NMR structure and dynamics of the Specifier Loop domain from the Bacillus subtilis tyrS T box leader RNA

Jiachen Wang¹, Tina M. Henkin² and Edward P. Nikonowicz*

¹Department of Biochemistry and Cell Biology, Rice University, Houston, TX 77251-1892 and ²Department of Microbiology and Center for RNA Biology, The Ohio State University, Columbus, OH 43210, USA

Received December 3, 2009; Revised January 7, 2010; Accepted January 8, 2010

ABSTRACT

Gram-positive bacteria utilize a tRNA-responsive transcription antitermination mechanism, designated the T box system, to regulate expression of many amino acid biosynthetic and aminoacyl-tRNA synthetase genes. The RNA transcripts of genes controlled by this mechanism contain 5' untranslated regions, or leader RNAs, that specifically bind cognate tRNA molecules through pairing of nucleotides in the tRNA anticodon loop with nucleotides in the Specifier Loop domain of the leader RNA. We have determined the solution structure of the Specifier Loop domain of the tyrS leader RNA from Bacillus subtilis. Fifty percent of the nucleotides in the Specifier Loop domain adopt a loop E motif. The Specifier Sequence nucleotides, which pair with the tRNA anticodon, stack with their Watson–Crick edges rotated toward the minor groove and exhibit only modest flexibility. We also show that a Specifier Loop domain mutation that impairs the function of the B. subtilis glyQS leader RNA from Bacillus subtilis. Fifty percent of the nucleotides in the Specifier Loop domain adopt a loop E motif. The Specifier Sequence nucleotides, which pair with the tRNA anticodon, stack with their Watson–Crick edges rotated toward the minor groove and exhibit only modest flexibility. We also show that a Specifier Loop domain mutation that impairs the function of the B. subtilis glyQS leader RNA from Bacillus subtilis. Fifty percent of the nucleotides in the Specifier Loop domain adopt a loop E motif. The Specifier Sequence nucleotides, which pair with the tRNA anticodon, stack with their Watson–Crick edges rotated toward the minor groove and exhibit only modest flexibility.

INTRODUCTION

The expression of many aminoacyl-tRNA synthetase genes and genes involved in amino acid metabolism and uptake in Gram-positive bacteria is regulated at the level of transcription attenuation by the T box riboswitch, an RNA element in the 5' untranslated (or leader) region of the gene (1,2). The expression of genes in this family is determined by the relative levels of charged and uncharged tRNAs in the cell, with each gene in the family responding individually to a specific tRNA species. If the corresponding tRNA is highly aminoacylated, a transcriptional attenuator in the leader region is active, and the gene is not expressed. Readthrough of the termination site is dependent on binding of uncharged tRNA to the T box element, which is 200–300 nt in length and contains several conserved primary and secondary structure elements, including three helices (Stems I, II and III), the Stem IIIA/B pseudoknot and the terminator and antiterminator structures (Figure 1A). The secondary structures of the terminator and antiterminator helices are mutually exclusive and the terminator helix is thermodynamically more favorable (3). Stabilization of the antiterminator helix, and subsequent transcriptional readthrough, is accomplished by pairing of the four 3' terminal nucleotides of the tRNA molecule with residues in a 7-nt bulge of the antiterminator helix (4).

Specific tRNA recognition is dependent primarily on the identity of three nucleotides, the Specifier Sequence, within the Specifier Loop domain, a conserved internal loop of variable size located in Stem I (3,5). The Specifier Sequence nucleotides are complementary to the anticodon nucleotides of the cognate tRNA. Changes in the Specifier Sequence can in some cases result in a switch in the specificity of the T box riboswitch to allow recognition of a new tRNA (3,5,6). The residue 3' to the Specifier Sequence is conserved as a purine, and pairs with the conserved U at position 33 in tRNA, 5' to the anticodon (7). The functional importance of these tRNA–leader RNA interactions have been demonstrated for the Bacillus subtilis tyrS and glyQS leader RNAs (3,8) and reproduced in vitro using a purified transcription assay and the glyQS leader RNA (9,7). In addition, interaction between glyQS Specifier Loop domain and tRNA Gly anticodon stem–loop was demonstrated with fluorescence assays (10).

In addition to the overall secondary structure and nucleotide conservation of T box leader RNAs (3), a
loop E (or S-turn) motif is predicted to form in the Specifier Loop (11). The loop E motif is a common RNA structural element found in many RNAs including ribosomal RNAs (12–14), the hairpin ribozyme (15), and the nucleolin-recognition element in eukaryotes (16,17). The sequences and structures of these motifs are very similar among different RNAs and feature a group of three non-canonical base pairs (18). This motif frequently is located proximal to multi-helix junctions and has crucial roles in protein-binding sites and in mediating RNA–RNA interactions (14). The high conservation of this motif within the Specifier Loop suggests an important role in leader RNA function, including tRNA binding.

The Specifier Loop domain of the *B. subtilis* tyrS leader RNA is an internal loop of 14 nt, including the 4 nt that pair with the tRNA\(^{3′}\) anticodon loop, flanked by two short helices. We have used heteronuclear nuclear magnetic resonance (NMR) spectroscopy to determine the solution structure and dynamics of this domain. Our results confirm the presence of the predicted loop E structural element within the internal loop, adjacent to the Specifier Sequence. We also show that a mutation within the loop E nucleotides that leads to loss of function in the context of the *B. subtilis* glyQS T box leader (Green,N.J., Grundy,F.J. and Henkin,T.M., unpublished data), destabilizes the motif, but does not significantly alter the conformation of the internal loop, confirming the importance of this motif for proper function.

**MATERIALS AND METHODS**

**Materials**

All enzymes were purchased from Sigma Chemical (St. Louis, MO, USA) except for T7 RNA polymerase, which was prepared as described (19). Deoxyribonuclease I type II, pyruvate kinase, adenylate kinase and nucleotide monophosphate kinase were obtained as powders, dissolved in 15% glycerol, 1 mM dithiothreitol and 10 mM Tris–HCl, pH 7.4, and stored at −20°C. Guanylate kinase and nuclease P1 were obtained as solutions and stored at −20°C. Unlabeled 5′-nucleoside triphosphates (5′-NTPs) were purchased from Sigma, phosphoenolpyruvate (potassium salt) was purchased from Bachem, and 99% [\(^{15}\)N]-ammonium sulfate and 99% [\(^{13}\)C]-glucose were purchased from Spectra Stable Isotopes (Branchburg, NJ, USA).

**Preparation of RNA samples**

The RNA sequence depicted in Figure 1B was prepared by *in vitro* transcription with T7 RNA polymerase using a synthetic DNA template (20) and either unlabeled
between corresponding proton doublets in HSQC spectra were determined from the measured frequency difference in D$_2$O. A 3D HCCH-TOCSY (56 ms DIPSI-3 spin lock) (16 ms and 24 ms DIPSI-3 spin lock) experiments collected spin systems were assigned using 3D HCCH-TOCSY of alignment was quantified using the quadrupole splitting of 19 mg/ml Pf1 and 0.35 mM RNA. The degree of alignment was determined using 13C/15N-labeled RNA for residual dipolar coupling measurements was achieved by adding RNA to concentrated Pf1 filamentous phage in 99.96% D$_2$O NMR buffer, yielding final concentrations of 1.5–2.5 mM). Partial alignment of 13C/15N-labeled RNA for residual dipolar coupling measurements was achieved by adding RNA to concentrated Pf1 filamentous phage in 99.96% D$_2$O NMR buffer, yielding final concentrations of 19 mg/ml Pf1 and 0.35 mM RNA. The degree of alignment was quantified using the quadrupole splitting of 19 mg/ml Pf1 and 0.35 mM RNA. The degree of alignment was quantified using the quadrupole splitting of 19 mg/ml Pf1 and 0.35 mM RNA. The degree of alignment was quantified using the quadrupole splitting of 19 mg/ml Pf1 and 0.35 mM RNA.

NMR spectroscopy

All spectra were acquired on a Varian Inova 500-MHz spectrometer equipped with a $^1$H–($^{13}$C, $^{15}$N, $^{31}$P) probe and Inova 600 and 800-MHz spectrometers equipped with cryogenically cooled $^1$H–($^{13}$C, $^{15}$N) probes. Solvent suppression for $^1$H homonuclear spectra collected in 90% H$_2$O was achieved using the WATERGATE scheme. Typically, the data points were extended by 25% using linear prediction for the indirectly detected dimensions. NMR spectra were processed and analyzed using Felix 2007 (Felix NMR Inc., San Diego, CA, USA).

Two-dimensional (2D) $^{13}$C–$^1$H HSQC spectra were collected to identify $^{13}$C–$^1$H chemical shift correlations. Sugar spin systems were assigned using 3D HCCH-TOCSY (16 ms and 24 ms DIPSI-3 spin lock) experiments collected in D$_2$O. A 3D HCCH-TOCSY (56 ms DIPSI-3 spin lock) was collected to establish the intra-base H2–C2–C8–H8 correlations in adenine residues. A 3D HCN experiment was collected to establish the intra-base H2–C2–C8–H8 correlations. 2D NOESY spectra (m = 120, 160 and 360 ms) and 3D $^{13}$C-edited NOESY spectra (m = 120 and 360 ms) were collected to identify purine N7 and adenine N1 and N3 resonances. For the exchangeable resonances, 2D $^{13}$N–$^1$H HSQC spectra were collected to identify $^{13}$N–$^1$H chemical shift correlations. 2D NOESY spectra (m = 180 and 400 ms) were acquired in H$_2$O and at 12°C to obtain distance restraints involving exchangeable protons.

$^1$H–$^{13}$C residual dipolar coupling constants (RDCs) were determined from the measured frequency difference between corresponding proton doublets in HSQC spectra acquired for isotropic and partially aligned samples. Thirty RDC values from base CH bond vectors and 16 from ribose 1' CH vectors were obtained in this manner. The axial and rhombic terms were determined within Xplor-NIH using an extensive grid search (24), and yielded values of $D_{\beta\gamma} = 31.67$ and $R_{\alpha\beta\gamma} = 0.19$.

$J_{\beta\gamma\gamma}$ coupling constants were estimated from DQF-COSY experiments. $J_{C-P}$ coupling constants were determined using the $^{13}$C–$^1$H ct-HSQC spin-echo difference method. $J_{P-H}$ couplings were estimated using $^{31}$P–$^1$H HetCor experiments.

$^{13}$C $T_1p$ relaxation times were measured using 2D $^{13}$C–$^1$H ct-HSQC-based experiments optimized for C2, for C1' and for C6 and C8 resonances. A 2.1-kHz $^{13}$C spin-lock field was used with delays of 5, 10, 15, 20, 30, 40, 50, 60, 70, 90 and 120 ms. The 5-ms experiment was collected twice to provide an estimate of the error of the measured intensities. The $^{13}$C–$^1$H cross-peak volumes were fit to a single exponential decay.

Distance and torsion angle constraints

Interproton distance estimates were obtained from cross-peak intensities in 2D NOESY and 3D $^{13}$C-edited NOESY spectra. Cross-peak intensities were calibrated using the pyrimidine H5–H6 fixed distance of 2.54 Å. NOE cross-peak intensities were classified into five categories assigned upper distance bounds of 3.0, 4.0, 5.0, 6.0 or 7.0 Å and a common lower bound of 1.8 Å. Base pairs were identified by direct detection of hydrogen bonds (25) or by observation of strong G=C NH–NH$_2$ or A=U H2–NH NOEs. Hydrogen bond distances restraints and planarity constraints were introduced for residues that form base pairs.

Ribose ring pucker and backbone dihedral constraints were derived from $J_{HHH}$, $J_{HP}$ and $J_{CP}$ couplings (26). Ribose rings with $J_{HH1-HH2}$ > 7 Hz and with C3' and C4' resonances between 76–80 and 85–86 ppm, respectively, were constrained to C2'-endo. Residues with $J_{HH1-HH2}$ < 5 Hz and couplings were constrained to C3'-endo. Residues with intermediate $J_{HH1-HH2}$ couplings were left unconstrained. For stem residues 1–6 and 33–38, γ was constrained to the gauche$^+$ conformation (60 ± 30°) (26); γ was left unconstrained for all other residues. For stem residues, β was constrained to the trans conformation (170 ± 40°) (26); β was loosely constrained to the trans conformation (160 ± 50°) for internal loop residues except G11 and A27 that were constrained loosely to gauche$^+$ (80 ± 50°) as determined from examination of loop E containing crystal structures (27,28). ε was constrained to exclude the gauche$^+$ conformation (−150 ± 50°) for residues with $J_{P-H1}$ > 5 Hz or $J_{P-C2}$ > 5 Hz. α and ζ were constrained to −70 ± 30° for the stem residues and were constrained to exclude the trans conformation (0 ± 120°) for residues 7–14 and 25–32 based on the absence of down-field shifted $^{31}$P resonances (29).

Structure refinement

Structure refinement was carried out with simulated annealing and restrained molecular dynamics (rMD) calculations were performed using Xplor-NIH v2.19 (24).
Starting coordinates for the tyrS<sub>SD</sub> model were generated using Insight II (Accelrys, San Diego, CA, USA) and were based on standard A-form helical geometry. The structure calculations were performed in two stages. Beginning with the energy-minimized starting coordinates, 100 structures were generated during the first round of structure calculation by 80 ps of rMD at 1200 K with hydrogen bond, NOE-derived distance and base-pairing restraints. The system then was cooled to 25 K in 47 cycles of rMD corresponding to a total of 12 ps. During this stage, RDC constraints and repulsive van der Waals forces were introduced into the system and the SANI force constant used for RDCs was gradually increased from 0.010 kcal mol<sup>−1</sup> Hz<sup>−2</sup> to 1.000 kcal mol<sup>−1</sup> Hz<sup>−2</sup>. Other force constants used for the calculations were increased—from 2 kcal mol<sup>−1</sup>A<sup>−2</sup> to 30 kcal mol<sup>−1</sup>A<sup>−2</sup> for the NOE and from 2 kcal mol<sup>−1</sup> rad<sup>−2</sup> to 30 kcal mol<sup>−1</sup> rad<sup>−2</sup> for the dihedral angle constraints. Once the temperature reached the target, each structure was then subjected to G<sup>33</sup> and the upper helix from G<sub>14</sub> to G<sub>21</sub>. The cytidine 15N-1H HSQC spectrum and the resonance linewidths are consistent with the predicted secondary structure, lending support to the prediction that the RNA is monomeric. The RNA molecule used in this study, tyrS<sub>SD</sub> (Figure 1B), corresponds to the Specifier Loop from Stem I of <i>B. subtilis</i> tyrS leader RNA. Cross-peaks in the NH<sup>15</sup>N-1H HSQC spectrum and the resonance linewidths are consistent with the predicted secondary structure, and the NH spectra at low concentration (5 μM) and high concentration (1.8 mM) are nearly identical, supporting the prediction that the RNA is monomeric. The chemical shifts of G<sub>14</sub>H<sub>1</sub> and residues G<sub>11</sub>, A<sub>13</sub>, G<sub>26</sub> and A<sub>28</sub> exhibit conformations intermediate between C2′-endo and C3′-endo.

The non-exchangeable 1H and 13C resonances of tyrS<sub>SD</sub> were assigned using standard heteronuclear techniques (32,33). Most of the base and ribose 1H-13C correlations are resolved, with four of the base resonances (A<sub>13</sub>, G<sub>14</sub> and G<sub>26</sub> H8-C8 and A<sub>13</sub> H2-C2) having spectral characteristics indicative of intermediate exchange (Figure 2A). All residues except U<sub>12</sub>, A<sub>13</sub> and G<sub>14</sub> yield base–ribose correlations in 3D HCN spectra, and 37 of the ribose spin systems were identified using 3D HCCH-TOCSY experiments (G<sub>1</sub> was only partially labeled).

Assignments for the non-exchangeable resonances were made using 2D NOESY (Figure 2B) and 3D 13C-edited NOESY experiments to identify sequential H6/8-H1′ NOE connectivities (32). The sequential H6/8-H1′ NOE connectivities are continuous in the 180-ms NOESY spectrum except at steps A<sub>10</sub>-G<sub>11</sub> and G<sub>11</sub>-U<sub>12</sub>. However, G<sub>11</sub>H<sub>8</sub>-U<sub>12</sub>H<sub>6</sub> and G<sub>11</sub>H<sub>4</sub>-U<sub>12</sub>H<sub>6</sub> cross-peaks are present in the spectrum. Interestingly, i to i+2 NOE cross-peaks between A<sub>10</sub>H1′/H2′ and U<sub>12</sub>H<sub>6</sub> are observed (Figure 2B) and suggest that the G<sub>11</sub> base bulges from the strand. The chemical shifts of G<sub>14</sub>H<sup>1</sup>′ (4.25 p.p.m.) and G<sub>11</sub>H<sub>4</sub>′ (5.85 p.p.m.) are unusual and are discussed below.

Most inter-nucleotide 31P resonances are dispersed between −3.4 and −5.1 p.p.m., but the G<sub>21</sub>pG<sub>22</sub> resonance has a chemical shift of −2.19 p.p.m. as previously noted for the UNCG tetraloop motif (34). Several 31P resonances could be assigned using the HCP experiment (35) or the H8/6-P and H1′-P correlations from 2D 31P-1H hetero-TOCSY-NOESY spectra (36). The chemical shifts for tyrS<sub>SD</sub> are listed in Supplementary Table S1.

**Structure of the tyrS<sub>SD</sub> molecule**

The structure of tyrS<sub>SD</sub> was calculated using a restrained molecular dynamics routine. The calculations used a total of 296 conformationally restrictive distance constraints, 189 dihedral angle constraints and 46 RDC constraints (Table 1) to produce 10 converged structures (Figure 3). The converged structures had an average of 5.8 distance constraint violations between 0.3 and 0.6 A. The heavy atoms of the converged structures superimpose on the average structure with an average root mean square deviation (RMSD) of 1.25 A. The local RMSDs for the loop (A<sub>7</sub>–A<sub>13</sub> and G<sub>26</sub>–A<sub>32</sub>) and stems are 0.54 A and 1.22 A, respectively.

The abundance of constraints for the internal loop and flanking stem nucleotides (residues C<sub>6</sub>-G<sub>14</sub> and C<sub>25</sub>-G<sub>33</sub>) defines the conformation of these nucleotides with good precision (0.60 A rmsd) (Figure 3B and Table 1). The ribose ring puckers of several residues in the internal loop adopt non-A-form, C′<sub>endo</sub>, conformations. Residues A<sub>9</sub>, A<sub>10</sub> and A<sub>32</sub> adopt C′<sub>endo</sub> ring puckers and residues G<sub>11</sub>, A<sub>13</sub>, G<sub>26</sub> and A<sub>32</sub> exhibit conformations intermediate between C2′<sub>endo</sub> and C3′<sub>endo</sub>.

**Chemical-shift assignments**

The sequence-specific resonance assignment of tyrS<sub>SD</sub> was accomplished using 1H-1H NOESY and 2D and 3D heteronuclear experiments. The NH resonances were assigned using the NOE connectivities between NH proton resonances of neighboring base pairs. These connectivities are continuous in the lower helix from G<sub>2</sub> to G<sub>33</sub> and the upper helix from G<sub>14</sub> to G<sub>21</sub>. The cytidine and adenosine NH<sub>2</sub> resonances were assigned using NOESY and HNCCH experiments (31). Additional weak NH resonances are present in the 1D 1H spectrum between 10.5 and 11.8 p.p.m., but cannot be uniquely assigned to U<sub>12</sub>, U<sub>19</sub>, U<sub>26</sub>, G<sub>1</sub>, G<sub>11</sub> or G<sub>26</sub>.
hairpin ribozyme (15,38). The core of this element typically contains three nonstandard base–base interactions, a sheared A–G pair, a U–A trans-Hoogsteen interaction and a parallel A–A base pair, that stack on one another (13). In some cases, a bulged G or C nucleotide is present between the non-canonical U–A and A–A interactions. The spectral data indicate that the loop E motif structure is present but adopts a more open conformation than is found in other contexts. A characteristic spectral signature of the sheared A–G pair is an upfield shift of the H1'H resonance of the residue 30 to the A (13,17,39). This H1'H resonance shift is observed for G14 and is consistent with a sheared A13–G26 orientation. The sheared A–G base pair forms two hydrogen bonds, AN6H2–GN3 and AN7–GN2H2 (Figure 4), and the A13 N7 and N6 15N chemical shifts provide additional support for the A13–G26 interaction. The trans-Hoogsteen U–A interaction adjacent to the sheared A–G pair also produces a very unusual spectral feature, an A10H8–U12H1' NOE cross-peak. The trans-Hoogsteen U–A arrangement rotates the uridine ribose ring so that the H1' resonance is repositioned proximal to the 5' (i-2) residue rather than the 3' (i + 1) residue. The U–A interaction involves hydrogen bonds between UO2 and AN6H2 and UN3H and AN7 (14). The U12N3H resonance is exchange-broadened and could not be observed, indicative of a weak hydrogen bond or solvent-accessible U12N3H. Adjacent to the trans-Hoogstein base pair is a parallel A–A interaction (Figure 4). The parallel A–A base pair is facilitated by an S-turn in the phosphate backbone and stabilized by two symmetric N6H2–N7 inter-base hydrogen bonds (14,28). The 15N chemical shifts are consistent with the A28N6H2–A10N7 hydrogen bond, but the chemical shifts do not support the second (symmetric) interaction (Supplementary Table S1; Figure 4). The G11 base is bulged and its conformation is supported by NOE data. Several i, i + 2 NOE cross-peaks connect A10 with U12 and no inter-residue NOE cross-peaks are observed between G11 and the A10 ribose. These NOE cross-peaks indicate stacking of the A10 and U12 bases. Efforts to confirm the predicted hydrogen bonds via through-bond correlation experiments (25) were unsuccessful. This result is consistent with the 15N chemical-shift data that suggest weak hydrogen bonds.

A hallmark of the loop E motif is the turn in the phosphate backbone 5' to the bulged nucleotide (between A10 and G11) (28). This turn permits the parallel A–A interaction without the adenine residues adopting the syn orientation about the glycosidic bond. The turn itself is accommodated by the flipping of the ribose moiety of one of the adenine residues relative to other ribose groups along the phosphate backbone. This ribose
Table 1. Summary of experimental distance and dihedral angle constraints and refinement statistics for tyrS\textsubscript{SD}

| Constraint                               | tyrS\textsubscript{SD} |
|------------------------------------------|-------------------------|
| NOE distance constraints                 |                         |
| Intra-residue\textsuperscript{c}         | 87                      |
| Inter-residue                            | 135                     |
| Mean number per residue                  | 16                      |
| NOE constraints by category              |                         |
| Very strong (1.8–3.0 Å)                  | 34                      |
| Strong (1.8–4.0 Å)                       | 180                     |
| Medium (1.8–5.0 Å)                       | 215                     |
| Weak (1.8–6.0 Å)                         | 117                     |
| Very weak (1.8–7.0 Å)                    | 24                      |
| Base pair constraints                    |                         |
| Total                                    | 32                      |
| Dihedral angle constraints               |                         |
| Ribose ring\textsuperscript{b}           | 104                     |
| Backbone                                 | 189                     |
| Mean number per residue                  | 7.7                     |
| Residual dipolar coupling constraints    |                         |
| Base CH                                  | 30                      |
| Ribose 1’ CH                            | 16                      |
| Violations                               |                         |
| Average distance constraints > 0.5 Å\textsuperscript{c} | 5.8                     |
| Average dihedral constraints > 0.5 Å\textsuperscript{d} | 1.6                     |
| RMSD from ideal geometry\textsuperscript{f} |                  |
| Heavy atoms (Å)                          | 1.24                    |
| Backbone atoms (Å)                       | 1.27                    |

\textsuperscript{a}Only conformationally restrictive constraints are included.
\textsuperscript{b}Three torsion angles within each ribose ring were used to constrain the ring to either the C2-endo or C3-endo conformation; the ring pucker of residues G\textsubscript{11}, A\textsubscript{13}, G\textsubscript{26} and A\textsubscript{28} were not constrained.
\textsuperscript{c}A distance violation of 0.5 Å corresponds to 5.0 kcal energy penalty.
\textsuperscript{d}A dihedral angle violation of 0.5° corresponds to 0.05 kcal energy penalty.
\textsuperscript{e}Calculated against the minimized average structure.

structure and participating directly in catalysis. Mg\textsuperscript{2+} is crucial for the correct folding of glyQ5 leader RNA (7), but the specific structural changes and exact sites of coordination associated with Mg\textsuperscript{2+} binding are not known.

The effects of Mg\textsuperscript{2+} on different RNAs that contain loop E motifs are variable. NMR spectra show that the loop E motif in eukaryotic 5S rRNA is stabilized by Mg\textsuperscript{2+} (40), but spectra of the same motif sequence located in the sNRE and the hairpin ribozyme are unaffected by Mg\textsuperscript{2+} (41–43). In addition, no metal ions are associated with the loop E motif in the crystal structure of the sarcin/ricin loop of 5S rRNA (44).

The NH and NH\textsubscript{2} chemical-shift perturbations by Mg\textsuperscript{2+} are limited to the G\textsubscript{2}–U\textsubscript{37} NH resonances in tyrS\textsubscript{SD}. However, the chemical shifts of several CH base resonances are altered by up to 0.15 p.p.m. (Supplementary Figure S1). These include the base C8–H8 and C6–H6 resonances of loop residues G\textsubscript{32}, C\textsubscript{26}, A\textsubscript{27}, A\textsubscript{9} and C\textsubscript{31} and adenine C2–H2 resonances A\textsubscript{13}, A\textsubscript{7}, A\textsubscript{9} and A\textsubscript{32}. The intensity of the A\textsubscript{13} H8–C8 resonance increases, but the G\textsubscript{11} H8–C8 resonance broadens. Significantly, the pattern and intensities of NOE cross-peaks of the Mg\textsuperscript{2+}-free sample are largely preserved upon addition of Mg\textsuperscript{2+}, indicating little or no structural impact.

\textsuperscript{13}C Relaxation measurements

The reorientation of a \textsuperscript{13}C–\textsuperscript{1}H bond vector on the picosecond timescale can be assessed through its carbon T\textsubscript{1p} relaxation: the longer the relaxation time, the more mobile the \textsuperscript{13}C–\textsuperscript{1}H pair. The T\textsubscript{1p} relaxation times for the base C6 and C8 and ribose C1’ positions of tyrS\textsubscript{SD} were measured. Cross-peak overlap and chemical exchange prevented accurate measurement of a few adenine C2 nuclei, pyrimidine C6 and several C1’ nuclei. The majority of base resonances from internal loop residues exhibited relaxation times between 40 and 50 ms, which are comparable to stem region nucleotides. However, the T\textsubscript{1p} values for nucleotides A\textsubscript{10}–G\textsubscript{26}, A\textsubscript{9} and A\textsubscript{28} within the loop E motif and the proximal stem nucleotides G\textsubscript{32} and C\textsubscript{25} were increased by 20%, indicative of increased mobility. Chemical exchange is dominant among the base resonances of G\textsubscript{11}, U\textsubscript{12} and A\textsubscript{13} and prevents accurate T\textsubscript{1p} measurements for these residues. A similar pattern is not exhibited by the C1’ nuclei. Only the A\textsubscript{9}, A\textsubscript{10}, A\textsubscript{28} and A\textsubscript{30} residues from the internal loop have relaxation times 30–50% longer than other residues in the internal loop and the stems. The increased relaxation times of nucleotide base resonances and exchange broadening for residues in the region of the loop E motif are consistent with a structural element that is moderately rigid. The addition of Mg\textsuperscript{2+} did not significantly alter the relaxation profile of the tyrS\textsubscript{SD} RNA.

**DISCUSSION**

The T-box mechanism for regulation of amino acid-related genes is widely utilized in Gram-positive bacteria. Extensive mutagenesis studies of the leader RNA sequence have defined sequence requirements for tRNA binding in the 5’ and 3’ regions of the leader RNA (4–6,11,45).
Residues within the Specifier Loop domain of the leader RNA bind directly to the anticodon nucleotides of cognate tRNA molecules and confer specificity to the leader RNA–tRNA interaction. The Specifier Loop domain also contains conserved nucleotides that correspond to the loop E RNA secondary structure motif.

Structure of conserved residues in the tyrS leader RNA Specifier Loop domain

The 14 nt corresponding to the Specifier Loop domain of the B. subtilis tyrS leader RNA are well ordered and only moderately dynamic. The results show that the nucleotides in the upper half of the domain adopt a loop E motif (Figure 3). This structural element includes an S-turn of the phosphate backbone at residue A10 resulting in an inverted orientation of the A10 ribose. Although well ordered, the loop E element in the Specifier Loop domain has a more open and possibly more flexible fold than loop E motifs in other contexts (13,17,44). The hydrogen-bonding pattern among the loop E motif bases could be inferred only from $^{15}$N chemical shifts and could not be directly confirmed through hydrogen bond-mediated inter-base correlations. These results point to a fold that is open and may be partially accessible to solvent (Figure 5). Flexibility on the intermediate timescale in this region is reflected in the exchange-broadened base resonances of residues G11–A13 and the relaxation data support the modest dynamic nature. The limited mobility of these residues may allow for an organized target to be presented to the incoming tRNA anticodon but still allow sufficient flexibility to optimize Specifier Sequence nucleotide–anticodon pairing.
Nucleotides at the 3’-end of the bulge are highly conserved and well ordered, and are stabilized primarily by base–base stacking. While the flexibility associated with the invariant residues is an important component of their tRNA binding function, the more rigid 3’-end of the bulge may limit the conformations sampled by the loop nucleotides and facilitate tRNA binding (46). The significant dynamics among the tRNA-pairing nucleotides in the antiterminator hairpin is not shared by the Specifier Sequence nucleotides. Although Specifier Sequence–tRNA pairing requires additional rotation of the Specifier Sequence bases away from the helix axis, these residues are well ordered and appear to be stabilized by stacking. The ribose groups of these residues also exhibit limited dynamics and adopt the regular A-form 3’-endo conformation. The three residues of the Specifier Loop domain that are dynamic (G_{11}–A_{13}) may provide flexibility to optimize tRNA binding.

**Role of the loop E motif in tRNA binding**

The loop E motif is a common RNA structural element that is well represented in ribosomal RNAs (13,37,47–50), loop B of the hairpin ribozyme (15,51), the specificity domains of some type B RNase P molecules (52,53), the internal ribosome entry site (IRES) element found in the 5′ untranslated region of hepatitis C virus (HCV) RNA (54,55), the central domain of potato spindle tuber viroid (PSTV) (37), and in in vitro selected nucleolin recognition element (sNRE) mutants (17). Loop E nucleotides participate directly in protein–RNA interactions in most of the rRNA molecules and in the sNRE RNA (13,16,17,47,49,56) and, in the hairpin ribozyme, the motif forms part of a corridor for active site water migration (57). In the case of the loop E structures found in domains IIb and IIIId of the IRES element of HCV RNA, it is still not clear whether the loop E structure provides a platform for protein or RNA binding (54,55).

The functional importance of the loop E motif in the Specifier Loop domain is suggested by nucleotide conservation and mutagenesis (11, Green,N.J., Grundy,F.J. and Henkin,T.M., unpublished data). Comparison of the Specifier Loop domains of several T-box leader RNAs across multiple bacterial species shows the conservation of two elements, the Specifier Sequence nucleotides at the 3′-end of the 3′ strand followed by an unpaired adenine nucleotide and a group of 7 nt in the upper part of the internal loop that correspond to the loop E motif (11,58). The number and identities of the remaining nucleotides in the Specifier Loop domain are variable, but nucleotide changes within the loop E motif can be deleterious. For example, an uridine-to-cytidine mutation in the Specifier Loop domain of the glyQS leader RNA at a position corresponding to residue 12 in the tyrS_{SD} RNA leads to loss of function in in vivo and in vitro transcriptional attenuation assays (Green,N.J., Grundy,F.J. and Henkin,T.M., unpublished data). Figure 6 shows that resonances primarily affected by the U-to-C mutation correspond to nucleotides that comprise the loop E motif and indicate that the structure of the motif has been disrupted. Interestingly, the structural

---

**Figure 5.** Stereoview of the minimized average structure of the Specifier Loop domain. The specifier nucleotide bases are colored green and the loop E motif nucleotide bases are colored red. The functional groups on the Watson–Crick edges of the specifier nucleotides are colored pink. The S-turn of the sugar-phosphate backbone can be seen between residues A_{32} and G_{11}.
perturbations appear limited to the loop E motif. The chemical shifts corresponding to the residues of the Specifier Loop domain outside of the loop E motif are not significantly affected by the mutation, suggesting minimal perturbation of their local structure.

The structure of tyrSSD and the effects caused by the U-to-C mutation suggest a model for tRNA binding to the Specifier Sequence on the minor groove side of the Specifier Loop domain (Figure 7). The Specifier Sequence bases are well ordered in the solution structure and are only moderately more dynamic than the helix residues, suggesting that they are primed for pairing with the tRNA anticodon loop. In order for the Specifier Sequence nucleotides to pair with tRNA, the bases must rotate out toward the major or minor groove to present their Watson–Crick edges. Rotation toward the major groove would result in under-winding of the sugar-phosphate backbone, whereas rotation toward the minor groove would entail over-winding of the backbone. In the structure of tyrSSD, the 4 Specifier Sequence bases turn toward the minor groove, whereas the 3 bases of the partner strand turn out toward the major groove.

We propose that the S-turn of the loop E motif is responsible for this base positioning and that it acts through the phosphate backbone. The S-turn, centered on A10, causes a differential shortening of the distance between the lower helix and the A10–A28 base pair of the loop. To accommodate the constraints on the phosphate backbone imposed by the S-turn and the asymmetric number of residues in this region, bases on the partner strand rotate toward the major groove and the Specifier Sequence bases rotate toward the minor groove. Although the U-to-C mutation does not appear to affect the local structure around the Specifier Sequence nucleotides, their propensity to rotate toward the minor groove would be impaired by loss of the S-turn at A10. Thus, in this model for tRNA binding, the loop E motif positions a specific structural component, the S-turn, that kinks the sugar-phosphate backbone of the partner strand and promotes rearrangement of the Specifier Sequence bases toward the minor groove for pairing with the tRNA (Figure 7). The loop E motif may additionally provide a platform to help stabilize the duplex formed by the Specifier Sequence and anticodon base pairs through stacking interactions.

**ACCESSION NUMBER**

Atomic coordinates for the refined structures have been deposited with the Protein Data Bank under accession code 2KHY.
SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Malgorzata Michnicka for preparation of the T7 RNA polymerase and synthesis of the labeled 5'-nucleotide triphosphates. The 800 MHz NMR spectrometer was purchased with funds from the W. M. Keck Foundation and the John S. Dunn Foundation.

CONFlict of interest statement. None declared.

REFERENCES

1. Henkin,T.M. (2008) Riboswitch RNAs: using RNA to sense cellular metabolism. Genes Dev., 22, 3383–3390.
2. Henkin,T.M. and Grundy,F.J. (2006) Sensing metabolic signals with nascent RNA transcripts: the T box and S box riboswitches as paradigms. Cold Spring Harbor Symp. Quant. Biol., 71, 231–237.
3. Grundy,F.J. and Henkin,T.M. (1993) tRNA as a positive regulator of transcription antitermination in B. subtilis. Cell, 74, 475–482.
4. Grundy,F.J., Rollins,S.M. and Henkin,T.M. (1994) Interaction between the acceptor end of tRNA and the T box stimulates antitermination in the Bacillus subtilis tyrS gene: a new role for the discriminator base. J. Bacteriol., 176, 4518–4526.
5. Grundy,F.J., Hodil,S.E., Rollins,S.M. and Henkin,T.M. (1997) Specificity of RNA-mRNA interactions in Bacillus subtilis tyrS antitermination. J. Bacteriol., 179, 2587–2594.
6. Grundy,F.J., Moir,T.R., Haldeman,M.T. and Henkin,T.M. (2002a) Sequence requirements for terminators and antiterminators in the T box transcription antitermination system: disparity between conservation and functional requirements. Nucleic Acids Res., 30, 1646–1655.
7. Yousef,M.R., Grundy,F.J. and Henkin,T.M. (2005) Structural transitions induced by the interaction between tRNA(Pro) and the Bacillus subtilis glyQS T box leader RNA. J. Mol. Biol., 349, 273–287.
8. Grundy,F.J. and Henkin,T.M. (1994) Conservation of a transcription antitermination mechanism in aminoaoyl-tRNA synthetase and amino acid biosynthesis genes in Gram-positive bacteria. J. Mol. Biol., 235, 798–804.
9. Grundy,F.J., Winkler,W.C. and Henkin,T.M. (2002b) tRNA-mediated transcription antitermination in vitro: codon-anticodon pairing independent of the ribosome. Proc. Natl Acad. Sci. USA, 99, 11121–11126.
10. Nelson,A.R., Henkin,T.M. and Agris,P.F. (2006) tRNA regulation of gene expression: interactions of an mRNA 5′-UTR with a regulatory tRNA. RNA, 12, 1254–1261.
11. Rollins,S.M., Grundy,F.J. and Henkin,T.M. (1997) Analysis of cis-acting sequence and structural elements required for antitermination of the Bacillus subtilis tyrS gene. Mol. Microbiol., 25, 411–421.
12. Szewczak,A.A. and Moore,P.B. (1995) The sarcin/ricin loop, a modular RNA. J. Mol. Biol., 247, 81–98.
13. Dallas,A. and Moore,P.B. (1997) The loop E-loop D region of Escherichia coli 5S rRNA: the solution structure reveals an unusual loop that may be important for binding ribosomal proteins. Structure, 5, 1639–1653.
14. Leontis,N.B. and Westhof,E. (1998) A common motif organizes the structure of multi-helix loops in 16S and 23S ribosomal RNAs. J. Mol. Biol., 283, 571–583.
15. Butcher,S.E. and Burke,J.M. (1994a) A photo-cross-linkable tertiary structure motif found in functionally distinct RNA molecules is essential for catalytic function of the hairpin ribozyme. Biochemistry, 33, 992–999.
16. Bouvet,P., Allain,F.H., Finger,L.D., Dieckmann,T. and Feigon,J. (2001) Recognition of pre-formed and flexible elements of an RNA stem-loop by nucleolin. J. Mol. Biol., 309, 763–775.
17. Johansson,C., Finger,L.D., Trantirek,L., Mueller,T.D., Kim,S., Laird-Offringa,I.A. and Feigon,J. (2004) Solution structure of the complex formed by the two N-terminal RNA-binding domains of nucleolin and a pre-rRNA target. J. Mol. Biol., 337, 799–816.
18. Brunel,C., Romby,P., Westhof,E., Ehresmann,C. and Ehresmann,B. (1991) Three-dimensional model of Escherichia coli ribosomal 5S RNA as deduced from structure probing in solution and computer modeling. J. Mol. Biol., 221, 293–308.
19. Davanloo,P., Rosenbaum,A.H., Dunn,J.J. and Studier,F.W. (1984) Cloning and expression of the gene for bacteriophage T7 RNA polymerase. Proc. Natl Acad. Sci. USA, 81, 2035–2039.
20. Milligan,J.F., Groebe,D.R., Witherrell,G.W. and Uitenbeck,O.C. (1987) Oligonucleotidyl synthesis using T7 RNA polymerase, and synthetic DNA templates. Nucleic Acids Res., 15, 8783–8798.
21. Nikonowicz,E.P., Srria,A., Legault,P., Jucker,F.M., Baer,L.M. and Pardi,A. (1992) Preparation of 13C and 15N labelled RNAs for heteronuclear multi-dimensional NMR studies. Nucleic Acids Res., 20, 4507–4513.
22. Hansen,M.R., Mueller,L. and Pardi,A. (1998) Tunable alignment of macromolecules by filamentous phage yields dipolar coupling interactions. Nat. Struct. Biol., 5, 1065–1074.
23. Marino,J.P., Diener,J.L., Moore,P.B. and Griesinger,C. (1997) Multiple-quantum Coherence dramatically enhances the sensitivity of CH and CH2 correlations in uniformly 13C-labeled RNA. J. Am. Chem. Soc., 119, 7361–7366.
24. Schwiers,C.D., Kuszewski,J.J., Tjandra,N. and Clore,G.M. (2003) The Xplor-NIH NMR molecular structure determination package. J. Magn. Reson., 160, 65–73.
25. Dingley,A.J. and Grzesiek,S. (1998) Direct observation of hydrogen bonds in nucleic acid base pairs by internucleotide 1H-1H couplings. J. Am. Chem. Soc., 120, 8293–8297.
26. Varani,G., Aboutoul-F. and Allain,F.H.T. (1996) NMR investigation of RNA structure. Prog. Nucle. Magn. Reson. Spec., 29, 51–127.
27. Szele,S., Wang,J. and Moore,P.B. (2003) The crystal structure of a 26-nucleotide RNA containing a hook-turn. RNA, 9, 44–51.
28. Correll,C.C., Freeborn,B., Moore,P.B. and Steitz,T.A. (1997) Metals, motifs, and recognition in the crystal structure of a 5S rRNA domain. Cell, 91, 705–712.
29. Gorenstein,D.G. (1984) Phosphorus-31 NMR: Principles and Applications. Academic Press, New York.
30. Clore,G.M. and Kuszewski,J. (2003) Improving the accuracy of NMR structures of RNA by means of conformational database potentials of mean force as assessed by complete dipolar coupling cross-validation. J. Am. Chem. Soc., 125, 1518–1525.
31. Simorre,J.-P., Zimmermann,G.R., Mueller,L. and Pardi,A. (1996) Triple-resonance experiments for assignment of adenosine base resonances in 13C/15N-labeled RNA. J. Am. Chem. Soc., 118, 5316–5317.
32. Pardi,A. (1995) Multidimensional heteronuclear NMR experiments for structure determination of isotopically labeled RNA. Methods Enzymol., 261, 350–380.
33. Dieckmann,T. and Feigon,J. (1997) Assignment methodology for larger RNA oligonucleotides: application to an ATP-binding RNA aptamer. J. Biol. Chem., 272, 259–272.
34. Allain,F.H.-T. and Varani,G. (1995) Divalent metal ion binding to a conserved wobble pair defining the upstream site of cleavage of group I self-splicing introns. Nucleic Acids Res., 23, 341–350.
35. Heus,H.A., Wijmenga,S.S., van de Ven,F.J.M. and Hilbers,C.W. (1994) Sequential backbone assignment in 13C/15N-labeled RNA via phosphorus-31 NMR heteronuclear multi-dimensional TOCSY. J. Am. Chem. Soc., 116, 4983–4984.
36. Kellogg, G.W. and Schweitzer, B.I. (1993) Two- and three-dimensional P-driven NMR procedures for complete assignment of backbone resonances in oligodeoxyribonucleotides. J. Biomol. NMR, 3, 577–595.
37. Branch, A.D., Benenfeld, B.J. and Robertson, H.D. (1985) Ultraviolet light-induced crosslinking reveals a unique region of local tertiary structure in potato spindle tuber viroid and HeLa SS RNA. Proc. Natl Acad. Sci. USA, 82, 6590–6594.
38. Butcher, S.E. and Burke, J.M. (1994b) Structure-mapping of the hairpin ribozyme. Magnesium-dependent folding and evidence for tertiary interactions within the ribozyme-substrate complex. J. Mol. Biol., 244, 52–63.
39. Tolbert, B.S., Kennedy, S.D., Schroeder, S.J., Krugh, T.R. and Turner, D.H. (2007) NMR structures of (rGCGAGGCUCG)2 and (rGCGGAUGCU)2: probing the structural features that shape the thermodynamic stability of GA pairs. Biochemistry, 46, 1511–1522.
40. Kellogg, G.W. and Schweitzer, B.I. (1993) Two- and three-dimensional P-driven NMR procedures for complete assignment of backbone resonances in oligodeoxyribonucleotides. J. Biomol. NMR, 3, 577–595.
37. Branch, A.D., Benenfeld, B.J. and Robertson, H.D. (1985) Ultraviolet light-induced crosslinking reveals a unique region of local tertiary structure in potato spindle tuber viroid and HeLa SS RNA. Proc. Natl Acad. Sci. USA, 82, 6590–6594.
38. Butcher, S.E. and Burke, J.M. (1994b) Structure-mapping of the hairpin ribozyme. Magnesium-dependent folding and evidence for tertiary interactions within the ribozyme-substrate complex. J. Mol. Biol., 244, 52–63.
39. Tolbert, B.S., Kennedy, S.D., Schroeder, S.J., Krugh, T.R. and Turner, D.H. (2007) NMR structures of (rGCGAGGCUCG)2 and (rGCGGAUGCU)2: probing the structural features that shape the thermodynamic stability of GA pairs. Biochemistry, 46, 1511–1522.
40. Serra, M.J., Baird, J.D., Dale, T., Fey, B.L., Retatagos, K. and Westhof, E. (2002) Effects of magnesium ions on the stabilization of RNA oligomers of defined structures. RNA, 8, 307–323.
41. Butcher, S.E., Allain, F.H. and Feigon, J. (2000) Determination of metal ion binding sites within the hairpin ribozyme domains by NMR. Biochemistry, 39, 2174–2182.
42. Rupert, P.B. and Ferre-D’Amare, A.R. (2001) Crystal structure of a hairpin ribozyme-inhibitor complex with implications for catalysis. Nature, 410, 783–786.
43. Finger, L.D., Trantirek, L., Johansson, C. and Feigon, J. (2003) Solution structures of stem-loop RNAs that bind to the two N-terminal RNA-binding domains of nucleolin. Nucleic Acids Res., 31, 4641–4672.
44. Correll, C.C., Wool, I.G. and Munishkin, A. (1999) The two faces of the Escherichia coli 23S rRNA sarcin/ricin domain: the structure at 1.11 Å resolution. J. Mol. Biol., 292, 275–287.
45. Grundy, F.J., Collins, J.A., Rollins, S.M. and Henkin, T.M. (2000) tRNA determinants for transcription antitermination of the Bacillus subtilis tyrS gene. RNA, 6, 1131–1141.
46. Gerdes, M.S., Henkin, T.M. and Hines, J.V. (2003) Solution structure of the Bacillus subtilis T-box antiterminator RNA: seven nucleotide bulge characterized by stacking and flexibility. J. Mol. Biol., 326, 189–201.
47. Ban, N., Nissen, P., Hansen, J., Moore, P.B. and Steitz, T.A. (2000) The complete atomic structure of the large ribosomal subunit at 2.4 Å resolution. Science, 289, 905–920.
48. Endo, Y., Gluck, A. and Wool, I.G. (1993) Ribosomal RNA identity elements for recognition by ricin and by alpha-sarcin: mutation in the putative CG pair that closes a GAGA tetraloop. Nucleic Acids Symp. Ser., 29, 165–166.
49. Schuwirth, B.S., Borovinskaya, M.A., Hau, C.W., Zhang, W., Vila-Sanjurjo, A., Holton, J.M. and Cate, J.H. (2005) Structures of the bacterial ribosomes at 3.5 Å resolution. Science, 310, 827–834.
50. Wimberly, B., Varani, G. and Tinoco, I. Jr (1993) The conformation of loop E of eukaryotic 5S ribosomal RNA. Biochemistry, 32, 1078–1087.
51. Butcher, S.E., Allain, F.H. and Feigon, J. (1999) Solution structure of the loop B domain from the hairpin ribozyme. Nat. Struct. Biol., 6, 212–216.
52. Krashnikov, A.S., Yang, X., Pan, T. and Mondragon, A. (2003) Crystal structure of the specificity domain of ribonuclease P. Nature, 421, 760–764.
53. Massire, C., Jaeger, L. and Westhof, E. (1998) Derivation of the three-dimensional architecture of bacterial ribonuclease P RNAs from comparative sequence analysis. J. Mol. Biol., 279, 773–793.
54. Klinck, R., Westhof, E., Walker, S., Afshar, M., Collier, A. and Abo-Ela, F. (2000) A potential RNA drug target in the hepatitis C virus internal ribosomal entry site. RNA, 6, 1423–1431.
55. Lukavsky, P.J., Kim, J., Otto, G.A. and Puglisi, J.D. (2003) Structure of HCV IRES domain II determined by NMR. Nat. Struct. Biol., 10, 1033–1038.
56. Selmer, M., Dunham, C.M., Murphy, F.V. t., Weixlbaumer, A., Petry, S., Kelley, A.C., Weir, J.R. and Ramakrishnan, V. (2006) Structure of the 70S ribosome complexed with tRNA and rRNA. Science, 313, 1935–1942.
57. Salter, J., Krucinska, J., Alam, S., Grum-Tokars, V. and Wedekind, J.E. (2006) Water in the active site of an all-RNA hairpin ribozyme and effects of Gua8 base variants on the geometry of phosphoryl transfer. Biochemistry, 45, 686–700.
58. Gutierrez-Preciado, A., Henkin, T.M., Grundy, F.J., Yanofsky, C. and Merino, E. (2009) Biochemical features and functional implications of the RNA-based T box regulatory mechanism. Microbiol. Mol. Biol. Rev., 36, 33–61.
59. Varani, G., Cheong, C. and Tinoco, I. Jr (1991) Structure of an unusually stable RNA hairpin. Biochemistry, 30, 3280–3289.