**Evolution of RNA-binding Specificity in T4 DNA Polymerase**

(Received for publication, January 10, 1997, and in revised form, May 12, 1997)

Chien-Chia Wang‡, Andrey Pavlov§, and Jim D. Karam¶

From the Department of Biochemistry, SL43, Tulane University School of Medicine, New Orleans, Louisiana 70112

---

DNA polymerase of phage T4 (T4 gp43), an essential component of the T4 DNA replicase, is a multifunctional single-chained (898-amino acid) protein that catalyzes the highly accurate synthesis of DNA in phage replication. The enzyme functions both as a DNA-binding replication protein and as a sequence-specific RNA-binding autogenous translational repressor. We have utilized a phylogenetic approach to study the relationships between the two nucleic acid-binding functions of the protein. We found that autogenous translational control of gp43 biosynthesis has been conserved in phage RB69, a distant relative of T4, although we also found that the RB69 system differs from its T4 counterpart in two regards: (a) nucleotide sequence and predicted secondary structure of the RNA target (translational operator), and (b) RNA specificity of the protein. T4 gp43 is specific to the RNA operator sequence of the T4 genome whereas RB69 gp43 can bind and repress operator RNA from both phages equally well. In studies with T4-RB69 gp43 chimeras, we mapped T4 gp43 RNA-binding specificity to a protein segment that also harbors important determinants for DNA binding and the polymerase catalytic function. Our results suggest that RNA functions as a regulator of both the dosage and activity of this DNA replication enzyme.

DNA polymerase of bacteriophage T4, product of phage gene 43 (T4 gp43), is a 103-kDa single-chained enzyme that catalyzes two types of reactions (1): DNA template-directed nucleotide additions to a primer (i.e. the polymerase or POL function) and 3′ exonucleolytic hydrolysis of unpaired DNA ends (i.e. the EXO function). The enzyme is also a sequence-specific RNA-binding autogenous translational repressor (2). In T4 DNA replication, initiation specificity (replication origin recognition) is attributed to proteins other than gp43, but the polymerase is essential for construction of a T4-specific multienzyme replicase and allows the replicase to interact with and copy any segment of the DNA genome regardless of the nucleotide sequence (3). In contrast, RNA recognition by the protein is highly specific to the polyribonucleotide chemistry and secondary structure of only one known phage encoded target, a translational operator that overlaps the ribosome-binding site of gene 43-specific mRNA (4–6). The affinity of T4 gp43 to operator RNA is 50–100-fold higher than to a DNA facsimile of the operator, or to RNA and DNA of generic nucleotide sequence (5). Clearly, DNA and operator RNA represent two distinct types of ligand for this enzyme.

The POL function of T4 DNA polymerase is indispensable for template-dependent DNA synthesis and, in addition, determines the fidelity of nucleotide precursor selection during replication (7, 8). The EXO function, on the other hand, can be eliminated by mutation without loss of replication and viability, although the proofreading (editing) activity of the 3′-exonuclease is also necessary for maintaining high fidelity in T4 DNA replication (9). Presumably, long-term stability of the phage genome depends on both the POL and EXO functions (44). Since interaction of T4 gp43 with DNA underlies the enzyme’s ability to support replication and express its fidelity functions, the determinants for DNA binding must have resisted change during evolution of this enzyme. In fact, many DNA polymerases share similar clusters of amino acid sequence (conserved motifs) at sites that have been shown, at least in some of these enzymes (including T4 gp43), to be required for the polymerization and exonucleolytic reactions (8, 10). In contrast, the need for conserving or evolving a sequence-specific RNA-binding function in such enzymes is not obvious. T4 gp43 is the only replication DNA polymerase known to possess RNA-binding specificity, but even with this genetically well characterized enzyme, it is not known if the capacity to bind specific RNA is crucial for long-term survival of the phage. Autogenous control of gp43 biosynthesis can be eliminated by mutations in the RNA operator without measurable consequences to phage replication and its fidelity under standard laboratory growth conditions (4, 6); however, all known amino acid substitutions in the protein that affect autogenous control also affect replication under some physiological condition, e.g. many exhibit temperature-sensitive phenotypes (8). We report here the results of a phylogenetic analysis which shows that RNA recognition is a conserved property of gp43 and suggests that the RNA and DNA-binding functions of gp43 are structurally linked by coevolution. Some of the RNA binding determinants of gp43 map in the POL domain of this protein. We propose that RNA is important for control of both the dosage (translational control) and enzymatic activities of this replication DNA polymerase.

**EXPERIMENTAL PROCEDURES**

**Phage and Bacterial Strains**—The RB phages used, including RB6, RB18, RB19, RB51, RB69, and RB70 (11), were obtained from W. B. Wood (University of Colorado, Boulder, CO). The T4 gene 43 double amber mutant 43amE4322-B22 has been described (12); it carries UAG codons for positions 386 and 731 of the gene product (8). It was grown on the amber-suppressing host Escherichia coli CR63 (sup D, serine insertion). The RB69 gene 43 mutant 43sacd has also been described (13); it carries a small frameshifting internal deletion which leads to the production of a truncated gene product that cannot support phage replication. RB69 43sacd was grown on bacterial hosts expressing a cloned wild-type RB69 gene 43.

**E. coli K802** (supE: glutamine insertion at UAG codons; Ref. 14) is highly competent for DNA transformations and was used as host for the
initial isolation of recombinant plasmids. *E. coli* BL21(DE3), which harbors a chromosomally integrated T7 RNA polymerase gene that can be expressed under *lac UV5* promoter control (15), was used as host for recombinant plasmids expressing gene 43 sequences under control of the T7 φ10 promoter of cloning vector pSP72 (Promega). BL21(DE3) was also used as host for qualitative spot tests and quantitative burst size measurements in plasmid-phage complementation assays (13). *E. coli* CAJ70 (supU UGA suppressor, tryptophan-insertion; Ref. 16) was used as host for phage infections when RB69-induced proteins were to be visualized by SDS-PAGE1 assays. Bacterial and phage growth conditions and complementation assays were as described previously (15, 17).

**Purification of Gp43—** Preparations of T4 DNA polymerase were the same as those used previously (5). The RB69 gp43 used for *in vitro* assays was purified from 2 liters of *E. coli* BL21(DE3) cells (at 5 × 10^10^ cells/ml) harboring recombinant plasmid pCW19, which expresses a translational operator constitutive (op*) RB69 gene 43 under T7 promoter control. Cells were induced with 10 mM isopropyl-thio-β-D-galactoside at 30°C for 2 h, harvested by centrifugation (total cell weight ~10 g), and resuspended in 25 ml of lysing buffer (Buffer A) containing 20 mM Tris, pH 7.5, 1 mM EDTA, 10 mM MgCl₂, 10 mM β-mercaptoethanol, 100 mM NaCl, and 30 mg/ml α₂-macroglobulin. The cell suspension was passed three times through a French Pressure Cell at 15,000 p.s.i. and the resulting lysate was cleared by two rounds of centrifugation at 30,000 × *g* for 30 min followed by one centrifugation at 100,000 × *g* for 30 min. The resulting supernatant (~50 ml) was dialyzed overnight against Buffer B (20 mM Tris, pH 7.5, 1 mM EDTA, 10 mM MgCl₂, 10 mM β-mercaptoethanol, and 10% glycerol) and then loaded onto a 40 × 1.6-cm column bed of Q Sepharose anion-exchanger (Pharmacia Biotech Inc.) equilibrated in the same buffer. The column was washed with 300 ml of Buffer B and then eluted with 400 ml of a 0–0.5 M KCl gradient made in buffer B. Fractions (4 ml each) containing gp43 were pooled and dialyzed against Buffer C (20 mM KH₂PO₄, pH 6.1 mM EDTA, 10 mM β-mercaptoethanol, 10% glycerol) before loading onto a 5 × 1-cm column bed of Phosphohtrolgel (IBF Biotechnics). The column was washed with 50 ml of Buffer C and then eluted with 50 ml of a 0–0.5 M KCl gradient made in Buffer C. Gp43-containing fractions were pooled, dialyzed, and then loaded onto a 2-ml column of single-stranded DNA-agarose gel (Life Technologies, Inc.). The column was washed with 20 ml of Buffer C and then eluted with a 0–0.5 M KCl gradient in Buffer C. The fractions containing gp43 were pooled, placed in dialysis tubing, and concentrated 2-fold by dialysis against solid PEG 8000. After retightening the tubing around its contents, additional dialysis was carried out against 1000 ml of Storage Buffer (200 mM K₂HPO₄, pH 6.5, 10 mM β-mercaptoethanol, 50% glycerol, 10% bovine serum albumin). The reaction mixture was incubated at 30°C for 15 min before being quenched by trichloroacetic acid precipitation. Quenched samples were washed on glass fiber filters (Gelman), first with 5% cold trichloroacetic acid and then with 95% ethanol, dried, and counted in a scintillation spectrometer. One polymerase unit is defined as the incorporation of 1 nmol of dAMP under the conditions described above. The values determined for various concentrations of gp43 were then averaged and entered as the “POL activity” of the gp43 preparation under study.

The 3'-exonuclease activity of gp43 preparations was determined by measuring nucleotide hydrolysis of a 5'-32P-end-labeled (dT)₆ oligonucleotide substrate (19) in 30-μl assay mixtures. Gp43 at 0.1 mM and (dT)₆ concentrations ranging from 0.5 to 64 μM were used. Assays were carried out under the same conditions as those used for POL activity measurements, except that dNTPs were omitted and the substrate was (dT)₆ rather than salmon sperm DNA. The amount of 32P-labeled (dT)₁₆ added was the same for all incubations (final concentration 0.005 μM, ~10⁶ cpm). Samples (5 μl) were withdrawn at different times of incubation and added to stop solution (also 5 μl) consisting of 5% formamide, 20 mM EDTA, 0.05% bromphenol blue, and 0.05% xylene cyanol. The quenched samples were then subjected to electrophoresis on a 1.5% Long Ranger (J. T. Baker) sequencing gel containing 7 M urea. The intensities of radiolabeled DNA bands in the gel were scanned and quantitated on a Fuji Bio-Imaging Analyzer (Fuji Photo Film Co., LTD.).

**RESULTS**

The Genomes of T4 and RB69 Are Related, but Highly Diverged from Each Other—Previously, we pointed out that the structural genes for DNA polymerases of T4 and RB69 differ at nearly 35% of their nucleotide positions (13). Comparisons between the electrophoretic patterns of T4 and RB69 induced proteins suggest that the entire genomes of the two phages had diverged significantly during their evolution. This is illustrated in Fig. 1, which compares SDS-PAGE profiles of prereplicative proteins for phages RB69, T2, T4, T6, and a number of other wild-type phage strains belonging to the phage phylogenetic group (11). Although electrophoretic mobility is not a direct or reliable measure of genetic similarity between proteins, the overall profile for RB69-induced proteins is clearly distinguishable from that of the other phages in the T4 group. Consistent with these results are previous observations that, except for RB69, all the phage strains used for Fig. 1 are at least partially compatible (and can recombine) with T4 in mixed infections (11, 20). Nevertheless, despite its many differences from this phage group, RB69 clearly resembles T4, T2, and T6 in general organization of the genetic map (21), resistance to restriction enzymes that also fail to cleave T4 genomic DNA (13), and primary structures and chromosomal order of several replication genes that have been well characterized in T4 (22). We focus here on the implications of divergence versus conservation of structure, function, and regulation of gp43 from T4 and RB69.

**T4 gp43 and RB69 gp43 Have Similar in Vitro Activities—** Identity of the RB69 gp43 band marked in Fig. 1 (which migrates more slowly than the T4 gp43 band) has been established by cloning wild-type and mutant alleles of the structural gene in expression vectors and demonstrating that the wild-type allele can support replication of T4 gene 43-defective phage mutants in phage-plasmid complementation assays (13). In addition, we have isolated and purified this RB69 encoded

---

1 The abbreviation used is: PAGE, polyacrylamide gel electrophoresis.

2 L. S. Yeh and J. D. Karam, manuscript in preparation.
Recently, the T4 and RB69 Gene 43 Translational Initiation Regions—The studies summarized above led us to consider that RB69 may provide generalized insights not only into the principles that govern evolutionary divergence of protein structure in relation to conservation of function, but also into the linkage between evolution of trans-acting regulatory proteins, such as gp43, and their cis-acting nucleic acid target sites, such as the translational operator for gene 43. The translational initiation region for gp43 translation resides in an untranslated leader segment of the mRNA and may be subject to different constraints with respect to evolutionary divergence than is the structural gene, which specifies an essential enzyme. We examined the nature of nucleotide sequence divergence among the T4-related phages listed in Fig. 1, focusing on the entire untranslated intercistronic region between genes regA and 43. On the basis of studies with T4, this region harbors the translational operator for gene 43 (2) as well as other control elements for transcription initiation and mRNA processing (22, 23).

Fig. 3 compares the nucleotide sequence of the T4 regA-43 intercistronic region (2) to that determined here for RB69. To obtain the RB69 sequence, we used RB69 gene 43-specific sequencing primers and several plasmid clones of RB69 genomic DNA segments encompassing portions of both the regA gene and gene 43. We were able to confirm the RB69 regA gene sequence that was reported by Miller and Jozwik (24) and most of the information relating to the putative RB69 operator sequence that was reported by Tuerk et al. (6), who used T4 gene 43-specific oligonucleotide primers and avian myeloblastosis virus reverse transcriptase to sequence RNA isolated from RB69-infected cells. We have corrected the few ambiguities in nucleotide assignments from the Tuerk et al. (6) study and been able to obtain the complete primary structure of the region from the RB69 regA termination codon (UAA) to the gene 43 initiator AUG of the RB69 genome. In Fig. 3A, we show a nucleotide sequence alignment that attempts to depict a maximal amount of similarity between the T4 and RB69 regA-43 intercistronic regions. At most, the two regions are 50% identical (47% for the operator segment) as compared with ~65% nucleotide identity in the protein encoding segments of the two polymerase genes (13). It appears that single and multiple base duplications and deletions (frameshift type mutations) have contributed significantly to the divergence of the two intercistronic regions over the course of evolution. Such mutations are much less likely to become established in open reading frames of essential genes like gene 43, although we note that the T4 and RB69 genes do differ in number of codons (5 more in RB69; Ref. 13) and that such differences could have originated from multiple (but in-frame) base additions or deletions.

The higher divergence in the untranslated upstream region (particularly the operator segment) as compared with the structural gene is consistent with the observation that T4 gene 43 operator mutations, many of which overproduce wild-type gp43, are not detrimental to phage growth within the confined region upstream of the gene 43 translation termination codon (UAA) to the gene 43 initiator AUG of the RB69 genome. In Fig. 3A, we show a nucleotide sequence alignment that attempts to depict a maximal amount of similarity between the T4 and RB69 regA-43 intercistronic regions. At most, the two regions are 50% identical (47% for the operator segment) as compared with ~65% nucleotide identity in the protein encoding segments of the two polymerase genes (13). It appears that single and multiple base duplications and deletions (frameshift type mutations) have contributed significantly to the divergence of the two intercistronic regions over the course of evolution. Such mutations are much less likely to become established in open reading frames of essential genes like gene 43, although we note that the T4 and RB69 genes do differ in number of codons (5 more in RB69; Ref. 13) and that such differences could have originated from multiple (but in-frame) base additions or deletions.

Binding of the RB69 Translational Operator to Gp43 from RB69 and T4—As proposed by Tuerk et al. (6) and diagrammed in Fig. 3B, the putative operator from RB69 can be fit into an RNA hairpin structure that resembles the T4 operator hairpin (2, 26); however, the predicted stabilities of the two hairpins are quite different from each other. In fact, based on mutational analyses of the T4 operator (4),3 we predicted that the operator sequence from RB69 would not bind effectively to T4 polymerases from T4 and RB69. The vertical scale marks the sizes of oligonucleotide products from the hydrolyses. Densities of bands were determined on a PhosphorImager and used for calculating $K_m$ and $K_{cat}$ values shown in Panel C. Assay conditions are described under "Experimental Procedures.

**Fig. 2.** Assay for POL and EXO activities of purified gp43. Panels A and B compare (dT)$_{16}$ hydrolysis by the EXO activities of DNA polymerases from T4 and RB69. The vertical scale marks the sizes of oligonucleotide products from the hydrolyses. Densities of bands were determined on a PhosphorImager and used for calculating $K_m$ and $K_{cat}$ values shown in Panel C. Assay conditions are described under "Experimental Procedures."
gp43. This prediction was borne out in experiments that measured in vitro binding affinities of the two RNA operator sequences to purified gp43 from the two phylogenetic relatives. Results are summarized in Fig. 4. T4 gp43 bound RNA corresponding to operator from the T4 source about 40 times more tightly than it did RNA corresponding to the putative RB69 operator. In contrast, the purified RB69 enzyme bound both RNA substrates equally well and with an affinity similar to that observed for T4 gp43 binding with its own operator, i.e. $K_d = 1$–2 nm. These results predicted that RB69 gp43 is, like T4 gp43, an autogenous translational repressor and that its repressor activity would be effective against the T4 as well as RB69 operators. In contrast, T4 gp43 would not be expected to repress the RB69 operator. The biological experiments described below confirmed these predictions.

**RB69 Gp43 Is Autogenously Regulated**—Autogenous regulation of RB69 gp43 biosynthesis was confirmed in two types of assays, one involving direct measurements of plasmid-mediated RB69 gene 43 expression and the other involving measurements of the effect of plasmid encoded RB69 gp43 on gp43 synthesis from infecting phage. The better understood T4 gp43 system was used for controls and to test for differences in RNA specificity between the two phage enzymes. The experiments described in Fig. 5 demonstrate that overproduction of plasmid encoded T4 as well as the RB69 gene 43 products occurred when either the operators or structural genes of the two phage phylogenetic relatives were altered by mutation. Other T4 gene 43 operator and repressor mutants have also been studied and observed to exhibit derepressed gp43 biosynthesis (2, 4, 6). Fig. 6 demonstrates that plasmid-encoded wild-type gp43 from either source can repress (in trans) the synthesis of gene 43 products from infecting phage strains and that the T4 protein represses the T4 operator specifically whereas RB69 gp43 can repress the gene 43 operators of both phages. That is, as an RNA-binding protein, RB69 gp43 has a broader RNA-binding specificity than its T4 counterpart.

**The RNA-binding Domain of Gp43**—To have conserved autogenous translational control of gp43 biosynthesis while diverging in operator sequence and specificity of the gp43-RNA interaction, T4 and RB69 must have each accumulated changes in gp43 primary structure that compensated for losses or changes in operator contact. We examined segments of amino acid sequence divergence between the two DNA polymerases as likely locations for the distinguishing determinants of specificity to RNA. We constructed several chimeric proteins consisting of exchanged segments of divergence between T4 gp43 and RB69 gp43 and determined repressor specificities of the T4-RB69 gp43 chimeras by in vivo assay. Although our assays were limited to chimeric constructs in which RNA binding activity could be demonstrated, we were able to correlate differential RNA recognition by T4 gp43 with a segment of the protein that specifies determinants for the POL function. Results are summarized in Fig. 7. For purposes of data interpretations, we define T4 RNA specificity as the ability to repress the T4 but not the RB69 operator, and RB69 RNA specificity as the ability to repress the T4 and RB69 operators equally well. The six chimeric gp43 constructs compared in Fig. 7 represent 3 pairs of reciprocal RB69-T4 gp43 exchanges involving the middle 400-amino acid segments of the two proteins. Other constructs were also studied, but results were too ambiguous to interpret because of lack of RNA binding activity in at least one critical member of the pair being compared. The CW95-CW96 comparison (Fig. 7) is an example. CW95 gp43 behaves just like wild-type RB69 gp43 (represses both the T4 and RB69 operators), but the expected reciprocal T4-like specificity could not be demonstrated in CW96 because the protein made by this construct is inactive. Nevertheless, the CW95-CW96 comparison suggests that the gp43 segment spanning from residue 331(T4)/334(RB69) to residue 388(T4)/391(RB69) does not bear RNA specificity determinants. This, of course, does not rule out the possibility that the segment may harbor determinants for nonspecific, and stabilizing, RNA contacts in...
both T4 gp43 and RB69 gp43. In the comparison between the chimeras CW4 and CW5, it is clear that gp43 segment 498(T4)/553(RB69)-552(T4)/555(RB69) also does not harbor determinants for distinguishing between the T4 and RB69 RNA specificities (Fig. 7). Also, deletion of the segment exchanged in CW4 and CW5 does not eliminate either binding or specificity to RNA (results not shown).

The results with chimeras CW131 and CW132 are perhaps the most revealing, although interpretations are somewhat complicated. CW132 gp43 appears to be largely inactive against both the T4 and RB69 operators; however, we do detect weak activity from this protein against the T4 operator (CW132 panel, Fig. 7). That is, CW132 gp43 appears to be weakly active against both the T4 and RB69 operators; however, we do detect weak activity from this protein against the T4 operator (CW132 panel, Fig. 7). On the other hand, CW131 gp43 bears the narrower RNA specificity of T4 gp43.

DISCUSSION

Our results indicate that autogenous translational repression has been conserved in evolution as a regulatory mecha-
nism for biosynthesis of the replicative DNA polymerase of the T4 family of phages. Comparisons between the DNA polymerase genes of T4 and RB69 suggest that in each case, the translational operator co-evolved with its cognate repressor to an extent where RNA specificity diverged while autogenous control was maintained. On the basis of the differences in primary structure between RB69 gp43 and T4 gp43, we suspect that RB69 gp43 evolved a capacity to bind a less stable RNA structure than did T4 gp43. In this context, we regard the RNA specificity of T4 gp43 to be narrowly limited to RNA structures that are at least as stable as wild-type T4 operator. Indeed, base substitutions that destabilize the T4 RNA hairpin lead to loss of binding and repression by T4 gp43 whereas base substitutions that further stabilize the hairpin maintain high affinity and repression by the T4 protein (4, 26, 27). T4 gp43 and RB69 gp43 bind and repress the T4 operator equally well (Figs. 4 and 6); however, it remains to be determined if the two proteins make identical contacts with this operator and if some of these contacts are missing in the interaction between RB69 gp43 and its own operator. Possibly, RB69 gp43 evolved RNA contacts that are unique to its operator, in addition to bearing contacts that are unique to its operator, in addition to bearing

### Figure 7

**In vivo assay of RNA-binding specificities of T4-RB69 gp43 chimeras.** *E. coli* CAJ70 cultures (at 2 × 10^9 cells/ml) expressing plasmid-borne wild-type or chimeric gene 43 or no gene 43 were infected with RB69 deletion mutant 43sacd or T4 double-amber mutant 43E422-B22 (multiplicity of infection = 10) and then labeled with [35S]methionine (5 μCi/ml) for 10 min beginning at 10 min postinfection. Samples of infected cultures were then analyzed by SDS-PAGE on a 10% gel. The amino acid residues bracketing the segments exchanged between the T4 and RB69 gp43 molecules are shown in the “gene configuration” column. In the autoradiogram panels, positions of the truncated gp43 fragments (gp43f) expressed from the mutant phages are marked with a ○ for RB69 gp43f, and a ● for T4 gp43f. The direction of protein migration in the gels is presented in left-to-right orientation. Support of replication of infecting phage mutant by the CW4, CW95, CW132, T4 WT, and RB69 WT plasmid constructs results in synthesis of phage late proteins (additional bands, e.g. those indicated by > in the respective autoradiogram lanes).

### Table

| Construct name | Gene 43 configuration (●: RB69; ○: T4) | RB69 43sacd infections | T4 43sacd infections |
|---------------|--------------------------------------|------------------------|----------------------|
| CW5           | 498 355                               |                        |                      |
| CW4           | 501 555                               |                        |                      |
| CW96          | 331 388                               |                        |                      |
| CW95          | 334 391                               |                        |                      |
| CW132         | 555 726                               |                        |                      |
| CW131         | 556 729                               |                        |                      |
| T4 WT         | no gene 43                            |                        |                      |
| CONTROL       |                                      |                        |                      |

**Fig. 7.** In vivo assay of RNA-binding specificities of T4-RB69 gp43 chimeras. *E. coli* CAJ70 cultures (at 2 × 10^9 cells/ml) expressing plasmid-borne wild-type or chimeric gene 43 or no gene 43 were infected with RB69 deletion mutant 43sacd or T4 double-amber mutant 43E422-B22 (multiplicity of infection = 10) and then labeled with [35S]methionine (5 μCi/ml) for 10 min beginning at 10 min postinfection. Samples of infected cultures were then analyzed by SDS-PAGE on a 10% gel. The amino acid residues bracketing the segments exchanged between the T4 and RB69 gp43 molecules are shown in the “gene 43 configuration” column. In the autoradiogram panels, positions of the truncated gp43 fragments (gp43f) expressed from the mutant phages are marked with a ○ for RB69 gp43f, and a ● for T4 gp43f. The direction of protein migration in the gels is presented in left-to-right orientation. Support of replication of infecting phage mutant by the CW4, CW95, CW132, T4 WT, and RB69 WT plasmid constructs results in synthesis of phage late proteins (additional bands, e.g. those indicated by > in the respective autoradiogram lanes).
essential enzyme for heme biosynthesis. UV cross-linking studies have revealed a major protein-RNA contact at a region near the enzyme’s catalytic center (31). Four arginine residues at the substrate (citrate)-binding site have been shown by site-directed mutagenesis to be directly involved in RNA binding (32). Several well-characterized metabolic enzymes, including thymidylate synthase (33), dihydrofolate reductase (34), and glyceraldehyde-3-phosphate dehydrogenase (35), have also been shown to bind RNA and to harbor overlapping binding sites for the cognate RNA and substrate (or cofactor) of the enzyme-catalyzed reaction. In addition, thymidine synthase and glyceraldehyde-3-phosphate dehydrogenase have been reported as RNA-binding translational repressors (34, 36). So, the T4 gp43 protein may be an example of a common theme in evolution where a protein utilizes RNA as regulator to control its own biosynthesis as well as catalytic functions.

Recently, RB69 gp43 was crystallized and its structure solved (43). The three-dimensional placement of linear segments of the amino acid sequence in the structure and the known biochemical properties of this and other DNA polymerases suggest that this family of enzymes makes at least three types of contact with the DNA primer-template complex during replication: the double-stranded DNA portion where DNA synthesis has just been completed, the single-stranded DNA template to be copied, and the single-stranded DNA portion of the primer when undergoing proofreading. Operator RNA may compete for DNA binding determinants at any (or all) of these components of the gp43-primer-template interaction. We should also mention that the spectrum of T4 gene 43 missense mutations that abolish translational repressor activity do not implicate any particular “domain” of gp43 in RNA binding. For example, amino acid substitutions near the protein’s N terminus (namely E26K; Ref. 8) may abolish autogenous control (17), by affecting protein structure globally rather than by removing RNA contact points. A site-directed mutational search for RNA-binding determinants and specific amino acid exchanges between the two gp43 phylogenetic relatives should ultimately identify two types of protein-RNA contacts: (a) those that determine high affinity and (b) a subset of affinity determinants that distinguishes between RNA specificities of the RB69 and T4 proteins.

We are particularly intrigued by the ubiquity of RNA binding functions among enzymes that have no other obvious nucleic acid targeted physiological role. In addition to the examples cited above, the list includes glutamate dehydrogenase (37) and lactate dehydrogenase (38). It has been proposed that the RNA-binding site and nucleotide-binding domain may be evolutionarily related in at least some of these enzymes (39). Also, in vitro selection techniques have shown that high-affinity RNA ligands can be found for virtually any protein (40, 41). For autoregulatory proteins, most, if not all, of the regulatory elements of the mRNAs reside in the less constrained untranslated (usually 5′-) regions, and are therefore amenable to maturation for high affinity and specificity. The conservation of autogenous translational control in the gene 43 system of the T4 family of phages suggest that control of synthesis of the protein has been an important contributing factor in the natural selection of this group of viruses. Regulation of enzyme activity may provide additional advantages. For example, a mutation in the protein that leads to a sublethal decrease in DNA polymerase activity may ultimately become established in the primary structure of the enzyme if it resulted in overproduction of the partially defective protein (due to translational derepression) and consequent compensation for the decreased activity, until compensatory mutations arise and reestablish enzyme activity and autogenous control to their normal levels. Operator mutations resulting in overproduction of wild-type enzyme, although seemingly harmless under laboratory growth conditions (4, 6), may have a selective disadvantage in the long term because of cumulative small side effects, e.g. interference from unused excess protein with assembly and functioning of processive replicase complexes and competition between the hypertranslated mRNA and other phage mRNAs for the limited number of ribosomes in a phage-infected cell. It is also possible that preferential binding of gp43 to a specific RNA ligand during the prereplicative phase serves as a means to immobilize the enzyme and facilitate assembly of the multiprotein complex that ultimately translates the polymerase to the sites of DNA replication initiation. Such roles for the gp43-operator interaction may not appear to be critical when examined under highly favorable growth conditions, but may have governed the natural selection of T4 and its related phages.

Acknowledgments—We thank Bill Konigsberg, Jan Drake, and Bob Weisberg for comments on the manuscript.

REFERENCES

1. Goulian, M., Lucas, Z. J., and Kornberg, A. (1968) J. Biol. Chem. 243, 627–638
2. Andrake, M., Guild, N., Hsu, T., Gold, L., Tuerk, C., and Karam, J. (1988) Proc. Natl. Acad. Sci. U. S. A. 85, 1974–1978
3. Kreuzer, K. N., and Morriss, S. W. (1994) in Molecular Biology of Bacteriophage T4 (Karam, J. D., Drake, J. W., Kreuzer, K. N., Mosig, G., Hall, D. W., Eislering, F. A., Black, L. W., Eddy, S., Parma, D., and Gold, L. (1980) J. Mol. Biol. 213, 749–761
4. Nossal, N. G. (1994) in Molecular Biology of Bacteriophage T4 (Karam, J. D., Drake, J. W., Kreuzer, K. N., Mosig, G., Hall, D. W., Eislering, F. A., Black, L. W., Eddy, S., Parma, D., and Gold, L. (1980) J. Mol. Biol. 213, 749–761
5. Reha-Krantz, L. (1994) in Molecular Biology of Bacteriophage T4 (Karam, J. D., Drake, J. W., Kreuzer, K. N., Mosig, G., Hall, D. W., Eislering, F. A., Black, L. W., Eddy, S., Parma, D., and Gold, L. (1980) J. Mol. Biol. 213, 749–761
6. Frey, M. W., Nossal, N. G., Capsen, T. L., and Benkovic, S. J. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 2579–2583
7. Joyce, C. M., and Steitz, T. A. (1990) J. Bacteriol. 177, 6321–6329
8. Russell, R. L. (1974) Genetics 78, 967–988
9. O’Donnell, P. V., and Karam, J. D. (1972) J. Virol. 9, 990–998
10. Miller, C. C., Yeh, L.-S., and Gold, L. (1985) J. Biol. Chem. 270, 26558–26564
11. Wood, W. B. (1966) J. Mol. Biol. 16, 118–133
12. Studier, F. W., and Moffatt, B. A. (1986) J. Mol. Biol. 189, 119–139
13. Samodel and Brenner, S. (1967) Nature 214, 452–453
14. Hughes, M. B., Yee, A. M. F., Dawson, M., and Karam, J. (1987) Genetics 115, 395–403
15. Gill, S. C., and von Hippel, P. H. (1989) Anal. Biochem. 182, 319–328
16. Lin, T.-C., Karam, G., and Konigsberg, W. H. (1994) J. Biol. Chem. 269, 12926–12930
17. Frey, M. W., Nossal, N. G., Capsen, T. L., and Benkovic, S. J. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 2579–2583
18. Russell, R. L., and Hosken, R. J. (1974) Genetics 78, 988–1014
19. Russell, R. L. (1967) Speciation among the T-even Bacteriophages, Ph.D. thesis, California Institute of Technology, University Microfilms Inc., Ann Arbor, Ml
20. Hsu, T., and Karam, J. D. (1990) J. Biol. Chem. 265, 5306–5316
21. Gold, N., Gayle, M., Sweeney, R., Hollingsworth, T., Modeer, T., and Gold, L. (1988) J. Mol. Biol. 199, 241–258
22. Miller, E. S., and Jozwik, C. E. (1990) J. Bacteriol. 172, 5180–5186
23. Jozwik, C. E., and Miller, E. S. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 5053–5057
24. Miller, C. C., and Gold, L. (1990) Science 249, 505–510
25. Miller, E. S., Karam, J. D., and Spizer, R. E. (1994) in Molecular Biology of Bacteriophage T4 (Karam, J. D., Drake, J. W., Kreuzer, K. N., Mosig, G., Hall, D. W., Eislering, F. A., Black, L. W., Eddy, S., Kather, E., Carlson, K., and Miller, E. S., eds) pp. 193–205, American Society for Microbiology, Washington, D. C.
26. Russell, M., Gold, L., Morissette, H., and O’Farrell, P. Z. (1976) J. Biol. Chem. 251, 7263–7270
27. Mudd, A. A., Prengi, P., Belin, D., and Kreisch, H. M. (1988) EMBO J. 7, 3601–3607
28. Halle, D. J., Roault, A. T., Tang, C. K., Chin, J., Harford, J. B., and Klauser, M. (1992) Proc. Natl. Acad. Sci. U. S. A. 89, 7346–7350
29. Basilion, J. P., Roault, A. T., Massinopat, C. M., Klauser, R. D., and Burgess, W. H. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 574–578
30. Philpot, C., Klauser, R. D., and Roault, A. T. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 7321–7325
31. Chu, E., Voeller, D., Koeller, D. M., Drake, J. C., Takimoto, C. H., Maley, G. F., Maley, F., and Allegre, C. J. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 517–521
RNA Specificity Variant of T4 DNA Polymerase

34. Chu, E., Takimoto, C. H., Voeller, D., Grem, J. L., and Allegra, C. J. (1993) *Biochemistry* **32**, 4756–4760.
35. Singh, R., and Green, M. R. (1993) *Science* **259**, 365–368.
36. Chu, E., Koeller, D. M., Casey, J. L., Drake, J. C., Chabner, B. A., Elwood, P. C., Zinn, S., and Allegra, C. J. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88**, 8977–8981.
37. Preiss, T., Hall, A. G., and Lightowlers, R. N. (1993) *J. Biol. Chem.* **268**, 24523–24526.
38. Grosse, F., Nasheuer, H.-P., Scholtissek, S., and Schomburg, U. (1986) *Eur. J. Biochem.* **160**, 459–477.
39. Hentze, M. W. (1994) *Trends Biochem. Sci.* **19**, 101–103.

40. Draper, D. (1995) *Annu. Rev. Biochem.* **64**, 593–620.
41. Gold, L., Polisky, B., Uhlenbeck, O., and Yarus, M. (1995) *Annu. Rev. Biochem.* **64**, 763–797.
42. Wang, C.-C. (1997) *Structural Organization and RNA-Binding Properties of the DNA Polymerases of Bacteriophages T4 and RB69*. Ph.D. thesis, Tulane University, New Orleans, LA (available from University Microfilms Inc., Ann Arbor, MI).
43. Wang, J., Sattar, A. K. M. A., Wang, C.-C., Karam, J. D., Konigsberg, W. H., and Steitz, T. A. (1997) *Cell*, in press.
44. Dressman, H. K., Wang, C.-C., Karam, J. D., and Drake, J. W. (1997) *Proc. Natl. Acad. Sci. U. S. A.*, in press.