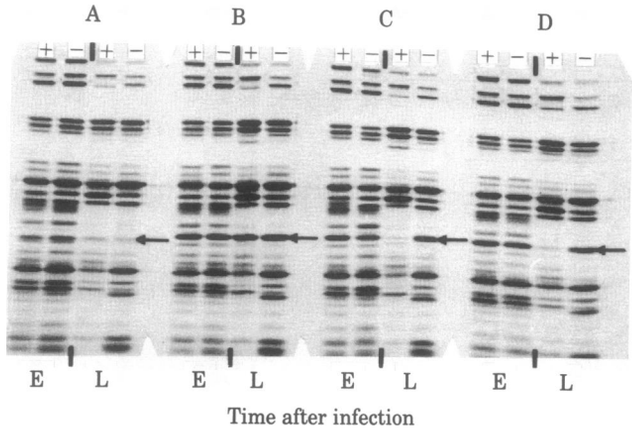

FIG. 3. Phage T4 *r*IIB mutations that render *r*IIB protein synthesis insensitive to control by the phage *regA* gene function. In addition to the *r*IIB lesions that were examined in the *regA*$^+$ and *regAR9* genetic backgrounds, the phage strains each carried nonsense lesions in genes *32*, *33*, *62*, *44*, and *55*. Gel electrophoretic assays were used to determine the amounts of T4-induced proteins synthesized at 5–10 min (E = early time period) and at 25–30 min (L = late time period) after infection. In the autoradiograph shown, the plus columns refer to infections with *regA*$^+$ phage and the minus columns to infections with *regA*$^-$ phage. The *r*II mutants used were: 326+FC6 (*A*); FC301+a_1+FC6 (*B*); FC6+FC31 (*C*); and *r*$^+$ (*D*). The horizontal arrows indicate the positions of the *r*IIB peptides synthesized in these infections.

genetics available at this moment. Recognition of a DNA operator sequence by a repressor protein involves scanning a known DNA structure in solution, whereas regA binding to specific mRNAs (if RNA is the direct target) may be complicated by the structural possibilities for RNA. Complicated models would allow a small protein to inspect a much larger domain of RNA than is proposed here. The interactions between the ribosomal proteins S4 and S8 with 16S RNA are metaphors for such complicated binding domains (53–56). However, given the genetics (Fig. 2) and the lack of biochemistry on the *regA* system, we propose for now that the regA target is small and contiguous with the initiation codon of the *r*IIB mRNA.

We imagine two mechanisms that might account for regA-mediated diminution of gene expression. On the one hand, the *regA* gene product could be a message-specific RNase. For many genes regulated by the *regA* protein, elimination of *regA* function is accompanied by both gene-specific protein overproduction *and* messenger stabilization (24, 27). One cannot know, however, if the instability directly results from regA action or from nucleolytic attack that follows simple translational repression. Trimble and Maley (26) showed that one regA-regulated gene, HMdCMP kinase (the phage T4 gene *1* product), yields equal amounts of mRNA *in vivo* whether the *regA* protein is present or not; their data suggest that, for at least one regA-sensitive mRNA, the means of repression is not mRNA destruction. If the *regA* protein regulates all target messengers by means of the same molecular mechanism, we would be forced to conclude that the *regA* protein is not an RNase. We note too that a model portraying regA as an RNase requires that regA-modulated mRNAs are not subject to degradation by the pathway(s) used to dispose of other T4 mRNAs because they are uniquely stable in a phage T4 *regA*$^-$ infection (24, 27).

A second potential means for regA-mediated regulation is simple translational repression. This model could have two forms, indicating different targets for the *regA* protein. In a model that reiterates transcriptional regulation by repressors, the *regA* protein would bind to the *r*IIB mRNA so as to prevent ribosomes from binding. This model would have a portion of the ribosome binding domain overlap the regA site. The data

Table 1. Effects of T4 *r*II mutations on *r*II protein synthesis in the *regA*$^+$ vs. the *regA*$^-$ genetic backgrounds

Mutant	Ratio *regA*$^+$/*regA*$^-$ *r*IIA protein	Ratio *regA*$^+$/*regA*$^-$ *r*IIB protein
ΔG49	0.17	0.16
ΔEM66	0.13	0.15
Δ1605	0.25	—
326+FC6	—	1.04
FC6+FC31*	—	0.13
326+a_2+FC9*	—	1.40
HD263	—	1.28
HD263rev10†	—	1.39
X504rev7	—	0.46
FC302+a_1+FC6‡	—	0.15
FC301+a_1+FC6	—	0.89
A31+$_1$a+FC0†	—	0.25
Δ196	—	0.25
ΔNB3034	—	0.11
r$^+$	0.15	0.17

E. coli B^E cultures that had been infected with the *r*II mutants were incubated with ^{14}C-labeled amino acids at 25–30 min after infection (a late time interval). Extracts of the labeled cultures were subsequently prepared and analyzed by gel electrophoretic assays. Densities of the relevant *r*II bands were then determined by scanning autoradiograms with a microdensitometer. In each case, the "*regA*$^+$/*regA*$^-$ *r*II protein synthesis ratio" refers to the ratio of densities of a particular *r*II protein band caused by the phage T4 *regA*$^+$ infection relative to the *regA*$^-$ infection.

* *FC31*, a_2, *FC9*, and *FCO* are mutations to the right of the expanded scale in Fig. 2. These mutations correct the translational reading frame (in the multiple mutants being examined) and eliminate an out-of-frame chain terminator (a_2). These sites have been located earlier (29, 31).

† This *r*IIB mutant is a pseudowild revertant of HD263 that has been described (44).

‡ This strain, by a mechanism that is not yet clear, gives *r*IIB translation through the a_1 codon in the plus frame (7, 51).

are consistent with this idea. A related model would have the ribosome itself capable of recognizing mRNA sequences 3′ to the initiation codon (1). Were this the case, message-specific repression could be accomplished if the *regA* protein were to bind directly to the ribosome so as to occlude ribosomal sites used for initiation of the subset of mRNAs regulated by regA. In this model regA would not possess RNA binding capacities on its own.

We cannot rigorously choose a model based on the available data [although the HMdCMP kinase data of Trimble and Maley (27) lead us away from RNA processing models]. If the *regA* protein were an RNase, the target sequence would have to be long to prevent unwanted destruction of all phage T4 mRNAs. An RNase site at any location on a mRNA can lead to functional inactivation, whereas simple repression models most probably utilize target sequences that closely flank the initiation codon. We are most intrigued by the repression model. Cardillo *et al.* (26), after reviewing the *regA* literature, suggested that simple repression by regA was most likely and also were unable to choose between a mRNA target and a ribosome target.

If regA is indeed a RNA-binding regulatory protein, it would be unique among such proteins. Previously studied proteins that bind to specific RNAs [such as EF-T_U, which must recognize tRNAs (57)] or to specific locations on some RNAs [such as RNase III binding to a variety of preferred cleavage sites (8, 9)] almost certainly recognize isomorphic *structural* elements shared by several target substrates. Even RNA binding proteins that have a single target site may scan for structural features: the R17/MS2 coat protein recognizes a local RNA hairpin that

includes the replicase initiation codon (9–12), *B. subtilis* RNase M5 recognizes a base-paired region at its processing target site on 5S rRNA (58), and the phage T4 gene 32 protein may scan for the absence of any RNA structure whatsoever (17, 18, 59). It is unlikely that the *regA* protein directly scans for local secondary structures on the *r*IIB mRNA: the target nucleotides cannot participate in a stable hairpin (G. D. Stormo, personal communication), and mutations [such as Δ9 (7)] that remove the only stable structure (60) near the *r*IIB initiation codon have no effect on *regA* function (7). Therefore, if the *regA* protein recognizes RNA directly, then its several target RNA molecules must be isomorphic in sequence, rather than in other structural features. Note that the regA site need not include the initiation codon itself; rather the model requires only that the protein binding site be overlapping with the ribosome binding site. In present formulations of ribosome binding sites (1), the regA site could be placed from 20 residues 5′ to 12 residues 3′ to the initiation codon itself. Gene *45* is among the genes affected by regA (Fig. 1). The ribosome binding domain of the gene *45* mRNA (E. Spicer and W. Konigsberg, personal communication) contains the sequence A-U-Ū-A-C-A-A-U-G-A-A-A [the initiation codon is underlined; the overscored sequence is found in the region we have identified as the regA site in *r*IIB mRNA (Fig. 2)].

We thank the National Institutes of Health and the American Cancer Society for support.

1. Gold, L., Pribnow, D., Schneider, T., Shinedling, S., Singer, B. S. & Stormo, G. (1981) *Annu. Rev. Microbiol.*, in press.
2. Ray, P. N. & Pearson, M. L. (1974) *J. Mol. Biol.* **85**, 163–175.
3. Ray, P. N. & Pearson, M. L. (1975) *Nature (London)* **253**, 647–650.
4. Steitz, J. A. (1980) in *Ribosomes: Structure, Function and Genetics*, eds. Chambliss, G., Craven, G. R., Davies, J., Davis, K., Kahan, L. & Nomura, M. (Univ. Park Press, Baltimore, MD), pp. 479–495.
5. Shine, J. & Dalgarno, L. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 1342–1346.
6. Dunn, J. J., Buzash-Pollert, E. & Studier, F. W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 2741–2745.
7. Singer, B. S., Gold, L., Shinedling, S., Colkitt, M., Hunter, L., Pribnow, D. & Nelson, M. A. (1981) *J. Mol. Biol.*, in press.
8. Dunn, J. J. & Studier, F. W. (1975) *J. Mol Biol.* **99**, 487–499.
9. Fallon, A. M., Jinks, C. S., Strycharz, G. D. & Nomura, M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 3411–3415.
10. Bernardi, A. & Spahr, P. F. (1972) *Proc. Natl. Acad. Sci. USA* **69**, 3033–3037.
11. Borisova, G. P., Volkova, T. M., Berzin, V., Rosenthal, G. & Gren, E. J. (1979) *Nucleic Acids Res.* **6**, 1761–1774.
12. Gralla, J., Steitz, J. A. & Crothers, D. M. (1974) *Nature (London)* **248**, 204–208.
13. Steitz, J. A. (1974) *Nature (London)* **248**, 223–225.
14. Kolakofsky, D. & Weissmann, C. (1971) *Biochim. Biophys. Acta* **246**, 596–599.
15. Krisch, H. M., Bolle, A. & Epstein, R. H. (1974) *J. Mol. Biol.* **88**, 89–104.
16. Gold, L., O'Farrell, P. Z. & Russel, M. (1976) *J. Biol. Chem.* **251**, 7251–7262.
17. Russel, M., Gold, L., Morrissett, H. & O'Farrell, P. Z. (1976) *J. Biol. Chem.* **251**, 7263–7270.
18. Lemaire, G., Gold, L. & Yarus, M. (1978) *J. Mol. Biol.* **126**, 73–90.
19. Yates, J. L., Arfsten, A. E. & Nomura, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 1837–1841.
20. Dean, D. & Nomura, M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3590–3594.
21. Brot, N., Caldwell, P. & Weissbach, H. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 2592–2595.
22. Dennis, P. P. & Fiil, N. P. (1979) *J. Biol. Chem.* **254**, 7540–7547.
23. Karam, J. D. & Bowles, M. G. (1974) *J. Virol.* **13**, 428–438.
24. Karam, J., McCulley, C. & Leach, M. (1977) *Virology* **76**, 685–700.
25. Wiberg, J. S., Mendelsohn, S., Warner, V., Hercules, K., Aldrich, C. & Munro, J. L. (1973) *J. Virol.* **12**, 775–792.
26. Cardillo, T. S., Landry, E. F. & Wiberg, (1979) *J. Virol.* **32**, 905–916.
27. Trimble, R. B. & Maley, F. (1976) *J. Virol.* **17**, 538–549.
28. Sauerbier, W. & Hercules, K. (1973) *J. Virol.* **12**, 538–547.
29. Barnett, L., Brenner, S., Crick, F. H. C., Shulman, R. G. & Watts-Tobin, R. J. (1967) *Philos. Trans. R. Soc. London Ser. B* **252**, 487–560.
30. Benzer, S. (1961) *Proc. Natl. Acad. Sci. USA* **47**, 403–415.
31. Pribnow, D., Sigurdson, D. C., Gold, L., Singer, B. S., Napoli, C., Brosius, J., Dull, T. J. & Noller, H. F. (1981) *J. Mol. Biol.*, in press.
32. Belin, D., Hedgpeth, J., Selzer, G. B. & Epstein, R. H. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 700–704.
33. Epstein, R. H., Bolle, A., Steinberg, C. M., Kellenberger, E., Boy de la Tour, E., Chevalley, R., Edgar, R. S., Susman, M., Denhardt, G. H. & Lielausis, A. (1963) *Cold Spring Harbor Symp. Quant. Biol.* **28**, 375–394.
34. Rabussay, D. & Geiduschek, E. P. (1977) in *Comprehensive Virology*, eds. Fraenkel-Conrat, H. & Wagner, R. R., (Plenum, New York), Vol. 8, pp. 1–196.
35. Bolle, A., Epstein, R. H., Salser, W. & Geiduschek, E. P. (1968) *J. Mol. Biol.* **31**, 339–362.
36. Wiberg, J. S., Dirksen, M., Epstein, R. H., Luria, S. E. & Buchanan, J. M. (1962) *Proc. Natl. Acad. Sci. USA* **48**, 293–302.
37. O'Farrell, P. Z. & Gold, L. M. (1973) *J. Biol. Chem.* **248**, 5502–5511.
38. Higgins, N. P., Geballe, A. P., Snopek, T. J., Sugino, A. & Cozzarelli, N. R. (1977) *Nucleic Acids Res.* **9**, 3175–3186.
39. Wood, W. B. & Revel, H. R. (1976) *Bacteriol. Rev.* **40**, 847–868.
40. Morris, C. F., Hama-Inaba, H., Mace, D., Sinha, N. K. & Alberts, B. (1979) *J. Biol. Chem.* **254**, 6787–6796.
41. Wood, W. B., Conley, N. P., Lyle, H. L. & Dickson, R. C. (1978) *J. Biol. Chem.* **253**, 2437–2445.
42. Snopek, T. J., Wood, W. B., Conley, M. P., Chen, P. & Cozzarelli, N. R. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 3355–3359.
43. David, M., Vekstein, R. & Kaufmann, G. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 5430–5434.
44. Singer, B. S. & Gold, L. (1976) *J. Mol. Biol.* **103**, 627–646.
45. Katz, E. R. & Brenner, S. (1975) *Mol. Gen. Genet.* **143**, 101–104.
46. Belin, D. (1979) *Mol. Gen. Genet.* **171**, 35–42.
47. Miller, J. H. (1979) *J. Mol. Biol.* **131**, 249–258.
48. Ogata, R. & Gilbert, W. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 5851–5854.
49. Pabo, C. O., Saver, R. T., Sturtevant, J. M. & Ptashne, M. (1979) *Proc. Natl. Acad. Sci. USA* **76**, 1608–1612.
50. Saver, R. T., Pabo, C. O., Meyer, B. J., Ptashne, M. & Backman, K. C. (1979) *Nature (London)* **279**, 396–400.
51. Johnson, A., Meyer, B. J. & Ptashne, M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 1783–1787.
52. Ptashne, M., Jeffrey, A., Johnson, A. D., Maurer, R., Meyer, B. J., Pabo, C. O., Roberts, T. M. & Saver, R. T. (1980) *Cell* **19**, 1–11.
53. Noller, H. F. (1980) in *Ribosomes: Structure, Function and Genetics*, eds. Chambliss, G., Craven, G. R., Davies, J., Davis, K., Kahan, L. & Nomura, M. (Univ. Park Press, Baltimore, MD), pp. 3–22.
54. Cole, M. D., Beer, M., Koller, T., Strycharz, W. A. & Nomura, M. (1978) *Proc. Natl. Acad. Sci. USA* **75**, 270–274.
55. Muller, R., Garrett, R. A. & Noller, H. F. (1979) *J. Biol. Chem.* **254**, 3873–3878.
56. Mackie, G. A. & Zimmerman, R. A. (1975) *J. Biol. Chem.* **250**, 4100–4112.
57. Altman, S., ed. (1978) *Transfer RNA* (MIT Press, Cambridge, MA).
58. Stahl, D. A., Mayhack, B. & Pace, N. R. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5644–5648.
59. Krisch, H. M., Duvoisin, R. M., Allet, B. & Epstein, R. H. (1980) in *Mechanistic Studies of DNA Replication and Genetic Recombination*, eds. Alberts, B. & Fox, C. F. (Academic, New York), pp. 517–526.
60. Tinoco, jun. I., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M. & Gralla, J. (1973) *Nature (London) New Biol.* **246**, 40–41.