**TRAP, the trp RNA-binding attenuation protein of Bacillus subtilis, is a toroid-shaped molecule that binds transcripts containing GAG or UAG repeats separated by two nucleotides**

(transcription regulation / protein–RNA recognition / protein structure)

**Paul Babitzke*,†, David G. Bear‡, and Charles Yanofsky***

*Department of Biological Sciences, Stanford University, Stanford, CA 94305; and ‡Department of Cell Biology and the Cancer Research and Treatment Center, University of New Mexico Health Sciences Center, Albuquerque, NM 87131

**Contributed by Charles Yanofsky, May 4, 1995**

**ABSTRACT** The trp RNA-binding attenuation protein of *Bacillus subtilis*, TRAP, regulates both transcription and translation by binding to specific transcript sequences. The optimal transcript sequences required for TRAP binding were determined by measuring complex formation between purified TRAP protein and synthetic RNAs. RNAs were tested that contained repeats of different trinucleotide sequences, with differing spacing between the repeats. A transcript containing GAG repeats separated by two-nucleotide spacers was bound most tightly. In addition, transmission electron microscopy was used to examine the structure of TRAP and the TRAP–transcript complex. TRAP was observed to be a toroid-shaped oligomer when free or when bound to either a natural or a synthetic RNA.

Expression of the tryptophan biosynthetic genes of *Bacillus subtilis* (1–6) is regulated in response to the intracellular concentration of L-tryptophan. When the trp RNA-binding attenuation protein of *B. subtilis*, TRAP, is activated by tryptophan, it binds to a segment of the nascent trpEDCFBA operon leader transcript that includes an RNA secondary structure, the antiterminator (see Fig. 1a). Disruption of the antiterminator favors formation of an overlapping factor-independent terminator structure (see Fig. 1a), resulting in transcription termination at a site just preceding the trp operon structural genes (3). In addition to regulating transcription, TRAP can also regulate translation. TRAP can bind to trp operon transcripts that escape termination and indirectly inhibit trpE translation (ref. 4; E. Merino, P.B., and C.Y., unpublished data). One trp gene, trpG, is located in a folic acid biosynthetic operon (7). trpG translation appears to be regulated by a TRAP binding site that overlaps the trpG ribosome binding site (see Fig. 1b) (2, 7). The trp operon leader binding site is composed of 11 (G/U)AG repeats (see Fig. 1b) (2). This recognition sequence correlates with the TRAP quaternary structure in which 11 identical subunits are arranged around an elevenfold axis of symmetry; one molecule of tryptophan is bound per subunit (8, 9). In the present study, we have determined the optimum binding site for TRAP and have used transmission EM to image the oligomeric structure of the protein in the presence and absence of RNA.

**MATERIALS AND METHODS**

**Plasmids.** Plasmid pTZ18U, containing a T7 RNA polymerase promoter (United States Biochemical), was used to construct the trp leader transcription templates used in these studies (2). Specific templates were constructed by ligating double-stranded oligonucleotide cassettes or single-stranded oligonucleotides flanked by EcoRI and Kpn I cohesive ends into the pTZ18U polylinker. Single-strand gap repair was performed with T4 DNA polymerase. The oligonucleotides used were designed so that the insert DNA would contain 6 trinucleotide repeats separated by appropriate spacer nucleotides (see Table 1). All pTZ18U-derived templates were confirmed by DNA sequencing. Transformation and DNA isolation were performed as described (1). Plasmid pH7050A, containing the *Escherichia coli* lac promoter, an 80- to 90-nt linker region, and 380 bp of the GAGA repeat, was constructed by cloning a satellite DNA sequence from *Drosophila melanogaster* (10) into the EcoRI and Sal I sites of the PHV100 polylinker (11).

**RNA Synthesis.** Labeled runoff transcripts were synthesized in vitro using T7 RNA polymerase according to the manufacturer’s protocol (New England Biolabs) from DNA templates linearized with the appropriate restriction enzyme. The labeling nucleotide was [5,6-3H]UTP (8 Ci/mmol; 1 Ci = 37 GBq). Unlabeled runoff transcripts were synthesized using the Ambion (Austin, TX) MEGAScript transcription kit. Transcripts were separated from unincorporated nucleotides by gel filtration. Unlabeled poly(GAGAA) transcript was synthesized by *E. coli* RNA polymerase from pH705A linearized with Sal I in a reaction mixture containing 40 mM Tris-HCl (pH 7.9); 10 mM MgCl2; 100 mM KCl; 0.1 mM dithiothreitol; 5% (vol/vol) glycerol; 200 μM CAMP; 1 μM *E. coli* catalytic repressor protein (gift of William McClure, Carnegie–Mellon University); 100 μM ATP, GTP, CTP, and UTP; and 118 mM *E. coli* RNA polymerase. The transcript was purified by phenol extraction and ethanol precipitation.

**Filter Binding Assays.** TRAP was purified as described (3). Standard filter binding reaction mixtures (0.1 ml) were used except that 2.5 pmol of [3H]RNA was used (2). Time course experiments were used to determine the stability of various TRAP–RNA complexes followed a published procedure (2). Experimental details are described in the appropriate figure legends.

**EM.** EM imaging was carried out with TRAP–tryptophan complexes, TRAP–trypnophan–RNA complexes, and TRAP–tryptophan–satellite RNA complexes. Binding was carried out in a vol of 100 μl in a buffer containing 40 mM HepesNaOH (pH 7.5), 1 mM dithiothreitol, 4 mM MgCl2, 250 mM KCl, and 1.25 mM L-tryptophan. Samples of TRAP used in the uranyl acetate staining method (see below) contained 212 nM TRAP protein (monomer concentration). The sample of TRAP–poly(GAGAA) complex used in the uranyl acetate staining method contained 212 nM TRAP and ~28 μM poly(GAGAA) (in terms of nucleotide monomer). The concentration of poly(GAGAA) was estimated by using an extinction coefficient of 9240 cm⁻¹·M⁻¹, which is the weighted average (3A/2G) of the extinction coefficients for poly(rA) and poly(rG) of monomer (Pharmacia P-L Catalog, 1984). In the rotary shadowing transmission EM studies of the TRAP–trp leader RNA complexes, the concentration of TRAP was 2.12 μM and the 737-nt trp leader RNA was 77 nM.

*Present address: Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802.
Two different methods for contrast enhancement in the imaging of TRAP and TRAP-RNA complexes were used: negative staining with uranyl acetate and rotary shadowing with tungsten. In the uranyl acetate method, we applied a 5-μl drop of either TRAP or TRAP-poly(GAGAA) to a glow-discharged carbon film-covered grid for 30 sec. Then 5 μl of a freshly made uranyl acetate solution (1% in H2O) was applied to the grid for 30 sec and removed. Fresh stain was applied and removed a total of seven times. The grid was air-dried and imaged immediately. Mounting of samples by the rotary shadowing method has been described (12). Imaging was carried out in a Hitachi H600-III transmission EM at 50 kV.

RESULTS

TRAP Binds to RNA Segments Consisting of GAG or UAG Repeats Separated by 2 nt. Previous studies revealed an RNA binding mechanism in which TRAP interacts with seven GAG and four UAG repeats in a segment of the trp operon leader transcript (2). Seven of the spacers separating adjacent repeats are 2 nt long, while the remaining three are 3 nt each (Fig. 1) (2). TRAP also binds to an RNA segment that includes the trpG ribosome binding site (2, 7). The trpG binding site contains seven GAGs and one UAG, although the nucleotide spacing between adjacent repeats is more variable than in the trp leader transcript: four of the trpG spacers are 2 nt, while the remaining three are 5, 7, and 8 nt (2). Within the 7-nt spacer, the central three residues are AAG, suggesting that in this arrangement AAG may serve as a TRAP recognition sequence (2, 13). One of the B. pumilus spacers is 3 nt, while the rest are 2 nt (2). The spacer nucleotides in all of the transcripts mentioned are predominantly A and U (2).

To define the transcript binding site requirements more definitively, we performed filter binding experiments with RNA targets that have defined sequences. Our strategy was to construct a series of plasmids that could serve as templates for in vitro synthesis of short RNAs. The transcripts contained six trinucleotide repeats, each separated by the 2-nt AU. AU was chosen as the spacer sequence for this analysis to prevent the inadvertent generation of other trinucleotide repeats containing AG in the second and third positions. Starting with the repeat trinucleotide sequence GAG, we varied each nucleotide one at a time, producing the series of artificial RNAs listed in Table 1. Each of these RNAs was then examined in filter binding experiments with tryptophan-activated TRAP. Our findings show that RNA containing GAG repeats separated by 2 nt serves as the best TRAP target. TRAP did not bind GAG repeat RNA when the spacers separating repeats were 1, 3, or 4 nt. The only additional trinucleotide repeat RNA that was bound appreciably contained UAG instead of GAG. Note that significant binding to TRAP did not occur with the AAG repeat transcript.

Previously, it was observed that an RNA species that contained the first 6 trinucleotide repeats of trp leader RNA, 3 GAGs and 3 UAGs, formed a TRAP complex with lower stability than the complex formed with wild-type trp leader RNA containing 11 repeats, 7 GAGs and 4 UAGs (2). A truncated trp leader transcript containing only the first three repeats, two GAGs and one UAG, did not form a stable complex with TRAP (2). To determine the relative stabilities

![Fig. 1](image-url)  
(a) Sequence of the B. subtilis trp operon leader transcript showing the presumed mutually exclusive antiterminator and terminator structures. Boxed nucleotides mark overlapping segments of the competing secondary structures. GAG and UAG repeat sequences are indicated by boldface type. (b) Sequence comparison of the TRAP binding sites in the B. subtilis trp leader and trpG transcripts and the putative binding site in the homologous Bacillus pumilus trp leader transcript (2, 13). GAG and UAG repeat sequences are indicated by boldface type. AAG repeat sequences in the B. subtilis trpG and B. pumilus trp leader transcripts are underlined. Positions of the trpG ribosome binding site (RBS) and translation start codon (Met) are shown (7).
Table 1. TRAP binding to artificial RNA targets

<table>
<thead>
<tr>
<th>Plasmid</th>
<th>RNA repeat sequence</th>
<th>RNA spacer sequence</th>
<th>cpm</th>
</tr>
</thead>
<tbody>
<tr>
<td>pPB90</td>
<td>GAG</td>
<td>AU</td>
<td>4117 ± 117</td>
</tr>
<tr>
<td>pPB100</td>
<td>UAG</td>
<td>AU</td>
<td>994 ± 89</td>
</tr>
<tr>
<td>pPB110</td>
<td>AAG</td>
<td>AU</td>
<td>95 ± 1</td>
</tr>
<tr>
<td>pPB120</td>
<td>CAG</td>
<td>AU</td>
<td>0</td>
</tr>
<tr>
<td>pPB140</td>
<td>GAA</td>
<td>AU</td>
<td>0</td>
</tr>
<tr>
<td>pPB150</td>
<td>GAC</td>
<td>AU</td>
<td>0</td>
</tr>
<tr>
<td>pPB160</td>
<td>GAA</td>
<td>AU</td>
<td>0</td>
</tr>
<tr>
<td>pPB170</td>
<td>GUG</td>
<td>AU</td>
<td>0</td>
</tr>
<tr>
<td>pPB180</td>
<td>GCG</td>
<td>AU</td>
<td>0</td>
</tr>
<tr>
<td>pPB190</td>
<td>GGG</td>
<td>AU</td>
<td>0</td>
</tr>
<tr>
<td>pPB132</td>
<td>GAG</td>
<td>U</td>
<td>114 ± 17</td>
</tr>
<tr>
<td>pPB202</td>
<td>GAG</td>
<td>AUU</td>
<td>0</td>
</tr>
<tr>
<td>pPB200</td>
<td>GAG</td>
<td>AUAU</td>
<td>0</td>
</tr>
</tbody>
</table>

Labeled transcripts produced from one of several plasmids that had been linearized with BamHI were used in each analysis. Reaction mixtures (0.1 ml) contained 0.27 pmol of TRAP 11-mer, 2.5 pmol of [3H]RNA, 10 units of RNasin, 1 mM L-tryptophan, and 1 mM dithiothreitol in filter binding buffer (2). Mixtures were incubated for 20 min at 37°C and then filtered through BA85 nitrocellulose membranes. The filters were washed twice with 0.1 ml of filter binding buffer, dried, and assayed in a liquid scintillation counter. cpm retained on the filter were corrected for the number of U residues present in each transcript, since UTP was used as the labeling nucleotide. Background cpm in the absence of L-tryptophan were subtracted from each value (generally 200–300 cpm). Each transcript contained six trinucleotide repeats of the indicated sequence separated by the spacer nucleotides listed in the table.

of TRAP–synthetic RNA complexes, we compared complexes formed with transcripts containing 6 GAGs or 6 UAGs (with AU spacers) with complexes formed with the trp leader transcripts containing 6 or 11 repeats, described above. Under the test conditions used (Fig. 2), the amount of labeled trp leader RNA bound in a complex with TRAP would be a function of the extent of initial saturation of TRAP by the unlabeled test RNA, the rate of dissociation of the TRAP–unlabeled RNA complex, and the ability of free unlabeled RNA to compete with the labeled RNA for TRAP. The results (Fig. 2) indicate that the artificial RNA target containing six GAGs (from pPB90) formed a TRAP–RNA complex that was more stable than the complex obtained with the trp leader transcript containing three GAGs and three UAGs (from pPB83). However, the 6 GAG repeat RNA–TRAP complex was less stable than the complex formed with the natural trp leader transcript containing 11 trinucleotide repeats (from pPB77) (Fig. 2). The TRAP–RNA complex formed with the transcript containing six UAGs (from pPB100) was unstable; the values obtained with this transcript were similar to those obtained with a control RNA that does not contain appropriately spaced (G/U)AG repeats (from pPB11).

TRAP Binds to RNA as a Toroid-Shaped Oligomer. Antson et al. (8, 9) have recently reported crystallographic data for complexes between TRAP and trp transcript. Their studies suggest that tryptophan-activated TRAP is a toroid-shaped 11-subunit oligomer with an elevenfold axis of rotation. To visualize the structure of the TRAP–trp RNA complex, we performed EM studies with uranyl acetate to obtain contrast enhancement. Fig. 3a shows a field of TRAP molecules in the presence of 1.25 mM trp. The TRAP molecules are quite uniform in size, but there are at least three different classes of shapes. In agreement with x-ray data (8, 9), we see many toroid-shaped molecules with diameters of 9 ± 1 nm (molecules labeled 1 in Fig. 3a). We also see molecules that are crescent-shaped (molecules labeled 2 in Fig. 3a); these may be due to partial dissociation of the subunits of some of the toroids during sample preparation (the samples were not fixed). Lastly, we see a class of round molecules with two small internal orifices of equal size (molecules labeled 3 in Fig. 3a). These molecules have an outside diameter that is the same size as the toroid. Since the monomeric molecular weight of TRAP is only 8328, and beyond the resolution of conventional transmission EM, all of the molecules in our images must be oligomers.

To determine the structure of tryptophan-activated TRAP in the presence of RNA, we initially used an RNA made from a template with 76 repeats of the 5-base sequence GAAGA (11). This template contains the E. coli lac promoter and was constructed from a clone of a D. melanogaster satellite DNA (10). Transcription of this template in vitro yields a set of long heterogeneous transcripts containing the GAAGA motif. EM studies of uranyl acetate-stained complexes of activated TRAP and poly(GAGAA) RNA revealed a complex fibrous network of filaments with diameters of 8–9 nm. A few partially saturated TRAP–RNA complexes were observed in which TRAP oligomers were present in an extended linear array. Fig. 3b shows one such field where several RNA molecules are highly oriented, presumably by surface tension forces on the grid during the mounting procedure. RNA is difficult to image by uranyl acetate staining (14); in only a few instances could we see the poly(GAGAA) RNA. However, TRAP oligomers bound along the poly(GAGAA) RNA clearly retain their rounded shape, although not all bound molecules appear to be perfect toroids. Again, this could be due to damage to some of the molecules during mounting.

To view the TRAP–RNA complex directly, we imaged glutaraldehyde-fixed complexes between TRAP and either poly(GAGAA) RNA or a 737-nucleotide transcript containing the B. subtilis trp leader sequence (Fig. 1) (3). We used tungsten rotary shadowing for contrast enhancement of the RNA (12, 14). Under nonsaturating conditions, we confirmed that TRAP retains a toroid shape when bound to poly(GAGAA).
however, the poly(GAGAA) was not readily visible (data not shown). We could readily see some DNA templates that were still present in our RNA preparations. Even at TRAP concentrations in excess of saturation, we saw no TRAP–DNA interactions.

In contrast to our poor shadowing results with TRAP–poly(GAGAA) complexes, we readily observed images of TRAP–trp leader RNA complexes with sufficient resolution to identify a long RNA tail emanating from one side of the TRAP oligomer (Fig. 4). The TRAP binding site in the 737-nt transcript containing the trp leader sequence is nt 451–506. We see numerous examples of complexes with a toroid-shaped TRAP molecule, consistent with the interpretation that the RNA molecule is wrapped around the toroid (see labeled molecules in Fig. 4). All bound TRAP molecules appear to be of uniform shape. We did not detect any RNA molecule with more than one TRAP oligomer bound. We conclude that in these images TRAP is interacting with its specific binding site in trp leader RNA and that TRAP retains its original toroid shape in this complex.

DISCUSSION

Analyses of the TRAP binding sites in trp operon leader RNA and trpG RNA led to the identification of several closely spaced repeated GAG and UAG sequences as the sites of TRAP interaction (2). Our in vitro binding studies with synthetic RNAs indicate that an RNA segment containing GAG repeats with 2-nt spacers serves as the best TRAP binding site (Table 1 and Fig. 2). TRAP bound more weakly to transcripts containing UAG repeats (Table 1 and Fig. 2). This result is in agreement with previous footprint data, which showed that each nucleotide in the GAG repeats was protected by bound TRAP, whereas only AG was protected in UAG repeats. While the spacer nucleotides were generally unprotected in footprint analyses (2), it is not known whether there is a preferred spacer sequence. TRAP bound poorly or not at all to similar transcripts containing AAG, CAG, GCG, GGG, GUG, GAA, GAC, or GAU repeats (Table 1). Although TRAP bound very poorly to the transcript that contained only AAG repeats (Table 1), an AAG might contribute to a transcript’s overall binding effectiveness if the adjacent repeats were GAG or UAG. The AAG present in the B. subtilis trpG transcript is surrounded by two GAG repeats, while the B. pumilus trpG leader transcript contains an AAG with a UAG repeat just upstream and a GAG repeat just downstream (Fig. 1). The context of a GAG, UAG, and AAG repeat may therefore affect a transcript’s interaction with TRAP.

We also found that an RNA containing six GAGs separated by AU spacers formed a stable complex with TRAP, whereas a TRAP complex formed with RNA containing six UAGs was relatively unstable (Fig. 2). An RNA species containing three GAGs and three UAGs formed a TRAP–RNA complex of intermediate stability (Fig. 2). These results indicate that GAG repeats contribute more significantly to the stability of TRAP–RNA complexes than UAG repeats, a result consistent with previous footprint data (2) and the data presented in Table 1. We also determined that the spacing between adjacent trinucleotide repeats is critical. TRAP bound very effectively to RNA containing GAGs separated by 2 nt. However, TRAP did not bind to transcripts containing GAGs that were separated by 1, 3, or 4 nt, indicating that 2-nt spacers are optimal for TRAP recognition (Table 1). It is unlikely that the inability of TRAP to bind tightly to transcripts containing GAG repeats
separated by 1, 3, or 4 nt is due to RNA secondary structure. The predicted stabilities of the secondary structures (20) of the tested transcripts containing 1-, 2-, 3-, and 4-nt spacers are -3.7, -3.7, -3.7, and -5.2 kcal/mol (1 cal = 4.184 J), respectively. In no case are any of the GAGs predicted to be base paired. However, it seems likely that GAGs that are separated by >2 nt in the \textit{B. subtilis} \textit{trp} leader and \textit{trpG} transcripts do contribute to TRAP binding (Fig. 1). Their respective contribution may depend on whether neighboring repeats are GAG or UAG and whether upstream and downstream repeats are optimally spaced.

We can infer several features of the structure of the TRAP–RNA complex from the transmission EM images. First, TRAP retains its toroidal shape upon binding to RNA. Since the hole is still quite apparent in many of the molecules that we have examined, it is quite likely that the RNA is bound around the perimeter of the toroid. Thus, the protein has the capacity to bind a very large contiguous segment of RNA. We estimate the diameter of the toroid molecule to be 90 Å; thus, the circumference is \( \approx 283 \) Å. Several single-stranded nucleic acid binding proteins have been found to increase the spacing of the bases upon binding; for example, the phage T4 gp32 expands the distance between bases in single-stranded DNA to 4.6 Å (15), while phage fd gp5 expands the distance between bases to an average of 5 Å (16). If we assume that TRAP (8.328 kDa per subunit), similar in size to gp5 (9.69 kDa per subunit), increases the base separation to \( \approx 5 \) Å, then TRAP should complex with 57 nt if the RNA is wrapped around its perimeter. The distance between the 1st nucleotide in the 1st repeat (UAG) to the 3rd nucleotide in the 11th repeat (GAG) in the \textit{trp} leader transcript is 56 nt (Fig. 1). A binding site size of \( \approx 5 \) nt for the TRAP monomer is quite reasonable, since fd gp5 has a monomeric binding site size of 4 nt for single-stranded DNA (16).

Our data are consistent with a model in which a continuous segment of \textit{trp} leader RNA is wrapped around the perimeter of TRAP. In our TRAP–RNA images, the bound \textit{trp} leader RNA appears to emanate from only one segment on the surface of the TRAP toroid. The 737-nt \textit{trp} leader transcript used in the transmission EM studies has its 1st UAG site at +451, and its 11th and last GAG site at +506. If resolution had permitted, we might have seen the 11-mer RNA entering and exiting from the TRAP oligomer at adjacent subunits. We suspect that sequences on the \( s' \) and \( s'' \) boundaries of the RNA binding site of the transcript are base paired in our images.

The \textit{E. coli} Rho protein and the \textit{B. subtilis} TRAP protein appear to share both structural properties and some biochemical activities. Both proteins are toroidal-shaped oligomers (9, 12). Rho can bind a 70- to 80-nt contiguous segment of poly(rC), and it retains its toroidal shape upon binding this RNA. Binding of ATP to Rho increases the protein’s affinity for structured RNA (17) and thereby facilitates opening of the RNA structure. In addition, Rho is an ATPase-dependent RNA–RNA helicase that can separate short segments of RNA–RNA hybrids (18). Similarly, tryptophan dramatically increases the affinity of TRAP for RNA (19). Tight binding could destabilize any RNA secondary structure containing (G/U)AG repeats. One such RNA structure that is crucial in attenuation regulation, the antiterminator, is predicted to form in the \textit{trp} operon leader transcript (Fig. 1).

We are indebted to Catherine Squires for her interest in this project and for her helpful suggestions. We also thank Philip C. Hanawalt and Peter Margolis for their comments on this manuscript. We thank Paul Gollnick for exchanging information prior to publication. This work was supported by National Institutes of Health Grant GM09738 to C.Y., by National Science Foundation Grant MCB 9408231 and Dedicated Health Research Funds of the University of New Mexico School of Medicine to D.G.B., and by institutional funds from The Pennsylvania State University to P.B. C.Y. is a Career Investigator of the American Heart Association.