Many amino acid-related genes in Gram-positive bacteria are regulated by the T box riboswitch. The leader RNA of genes in the T box family controls the expression of downstream genes by monitoring the aminoacylation status of the cognate tRNA. Previous studies identified a three-nucleotide codon, termed the “Specifier Sequence,” in the riboswitch that corresponds to the amino acid identity of the downstream genes. Pairing of the Specifier Sequence with the anticodon of the cognate tRNA is the primary determinant of specific tRNA recognition. This interaction mimics codon-anticodon pairing in translation but occurs in the absence of the ribosome. The goal of the current study was to determine the effect of a full range of mismatches for comparison with codon recognition in translation. Mutations were individually introduced into the Specifier Sequence of the glyQS leader RNA and tRNA<sub>Gly</sub> to test the effect of all possible pairing combinations on tRNA binding affinity and antitermination efficiency. The functional role of the conserved purine 3′ of the Specifier Sequence was also verified in this study. We found that substitutions at the Specifier Sequence resulted in reduced binding, the magnitude of which correlates well with the predicted stability of the RNA-RNA pairing. However, the tolerance for specific mismatches in antitermination was generally different from that during decoding, which reveals a unique tRNA recognition pattern in the T box antitermination system.

The T box riboswitch regulates expression of amino acid-related genes in Gram-positive bacteria (1–4). Most genes in the T box family are regulated at the level of transcription attenuation and contain an intrinsic transcriptional terminator in the leader region of the transcript upstream of the start of the regulated coding sequence. Binding of a specific uncharged tRNA to the nascent transcript directs formation of an alternate antiterminator structure and results in readthrough of the termination site and expression of the downstream gene (5, 6). Expression of each gene in the T box family is induced when aminoacylation of the cognate tRNA is reduced; this allows the cell to increase the aminoacylation status of that tRNA by increasing the appropriate amino acid pool or the levels of the cognate aminoacyl-tRNA synthetase. tRNA binding and antitermination occur in the absence of protein factors, indicating that the T box leader RNA is competent for specific tRNA recognition (7, 8).

The specificity of tRNA recognition is mediated in large part by base pairing between the tRNA anticodon and a triplet sequence, designated the “Specifier Sequence” (5), which is located within an internal loop in the Stem I domain of T box leader RNAs (Fig. 1A). The Specifier Sequence triplet corresponds to a codon with the amino acid identity of the downstream coding sequences, and mutational analysis of multiple T box family genes confirmed the model that specificity is determined by specific recognition of the tRNA anticodon in the absence of leader RNA translation (5, 9). In addition, specific tRNA-dependent antitermination in a purified in vitro transcription system (7, 10), binding and structural mapping assays (8), fluorescence spectroscopy assays (11), and structural analyses (12–15) clearly demonstrated codon-anticodon recognition in the absence of the ribosome. Discrimination between uncharged and charged tRNA is mediated by base pairing between the unpaired 3′-end of the tRNA (5′-NCCA-3′) and residues (5′-UGGN-3′) within a bulge in the antiterminator element; the variable residues in the tRNA and antiterminator covary to provide additional specificity (6, 7).

Most of the biochemical studies in this system used the Bacillus subtilis glyQS gene, which encodes glycyl-tRNA synthetase, as a model. The glyQS RNA is a member of a subclass of T box family leader RNAs that lack two major structural elements (Stem II and the Stem IIA/B pseudoknot) that are conserved in the majority of T box leader sequences (Fig. 1A). Efficient tRNA<sub>Gly</sub>-dependent antitermination and leader RNA-tRNA<sub>Gly</sub>
binding are dependent on a match between the Specifier Sequence and the tRNA anticodon (7, 8). Addition of an extra residue at the 3′-end of the tRNA generated a mimic of charged tRNA\(^{\text{Gly}}\) that acts as a competitive inhibitor for binding and antitermination (16). Stem I alone is capable of specific tRNA binding, and the terminal region of Stem I was shown to form an additional interaction with the elbow region of tRNA\(^{\text{Gly}}\) (13, 14). An RNA element containing only the bottom portion of Stem I, including the Specifier Loop, also is capable of specific tRNA recognition (11). Based on these results, it was proposed that the Specifier Sequence-tRNA anticodon pairing is crucial for discrimination between cognate and non-cognate tRNA. Because the correct tRNA must be selected from a pool of incorrect tRNAs \textit{in vivo}, it is likely that the Specifier Sequence is “probed” by both matching and non-matching tRNAs by analogy with aminoacyl-tRNA selection during translation, but only the matching tRNA is capable of a productive interaction. It is unclear, however, whether simple base-pairing stability is sufficient for this discrimination.

The residue 3′ to the Specifier Sequence is conserved as a purine (A or G); this residue was predicted to form a fourth pairing interaction with the conserved U33 tRNA residue based on structural mapping results (8). However, structural studies demonstrated that this purine residue stacks underneath the Specifier Sequence and stabilizes the Specifier Sequence-anticodon helix (14).

Previous analyses of T box gene expression \textit{in vivo} using the \textit{B. subtilis} tyrS gene, which encodes tyrosyl-tRNA synthetase, showed that induction occurs only in response to increased levels of uncharged tRNA\(^{\text{Ty}}\) (5, 9). Accumulation of a noncognate uncharged tRNA, which results from limitation for an amino acid different from that corresponding to the Specifier Sequence, has little effect on expression. Substitution of a different Specifier Sequence in conjunction with the appropriate change in the antiterminator to provide a match at the tRNA 3′-end sometimes allows a response to the corresponding tRNA, but the resulting efficiency of antitermination is generally lower than that observed with the cognate tRNA (9). The \textit{in vivo} studies are complicated by the fact that the cognate and noncognate tRNAs vary at many positions in addition to the Specifier Sequence (7) in which the T at position +2 relative to the transcription start point was substituted with C to allow initiation with the dinucleotide ApC \textit{in vitro}. The resulting plasmid variants were sequenced and used as templates for PCR to generate 440-bp products for use as templates for \textit{in vitro} transcription by \textit{B. subtilis} RNA polymerase (7). Templates for T7 RNA polymerase transcription (200 bp) were generated by using a 5′-primer containing a T7 RNA polymerase promoter initiating with tandem G residues fused to position +1 of the glyQS leader RNA and a 3′-primer ending at glyQS position +183 (3′ of the antiterminator helix). DNA templates for generation of variants of tRNA\(^{\text{Gly}}\) (Table 1) were generated as described previously (7) using appropriate oligonucleotides to obtain the desired sequence changes.

T7 RNA polymerase transcription was carried out with an Ampliscribe T7 transcription kit (Epigenic Biotechnologies). \(^{32}\text{P}\)-Labeled tRNA was generated by addition of \([\alpha-^{32}\text{P}]\text{UTP}\) (GE Healthcare; 800 Ci/mmol) at 0.63 \(\mu\text{M}\) final concentration. RNA products were purified by electrophoresis in a denaturing 6% (w/v) polyacrylamide gel and refolded by incubation at 80 °C and slow cooling to room temperature. RNA concentration was determined using an ND-1000 spectrophotometer (NanoDrop Technologies, Inc.). Quantitation of \(^{[32}\text{P}]\text{tRNAGly}\) was carried out by spotting 1-μl samples of dilutions of labeled tRNA onto a 1% agarose slab gel containing ethidium bromide (0.5 μg/ml); dilutions of a tRNA preparation of known concentration was used as a standard. The Bio Doc-It Imaging System (UVP, Upland, CA) and ImageQuant software (Molecular Dynamics) were used to compare fluorescence intensities and estimate tRNA concentrations.

\textbf{tRNA Binding Assays—}Binding assays were performed by mixing a 50 nM concentration of each \(^{32}\text{P}\)-labeled tRNA\(^{\text{Gly}}\) variant with different concentrations of the leader RNA variants in 1× transcription buffer as described previously (8). The reaction mixture was heated to 65 °C for 5 min and slow cooled to 40 °C. Samples (100 μl) were filtered through a Nanosep 30,000 molecular weight cutoff Omega filter microcentrator (Life Sciences) and washed six times with 1× transcription buffer (150 μl/wash). Material retained by the filter was mixed with Packard BioScience Ultima Gold scintillation fluid and counted in a Packard Tri-Carb 2100TR liquid scintillation counter. Each experiment was repeated twice with the experimental points run in duplicate. The specific activity (cpm/pmol) of each \(^{[32}\text{P}]\text{tRNAGly}\) variant was used to calculate the picomoles retained. Nonlinear regression analysis was performed using...
GraphPad Prism version 4.0 (GraphPad Software). The $K_d$ (μM) was calculated using the following equation.

$$K_d = [\text{glyQS leader RNA}] \times ([\text{tRNA}_{\text{bound}}] - [\text{tRNA}_{\text{bound}}]) / [\text{tRNA}_{\text{bound}}]$$  

(Eq. 1)

**In Vitro Transcription Antitermination Assays**—Single round in vitro transcription assays were carried out with *B. subtilis* RNA polymerase as described previously (7). The initiation reaction included template DNA (3 nM), *B. subtilis* RNA polymerase (6 nM), ApC (150 μM, Sigma), ATP and GTP (2.5 μM each), UTP (0.75 μM), and [$\alpha$-32P]UTP (0.25 μM; 800 Ci/mmol); omission of CTP from the initiation reaction resulted in a halt at position +17. The initiation mixtures were incubated at 37 °C for 15 min, heparin (25 nM; Sigma) was added to block reinitiation, and MgCl2 was added to a final concentration of 40 mM. The tRNA$^{Gly}$ variants were added to the reaction mixture as indicated, and elongation was restarted.
by the addition of NTPs to 10 μM final concentration. Transcription reactions were terminated by extraction with phenol, and the products were resolved by denaturing 6% polyacrylamide gel electrophoresis and visualized and quantitated by PhosphorImager analysis (GE Healthcare). Antitermination efficiency was calculated from the fraction of product at the position of the readthrough band relative to the total of the terminated and readthrough products. Each experimental point was a result of three independent experiments. Nonlinear regression analysis was performed using GraphPad Prism version 4.0. The $K_{1/2}$ was calculated using the following equation:

$$K_{1/2} = \frac{[\text{tRNA}] \times (RT_{\text{max}} - [%\text{RT}])}{[%\text{RT}]} \quad (\text{Eq. 2})$$

The $RT_{\text{max}}$ (%) is defined as the maximum amount of percent readthrough estimated by extrapolation of a hyperbolic curve, and the $K_{1/2}$ (μM) is defined as the tRNA concentration that is required to reach half of the maximum readthrough. $K_{1/2}$ is used as a parameter to evaluate the antitermination efficiency.

Results

Mutations in the Specifier Sequence and tRNA Anticodon—Codon-anticodon recognition in the T box system was tested by generation of all single change Specifier Sequence variants in the glyQS leader and tRNA<sup>Gly</sup> anticodon and testing leader RNA/tRNA combinations in binding and in vitro transcription antitermination assays. Each set of variants was compared with the wild-type combination (GGC Specifier Sequence/GCC anticodon), which showed the highest activity in both tRNA binding and antitermination. Binding was quantitated by measurement of the dissociation constant ($K_d$), and antitermination efficiency ($K_{1/2}$) was determined as the amount of tRNA required for half-maximal antitermination (Fig. 2, B and D).

An example of this analysis is shown in Fig. 2. Introduction of a 34G→A mutation in tRNA<sup>Gly</sup> resulted in a C/A mismatch at position 3 of the Specifier Sequence, which reduced the binding affinity and antitermination efficiency by 3.8- and 3.4-fold, respectively (see supplemental Table S1). The 34G→C and 34G→U mutations caused a more dramatic reduction in activity, particularly in antitermination efficiency (11–12-fold reduction). The combinations of variants were grouped according to the position of the mismatch in the Specifier Sequence (Fig. 2, C and D, and supplemental Table S1). The binding and antitermination activity patterns were complex, and the effect of different mismatches depended on the position and type of mismatch. For both assays, mismatches at position 2 of the Specifier Sequence resulted in the most severe reduction in activity. Mismatches at positions 1 and 3 in general behaved similarly in the binding assay, whereas larger differences were observed in the antitermination assay. The magnitude of the
Codon-Anticodon Recognition by the T Box Riboswitch

reduction in function also was more extreme in the antitermination assay, indicating that this is a more sensitive measurement of tRNA interaction. This is consistent with the observation that glyQ’s leader RNA-tRNA^{Gly}_U binding is strongly dependent on the Stem 1-tRNA interaction, whereas antitermination is dependent not only on that interaction but also on proper positioning of the acceptor end of the tRNA for pairing with the antiterminator bulge (8). This is also consistent with the fact that antitermination requires the interaction of the ligand to occur during transcription, whereas the binding assay allows a longer time for interaction.

**Mutations at Position 3 of the Specifier Sequence-Anticodon Pairing—** In translation, mismatches at the third position of the codon are often tolerated, allowing a particular tRNA to decode more than one codon; in particular, wobble pairs (U-G or G-U) are efficiently recognized at this position. Recognition of a UAU tyrosine codon in *B. subtilis* in vivo utilizes a U-G pair at this position as *B. subtilis* lacks a variant of tRNA^{Tyr} with an AUA anticodon (17); the pairing at this position in vivo is further complicated by tRNA modification. Phylogenetic analysis of T box leader sequences revealed a strong preference for a C residue at position 3 of the Specifier Sequence (the “C rule”); this preference is independent of codon preferences in the overall genome (2, 5). Substitution of a U for the C at position 3 of the UAC Specifier Sequence in the *B. subtilis* tyrS leader sequence resulted in decreased expression in vivo (18).

In glyQ’s, the GGG/CCC combination in which the position 3 pairing is switched from the wild-type C-G to G-C was the only combination that resulted in binding affinity similar to the wild-type C-G pair. The pairing at this position in vivo utilizes a U-G pair at this position as *B. subtilis* lacks a variant of tRNA^{Tyr} with an AUA anticodon (17); the pairing at this position in vivo is further complicated by tRNA modification. Phylogenetic analysis of T box leader sequences revealed a strong preference for a C residue at position 3 of the Specifier Sequence (the “C rule”); this preference is independent of codon preferences in the overall genome (2, 5). Substitution of a U for the C at position 3 of the UAC Specifier Sequence in the *B. subtilis* tyrS leader sequence resulted in decreased expression in vivo (18).

In glyQ’s, the GGG/CCC combination in which the position 3 pairing is switched from the wild-type C-G to G-C was the only combination that resulted in binding affinity similar to the wild-type combination (GGC/CCG) (Fig. 2C). The other combinations that retained Watson-Crick (W-C) pairing (A-U and U-A) showed reduced binding (3.6- and 3.4-fold relative to wild-type, respectively). The remaining combinations (with the exception of the pyrimidine-pyrimidine (Y-Y) pairs) exhibited a 2.2-4.2-fold reduction of binding. The wobble pairs (G-U and U-G) were slightly more effective than the W-C A-U and U-A pairings and G-G, A-G, and G-A purine-purine (R-R) pairings exhibited binding similar to that of the A-U and U-A W-C pairings. The remaining non-W-C pairings exhibited lesser effects in binding (3.8-6.4-fold) with the Y-Y combinations exhibiting the largest effects. These results show that R-R pairings are surprisingly well tolerated at position 3.

Although the *K_a* values for binding varied over a 6.4-fold range relative to the wild-type pairing, the *K_{1/2}* obtained by measurement of antitermination activity had a much broader range of effects to over 20-fold (Fig. 2D). Consistent with the binding assays, replacement of the wild-type C-G pair with the G-C W-C pair was the most efficient (2.0-fold reduction); however, the A-U pairing was now more efficient than U-A (2.2-5.2-fold reduction) and the wobble pairs (3.8-fold reduction for G-U and 5.8-fold for U-G). The A-G and G-G pairings, which had a modest effect on binding, resulted in a more serious defect in antitermination (11- and 14-fold reduction) similar to that of the Y-Y pairings, which had the greatest defect in activity. The remaining R-R pairs (A-A and G-A) together with C-A had a modest effect on antitermination. C at position 3 had a greater defect in antitermination versus binding such that C-A and C-U exhibited higher activity than A-C and U-C, respectively.

The relative efficiencies for each of the position 3 pairings in binding and antitermination are shown in Table 2. The general order follows a similar pattern in the two assays with notable differences in the defects for Y-R pairings (U-G and U-A), which are better tolerated for binding than for antitermination (2.4-3.4-fold reduction in binding and 5.2-5.8-fold reduction in antitermination; Table S1). The order for binding in general follows what would be expected based on predicted base pair stability (see below), whereas this is less apparent in the antitermination results. Binding was also not very sensitive to the polarity of the base pair (e.g. R-Y versus Y-R); in contrast, in antitermination, the wild-type C-G pair was better than G-C, but A-U > U-A, and G-U > U-G, indicating that the wild-type Y-R polarity does not correlate with a general Y-R preference (Table 2).

**Mutations at Position 1 of the Specifier Sequence-Anticodon Pairing—** Mispairing at position 1 of a codon is less well tolerated during translation than is mispairing at position 3 (19). This differential sensitivity was not apparent in the binding assay (Fig. 2C and supplemental Table S1). As was observed for the position 3 variants, substitutions that maintained W-C or wobble pairing generally had modest effects on binding with only slight differences in the order of binding for the W-C and wobble pairs (Table 2). Again, Y-Y mispairs exhibited the weakest binding with the exception of the C-U pair, which showed a defect similar to that of the U-A pair. This differed from position 3 where C-U had a more severe defect in binding than was observed for the U-A pair.

The antitermination assays were more sensitive to the effects of codon-anticodon mismatches in agreement with the results for position 3 variants (Fig. 2D and supplemental Table S1). The most notable exceptions are three of the Y-Y pairs (U-U, C-U, and C-C), which were better tolerated at position 1; however, the U-C mismatch had the greatest defect at both positions 1 and 3. Substituting the wild-type G-C pair with a C-G pair led to a 3.1-fold reduction in activity despite the fact that this change is not predicted to affect the stability of the codon-anticodon interaction (see below). This was similar to the situation at position 3 when C-G was switched to G-C. Activity of the C-G pair at position 1 was comparable with that of a U-A or G-U pair; in contrast, an A-U or U-G pairing resulted in a more serious defect in antitermination. Interestingly, A-U was about 2-fold better tolerated at position 3 than at position 1, whereas U-A had the opposite effect even though the wild-type interaction has Y-R at position 1 and Y-R at position 3. The R-Y pair G-U was more effective than the Y-R pair U-G at both positions. All of the mispairs involving tRNA^{Gly}_{36C}→A (C-A, G-A, and A-A) had a worse defect at position 1 versus position 3.

The overall pattern for position 1 effects is summarized in Table 2. These results indicate that although there are specific differences, mismatches at positions 1 and 3 of the Specifier Sequence generally exhibited a similar pattern of effects. This differs markedly from what is observed during translation.
Mutations at Position 2 of the Specifier Sequence—Mismatches at position 2 of the Specifier Sequence were expected to have more severe defects than mismatches at position 1 or 3 as they would have weaker stacking interactions of adjacent base pairs and could disrupt pairing at all three positions. Consistent with this prediction, we observed a broader range of effects on tRNA binding. Substitution of the wild-type G-C pair with a C-G pair had no effect on binding (Fig. 2C and supplemental Table S1) in agreement with the analogous variants at positions 1 and 3. A-U and U-A W-C pairings resulted in reduced binding as did the G-U and U-G wobble pairings. In contrast to what was observed at positions 1 and 3, R-R pairings at position 2 had a major negative effect on binding and were comparable with the Y-Y pairings. The Y-R mismatched pairs (A-C and C-A) had an intermediate effect on binding relative to the wobble pairs and the R-R or Y-Y pairs.

In antitermination assays, substitution of the G-C pair with a C-G resulted in a significant defect (Fig. 2D and supplemental Table S1) similar to that observed at position 1. Different from what was observed at positions 1 and 3, the U-A pair exhibited activity similar to that of the A-U W-C pair. G-U and U-G wobble pairs were less well tolerated at position 2. All mismatch pairs (R-R, Y-Y, and C-A/A-C) resulted in a more severe defect in antitermination efficiency than was observed at positions 1 and 3. The general pattern for effects of substitutions at position 2 is summarized in Table 2. Effects on binding and antitermination were more consistent at position 2 than for positions 1 and 3. Comparison of Effects at Each Position of the Specifier Sequence—Substitutions at positions 1 and 3 of the Specifier Sequence generally had similar effects on binding (Table 2 and Fig. 2C). Position 2 substitutions caused more severe defects, although the pattern is generally similar with the exception of the R-R pairs, which were more detrimental at position 2 than at positions 1 and 3, and C-U, which had a larger effect at positions 2 and 3 than at position 1. The R-R result could be due to the greater destabilization caused by this mismatch at position 2 as compared with position 1 or 3 (see below). The antitermination activity comparison is more complex (Table 2 and Fig. 2D). At each position, the native C-G pairing is better than any other W-C pairing, including the reverse G-C pair. Any tRNA variant with A at position 1 of the anticodon (which pairs with position 3 of the Specifier Sequence) exhibited more efficient antitermination than the corresponding combinations at the other positions. Although Y-Y pairs at any position resulted in decreased activity, the U-U, C-U, and C-C pairs were better tolerated at position 1 than at position 2 or 3. Also, A-G and G-G exhibited higher activity at position 1 than at position 3. At position 2, only W-C or G-U and U-G wobble pairs gave effective antiter-

<table>
<thead>
<tr>
<th>Position 1</th>
<th>Position 2</th>
<th>Position 3</th>
<th>Reduction in activity</th>
</tr>
</thead>
<tbody>
<tr>
<td>G-C</td>
<td>C-G</td>
<td>G-C</td>
<td>low</td>
</tr>
<tr>
<td>C-G</td>
<td>G-C</td>
<td>C-G</td>
<td></td>
</tr>
<tr>
<td>A-U</td>
<td>U-A</td>
<td>G*U</td>
<td></td>
</tr>
<tr>
<td>G*U</td>
<td>A-U</td>
<td>U*G</td>
<td></td>
</tr>
<tr>
<td>G*G</td>
<td>G*U</td>
<td>G*G</td>
<td></td>
</tr>
<tr>
<td>A*G</td>
<td>U*G</td>
<td>A*G</td>
<td></td>
</tr>
<tr>
<td>U-A</td>
<td>C*C</td>
<td>A-U</td>
<td></td>
</tr>
<tr>
<td>C*U</td>
<td>C*A</td>
<td>U-A</td>
<td></td>
</tr>
<tr>
<td>G*A</td>
<td>A*C</td>
<td>G*A</td>
<td></td>
</tr>
<tr>
<td>A*C</td>
<td>U*C</td>
<td>C*A</td>
<td></td>
</tr>
<tr>
<td>U*G</td>
<td>G*G</td>
<td>A*C</td>
<td></td>
</tr>
<tr>
<td>C*A</td>
<td>A*G</td>
<td>A*A</td>
<td></td>
</tr>
<tr>
<td>A*A</td>
<td>U*U</td>
<td>C*C</td>
<td></td>
</tr>
<tr>
<td>U*U</td>
<td>C*U</td>
<td>C*U</td>
<td></td>
</tr>
<tr>
<td>C*C</td>
<td>G*A</td>
<td>U*U</td>
<td></td>
</tr>
<tr>
<td>U*C</td>
<td>A*A</td>
<td>U*C</td>
<td></td>
</tr>
</tbody>
</table>

*a Reduction in $K_d$ and $K_{1/2}$ at each position is ranked based on results in supplemental Table S1. W-C pairs are connected by “-” and bold, and non-canonical W-C pairs are connected by “●”.

TABLE 2

Effect of pairing interactions at positions 1, 2, or 3 of the Specifier Sequence on binding affinity ($K_d$) and antitermination efficiency ($K_{1/2}$)
Correlation of Codon-Anticodon Stability with Effects on Leader RNA-tRNA Interactions—
The simplest prediction is that stability of the Specifier Sequence-anticodon interaction is the primary determinant for both tRNA binding and antitermination. To determine how the strength of the codon-anticodon interaction can influence the initial selection of the tRNA, we estimated the stability for each Specifier Sequence-anticodon combination (supplemental Table S1) according to the rules derived by using tRNA anticodon-anticodon complexes as a model system (20, 21). The correlation between the pairing stability and the loss of binding or antitermination was determined using Pearson’s correlation coefficient ($r$; Fig. 3). As pointed out by Curran (21), this model system is limited in that it involves interaction of two highly structured molecules and may not represent all RNA-RNA interactions. However, the Specifier Sequence is also presented within the structured context provided by the S turn within the Specifier Loop, lending credibility to application of this model.

The estimated stabilities correlated well with tRNA binding affinities (Fig. 3, A–C) with the highest correlation at position 2 ($r = 0.83$) and similar correlations at positions 1 and 3 ($r = 0.71$ and 0.74, respectively). The stabilities also correlated well with antitermination efficiency at positions 1 and 2 ($r = 0.81$ and 0.91, respectively; Fig. 3, D and E); however, lower correlation was observed at position 3 ($r = 0.59$; Fig. 3F). This is consistent with the observation that mismatches at position 3 showed variable effects on antitermination, although all but the C-C pair are predicted to have similar effects on stability. These results suggest that antitermination involves constraints in addition to base pair stability, particularly at position 3 of the Specifier Sequence.

The more severe defects caused by substitutions at position 2 are consistent with the effects of mispairing at this position on the estimated stability of the codon-anticodon interaction (20, 21). Position 2 mismatches resulted in more severe defects in binding than did mismatches at position 1 or 3 (Fig. 2C), resulting in tighter correlation with stability effects (Fig. 3B). Similarly, position 2 substitutions caused a greater loss of antitermination efficiency relative to that observed for positions 1 and 3 (Fig. 2D) so that the effects on antitermination are more tightly correlated with effects on stability (Fig. 3E). This indicates that the pairing stability at position 2 of the Specifier Sequence directly affects both binding and antitermination.

Effect of the Position 3’ to the Specifier Sequence—
The nucleotide located immediately 3’ of the Specifier Sequence is conserved as a purine, and this position (A102) in the glyQS leader RNA along with the Specifier Sequence is protected from cleavage upon tRNA binding (8). This protection was predicted to be due to pairing between this residue and the base 5’ to the anticodon, which is highly conserved in tRNA as a U, resulting in extension of the Specifier Sequence-anticodon pairing to four residues. However, a base pair at this position was not observed in the crystal structure and NMR studies (14, 15). We tested the extended pairing by varying A102 of the glyQS leader RNA and U33 of tRNAGly. Substitution of the wild-type A-U with a G-U or C-U had little effect on binding, whereas substitution with a U-U resulted in a reduction in binding similar to that observed at positions 1 and 3 (Fig. 4 and supplemental Table S1).
substitution had little effect in the more sensitive antitermination assay, consistent with the phylogenetic data that show equal distribution of A or G at this position (2), whereas both C-U and U-U had reduced activity.

When other combinations were introduced by changing the tRNA sequence, maintenance of A or G at position 102 allowed efficient antitermination regardless of the identity of the corresponding nucleotide in the tRNA with at most a 2-fold effect. Although alteration of the conserved U33 in tRNA can alter the structure of the anticodon loop (22), the tRNA variants used here were able to promote antitermination with high efficiency if position 102 in the leader RNA was a purine, indicating that this change in the anticodon loop structure is tolerated by the T box mechanism at least in the context of tRNA<sup>Gly</sup>. C at position 102 in the C-A, C-U, and C-C pairs led to a 3.4–7.0-fold reduction in antitermination, whereas replacement of U33 with G (to allow pairing with C at position 102) resulted in a more substantial decrease despite the predicted restoration of base pairing. Placement of a U at position 102 resulted in poor antitermination regardless of the identity of the corresponding position in the tRNA. These results are consistent with the observations in the crystal structure and NMR studies in which the conserved purine 3′ of the glycine Specifier Sequence contributes to a stacking interaction to stabilize the Specifier Sequence-anticodon helix rather than formation of a base pair interaction with the conserved U33 of tRNA<sup>Gly</sup>.

**Discussion**

The goal of this study was to determine requirements for Specifier Sequence-anticodon pairing for the glyQS T box riboswitch. Changes were limited to single positions in the Gly GGC codon and the tRNA<sup>Gly</sup> anticodon, and the tRNA and leader RNA backbones outside the anticodon and Specifier Sequence were maintained. The mutant anticodons are therefore not in their natural context and could therefore result in structural perturbations of either the Specifier Loop or tRNA anticodon loop, but these data provide a simple comparison of the effects of various mismatches within a single context.

The previously described strong preference for C at position 3 of the Specifier Sequence was substantiated by the observation that C-G was the most active pairing at this position in glyQS. Interestingly, codons having C in the third position exhibit better relative rate constants for aminoacyl-tRNA selection during protein synthesis (21). *B. subtilis* exhibits some preference for codons with C at the third position (17), but genes regulated by the T box riboswitch often do not use the most common codon as the Specifier Sequence.

The binding data generally fit the prediction that binding efficiency should be directly related to the stability of the Specifier Sequence-anticodon interaction (Fig. 3, A–C). The order of binding preferences exhibited at positions 1 and 3 can be partially explained by the observation that R-R mispairs are more stable than Y-Y mispairs (23). In particular, Y-Y mispairs cause a backbone distortion that would enhance electrostatic repulsion of the phosphates (24). Conversely, for position 2, it is possible that accommodation of the compact Y-Y mispairs could be favored over the bulky R-R mispairs. The tolerance of Y-R mispairs at each position could be due to relatively high stability and reduced distortion of the codon-anticodon mini-helix from W-C geometry. The contribution of flanking leader RNA and tRNA sequences to the final stability is unknown, but Nelson *et al.* (11) observed that interaction between the Specifier Sequence of the glyQS leader RNA and tRNA<sup>Gly</sup> is more stable than expected for codon-anticodon interaction. Recognition of other tRNA<sup>Gly</sup> determinants by the glyQS leader RNA is likely to partially mitigate the adverse effect of mispairing at least in the experimental conditions used during the tRNA binding assay.

Another simple prediction is that antitermination efficiency should be directly correlated with binding efficiency as binding must occur first. However, antitermination efficiency did not correlate well with stability for position 3 substitutions (Fig. 3F). The effect of substitutions on antitermination relative to that on binding identified combinations that differentially affect the two processes (Fig. 5). A high value for this ratio indicates variants that antiterminate less efficiently than would be expected due to decreased binding alone. Most combinations exhibited a more significant defect in antitermination efficiency than in binding (ratio >1.0). There were only a few cases (U-A at posi-
Codon-Anticodon Recognition by the T Box Riboswitch

Codon minihelix, and kinetic proofreading (35, 40, 41). These events ensure the accuracy of translation by amplifying the destabilizing effects of mispairing in the codon-anticodon interaction (31, 42–45). The selection mechanisms act at positions 1 and 2 of the codon where bases A1492, A1493, and G530 of 16S rRNA probe for the correct geometry of a W-C pairing, leaving position 3 free to accept non-W-C interactions. Mispairing rules and extended wobble rules for third position pairing allow recognition of C-A and the Y-Y mismatches (U-U and C-U) in addition to G-U and U-G, whereas A-C, U-C, and the R-R mismatches are recognized at much lower levels, and C-C is the only excluded pair (46). At the first and second positions, Y-R and R-Y mismatches have reversed acceptability compared with the translational rules derived by Lim and Curran (46). The T box system also exhibits a preference for G-U over U-G, which is the opposite of what is found in translation (47–49). However, as in translation (50, 51), mismatches at position 2 were the most deleterious.

Interestingly, the effects of mismatches on antitermination more closely followed the misreading rules obtained with a simplified cell-free translational system (52), which result in lower fidelity conditions (high magnesium and absence of cognate tRNA competition). In this system, R at position 1 or 2 of the codon could be misread more easily than Y, and G was most easily misread. Mispairings at position 2 were better tolerated in the presence of a 5′-U-A base pair as compared with a G-C base pair, a phenomenon known as “context effect” (53–55). However, even under these conditions, the ribosomal A site is more selective than the T box system and exhibits higher discrimination against position 1 mispairing (19, 44).

The P site in which active tRNA selection occurs only during translation initiation can accept mispairing at position 1 and to a smaller extent at position 3 (56, 57). Recognition of unique determinants of the initiator tRNA by 16S rRNA nucleotides in the P site appears to overcome codon-anticodon mispairing (39, 39, 58, 59). Altering the anticodon sequence of the initiator tRNA while maintaining W-C base pairing allows for complete recovery of activity in vivo (60). The T box system exhibited a similar low tolerance for position 3 mispairing, but restoration of base pairing generally did not result in complete restoration of function.

The E site interacts only with decylated tRNAs (61–63). The primary contacts of E site tRNA with the ribosome are with the acceptor end and with bases in the D and T loops. A similar interaction occurs in tRNA recognition by RNase P (64) and the T box leader RNA (13, 14), indicating a general tRNA binding platform. Although there is little codon-anticodon interaction within the E site, the stability of codon-anticodon interactions in the P and E sites significantly contributes to the global accuracy of the elongation process by affecting the occurrence of frameshifting (21, 62, 65–71). Whether any of the additional features of tRNA recognition in the ribosome are in common with those of the T box system remains to be determined.
Any alteration of the Specifier Sequence led to a decrease in antitermination activity even in the presence of the matching mutant tRNA. This could result from a combination of detrimental effects due to the introduction of mutations in the tRNA anticodon and in the Specifier Sequence. A given anticodon is more functional in the presence of the native tRNA structural context (54, 72), and Olejniczak and co-workers (73, 74) suggest that tRNAs have been tuned to their anticodons to reduce incorrect codon-anticodon recognition. Furthermore, whereas the tRNA Glu anticodon is naturally unmodified, modification in vivo might affect the structure of other anticodon sequences. The effect of the alteration of the anticodon alone cannot explain cases where binding was unaffected but antitermination was reduced (e.g. G-C at position 3). During translation, codon-anticodon pairing generates a signal that is transmitted through the tRNA via structural alterations to the tRNA acceptor end, leading to activation of elongation factor Tu and tRNA accommodation in the large ribosomal subunit. Triggering these events through tRNA deformation relies on the proper alignment of cognate codon-anticodon pairings to orient the acceptor end of the tRNA (42). An altered Specifier Sequence-anticodon interaction could lead to altered presentation of the acceptor end to the antiterminator bulge. Support for this hypothesis is provided by in vitro transcription experiments that demonstrated sensitivity of antitermination activity to the orientation of the tRNA acceptor end (75). Differences between binding and antitermination also could be due to the different kinetic constraints as antitermination is likely to require interaction with the tRNA under a much tighter time constraint due to the requirement for co-transcriptional binding as compared with the binding assay, which is more likely to be closer to equilibrium conditions.

The Specifier Loop contains a variant of the S turn RNA structural element (Fig. 1A and Refs. 26 and 76), which might be important for presentation of the Specifier Sequence residues for interaction with the tRNA anticodon (8, 12). Alterations in the Specifier Sequence could affect the optimal arrangement of this domain, therefore affecting presentation of both the Specifier Sequence itself and the acceptor end of the tRNA for interaction with the antiterminator. This could be additionally affected by the conserved purine 3′ to the Specifier Sequence. Although this residue was originally predicted to base pair with the conserved U33 of the tRNA based on structural mapping experiments (8), the x-ray crystal structure and NMR studies show that this base stacks underneath the Specifier Sequence-anticodon helix to stabilize the complex (14, 15). The current study further confirmed the functional role of a purine at this position regardless of the identity of the nucleotide at position 33 of tRNAGlu.

It is highly likely that the final accuracy of tRNA selection by the T box mechanism cannot rely only on the Specifier Sequence-anticodon and acceptor stem-antiterminator bulge interactions. Other portions of the leader RNA and recognition of other tRNA determinant(s), including the Stem I terminal domain/tRNA elbow region, are likely to contribute to selectivity. It is clear, however, that the Specifier Sequence-tRNA anticodon interaction is a dominant feature of T box riboswitch function and that tRNA recognition utilizes mechanisms distinct from those used during translation.

### References

8. Yousef, M. R., Grundy, F. J., and Henkin, T. M. (2005) Structural transitions induced by the interaction between tRNA(Glu) and the *Bacillus subtilis* gysQ5 T box leader RNA. *RNA* 11, 7240–7245
Codon-Anticodon Recognition by the T Box Riboswitch


in complex with tRNA. Nature 468, 784–789
69. Bucklin, D. J., Wills, N. M., Gesteland, R. F., and Atkins, J. F. (2005) P-site pairing subtleties revealed by the effects of different tRNAs on programmed translational bypassing where anticodon re-pairing to mRNA is separated from dissociation. J. Mol. Biol. 345, 39–49
74. Olejniczak, M., and Uhlenbeck, O. C. (2006) tRNA residues that have coevolved with their anticodon to ensure uniform and accurate codon recognition. Biochimie 88, 943–950