Published in final edited form as:

Microbiol Spectr. 2018 July ; 6(4): . doi:10.1128/microbiolspec.RWR-0028-2018.

The T box riboswitch: tRNA as an effector to modulate gene regulation

Kiel D. Kreuzer^{a,b} and Tina M. Henkin^{a,*}

^aDepartment of Microbiology and Center for RNA Biology, The Ohio State University, Columbus, OH 43210

^bMolecular, Cellular and Developmental Biology Graduate Program, The Ohio State University, Columbus, OH 43210

Abstract

The T box riboswitch is a unique RNA-based regulatory mechanism that modulates expression of a wide variety of amino acid-related genes, predominantly in Firmicutes. RNAs of this class selectively bind a specific cognate tRNA, utilizing recognition of the tRNA anticodon and other tRNA features. The riboswitch monitors the aminoacylation status of the tRNA to induce expression of the regulated downstream gene(s) at the level of transcription antitermination or derepression of translation initiation in response to reduced tRNA charging via stabilization of an antiterminator or antisequestrator. Recent biochemical and structural studies have revealed new features of tRNA recognition that extend beyond the initially identified Watson-Crick base-pairing of a codon-like sequence in the riboswitch with the tRNA anticodon, and residues in the antiterminator or antisequestrator with the tRNA acceptor end. These studies have revealed new tRNA contacts, and new modes of riboswitch function and ligand recognition that expand our understanding of RNA-RNA recognition and the biological roles of tRNA.

Keywords

RNA-mediated regulation; transcription attenuation; translational regulation; RNA structure; tRNA; gene expression

INTRODUCTION

Bacteria have evolved a wide array of mechanisms to control gene expression in response to environmental changes. These regulatory mechanisms ensure that specific genes are expressed under the appropriate physiological conditions, and they regulate every step of expression from transcription initiation to post-translational modification and protein stability. The discovery of the T box mechanism revealed that an uncharged tRNA can interact with an mRNA to regulate expression of the downstream coding region (1) (Figure 1). This mechanism was the first of many regulatory systems in which *cz'*s-encoded RNA responds directly to a physiological signal to control gene expression through structural

*Corresponding author contact information: Department of Microbiology, The Ohio State University, 484 W. 12th Avenue, Columbus, OH 43210, 614-688-3831, henkin.3@osu.edu.

rearrangements. Regulatory RNAs of this type, termed riboswitches, have become an intense focus of research, and to date dozens of riboswitch classes that respond to various signals have been identified and characterized, including those that respond to temperature, pH, and metabolites such as enzyme cofactors, amino acids, and nucleotides (Chapter by Lotz and Suess).

The T box riboswitch is a unique class of regulatory RNA in part because its tRNA ligand is itself a complex structured RNA. All non-mitochondrial tRNAs share a similar L-shaped tertiary structure, but they vary considerably in sequence, stability, flexibility, and post-transcriptional modifications both across species and within the same organism (2). Unlike metabolite-binding riboswitches, which generally fold to form a compact binding pocket that recognizes specific chemical functional groups, T box RNAs contact their cognate tRNA ligand at several distant sites through different RNA-RNA interactions. Binding specificity and affinity for cognate tRNA, and discrimination against non-cognate tRNAs, are major challenges for T box riboswitches that mirror those for aminoacyl-tRNA synthetases (aaRSs). Significant progress has been made toward understanding the biochemical and structural features of T box regulation and tRNA binding, yet many aspects are not fully understood. Here, the current knowledge about the T box riboswitch is reviewed with an emphasis on recent developments and discoveries.

DISCOVERY AND MECHANISM

The T box regulatory system was first identified by recognition of a conserved 14 nt sequence in the upstream region of several aaRS genes in *Bacillus subtilis* (3). This sequence precedes an intrinsic transcription terminator helix in the 5' untranslated region, or leader region, of these genes, which suggested a common regulatory mechanism involving transcription attenuation (Chapter by Winkler). Mutational analysis of this conserved sequence, designated the T box, revealed that it was required for expression of the downstream genes. Intriguingly, expression of each gene could be induced only by limitation for its cognate amino acid, and not by general amino acid starvation. The key discovery in understanding the specificity of this regulatory mechanism was identification of a three nt codon-like sequence that corresponds to the amino acid identity of the downstream gene (1) (Figure 2A). This element was termed the Specifier Sequence, and mutation of the UAC sequence in a *tyrS-lacZ* transcriptional fusion to match the UUC phenylalanine sequence was sufficient to promote expression in response to limitation for phenylalanine instead of tyrosine. This result provided the first evidence that tRNA was likely to be the effector in the T box mechanism, and the tRNA anticodon was proposed to base pair with the Specifier Sequence (Figure 2B). Mutation of the UAC sequence to amber (UAG) and ochre (UAA) nonsense codons resulted in expression dependent on a suppressor tRNA, which further supported the crucial role for tRNA as the effector (4). To test if tRNA interacts with the leader RNA in the absence of the ribosome, a single nucleotide was inserted immediately upstream of the Specifier Sequence, which would cause a frameshift in any potential open reading frame. This mutation did not significantly affect *tyrS-lacZ* expression or induction, indicating that tRNA promotes antitermination independent of Specifier Sequence translation (1).

The T box sequence was predicted to form the 5' side of an antiterminator helix, comprised of two short helices flanking a 7 nt bulge, that competes with formation of the more thermodynamically stable terminator helix. The antiterminator element was predicted to be stabilized when the bulge nucleotides interact with the acceptor arm of uncharged tRNA, thereby preventing formation of the terminator helix (4) (Figure 1). Interaction of charged tRNA promotes termination indirectly as it is unable to stabilize the antiterminator due to the amino acid at the 3' acceptor end but can compete with uncharged tRNA for binding to the Specifier Sequence (5). The T box system therefore can monitor the charging ratio of a tRNA, such that only uncharged tRNA can stabilize the antiterminator and promote expression. Subsequent studies demonstrated tRNA-dependent antitermination in purified in vitro systems for *B. subtilis glyQS* and *Clostridium acetobutylicum alaS*, which revealed that regulation of these genes can occur without additional cellular factors (6, 7).

Whereas most T box riboswitches regulate expression at the level of premature transcription termination, a subclass regulates at the level of translation initiation (8). Sequence analysis indicated that the T box leader regions in some Actinobacteria do not contain a terminator helix; instead, the leader RNA was predicted to form a sequestrator helix that sequesters the Shine-Dalgarno (SD) sequence to prevent translation initiation (9). In this regulatory model, the acceptor arm of uncharged tRNA binds and stabilizes an anti sequestrator element, which prevents formation of the competing sequestrator helix and therefore promotes translation. Biochemical studies of the *Nocardia farcinica ileS* leader RNA provide evidence for specific tRNA binding and tRNA-dependent structural changes that are consistent with the proposed regulatory model (8).

CONSERVED ELEMENTS AND STRUCTURAL ORGANIZATION

T box leader RNAs are characterized by three helical regions (Stems I, II, and III) and a pseudoknot (Stem IIA/B) preceding the mutually exclusive terminator and antiterminator helices (or sequestrator and anti sequestrator helices) (1) (Figure 2A). Within Stem I, the Specifier Sequence is found in an internal loop called the Specifier Loop, which in most T box RNAs contains an S-turn (Loop E) element. The distal region of Stem I includes conserved sequence motifs in the AG bulge and terminal loop. In *glyQS* leader RNAs, these motifs interact to form a platform that contacts the tRNA D-/T-loops through a stacking interaction (10, 11). A structurally conserved internal loop or small bulge between the Specifier Loop and terminal region acts as a hinge to allow Stem I to bend and contact the tRNA at the anticodon and the D-/T-loop region. Below the Specifier Loop is a kink-turn (K-turn or GA motif) (12), which is a common RNA structural element that introduces a ~120° angle in the backbone (13). The Stem II domain, which usually stacks onto the base of the K-turn, varies in size and typically contains an S-turn within an internal loop (9). A pseudoknot element is usually found immediately downstream of Stem II, and is formed when nucleotides in the loop of Stem IIA pair with the downstream single stranded region to form Stem IIB (14). In *tyrS*, both the S-turn in Stem II and the pseudoknot are required for efficient antitermination, but their roles in the T box mechanism are not fully understood. Mutational analysis indicated that both the secondary structure and primary sequence of the pseudoknot element are required for antitermination (14). Stem III is found just upstream of

the antiterminator element, and varies considerably in sequence and length but is conserved in nearly all T box leader RNAs.

Although T box RNAs contain many conserved structural elements, not all elements are found in every leader RNA. All glycyl T box genes lack Stem II and the pseudoknot, which are replaced by a single stranded linker region between Stem I and Stem III (6). Alanyl T box genes are split into two classes that either lack (e.g., *C. acetobutylicum alaS*) or include (e.g., *B. subtilis alaS*) Stem II and the pseudoknot (7). Threonyl genes lack the S-turn motif in the Specifier Loop, which suggests an alternate presentation of the Specifier Sequence. The *ileS* genes in Actinobacteria contain Stem I variants that lack the motifs in the Stem I terminal region that comprise the tRNA-binding platform (8). In the ultrashort class of *ileS* RNAs, the Specifier Sequence is in the terminal loop, while another class contains sequences above the Specifier Loop that do not resemble those in the canonical Stem I. Studies of these structural variants demonstrated that all are functional in vitro, in vivo, or both. A structural element that is absent from one T box gene is often essential in another, which suggests that the variant classes achieve proper tRNA binding and regulation through other compensatory elements.

PHYLOGENETIC DISTRIBUTION

Transcriptional units regulated by the T box system have been identified based on the conservation of the T box sequence as well as other structural elements (9, 15). The T box mechanism is proposed to have arisen in a common ancestor to the Firmicutes, Actinobacteria, Chloroflexi, and Deinococcus-Thermus groups, with scattered examples of horizontal gene transfer to members of Deltaproteobacteria. This regulatory mechanism is most prevalent in Firmicutes, where many species contain multiple genes and operons under T box regulation, in some cases representing over 1% of transcriptional units. Of the gene classes that are regulated by the T box system, aaRS genes comprise the vast majority, followed by amino acid biosynthesis, amino acid transport, and regulatory genes. All 20 canonical amino acids are represented in the T box family, and this mechanism is the most common way of regulating aaRS genes in Firmicutes.

T BOX LEADER RNA-tRNA INTERACTIONS

Specifier Sequence-tRNA Anticodon Interaction

The main specificity determinant of the T box system is the base pairing of the Specifier Sequence with the tRNA anticodon. This leader RNA-tRNA interaction was first demonstrated genetically in the initial study of the *tyrS* T box mechanism (1, 4). Further studies examined tRNA specificity in more detail by testing additional mutations in the Specifier Sequence (and antiterminator bulge) in a *tyrS-lacZ* fusion, which enabled response to many different uncharged tRNA species (16). Several non-cognate tRNAs can promote *tyrS* antitermination (with reduced efficiency relative to tRNA^{Tyr}) with the appropriate changes to the Specifier Sequence and antiterminator bulge, but some tRNAs did not promote readthrough at a detectable level. Of the tRNAs that can promote readthrough, none are able to reach the induction levels of wild-type tRNA^{Tyr}, which indicates that there are additional specificity determinants in the tRNA and/or leader RNA that contribute to

maximal antitermination. Studies of the *B. subtilis* *ilv-leu*, *valS*, and *proBA* systems yielded similar results, where mutations to the Specifier Sequence resulted in a response to limitation for a different amino acid (17–19). Additionally, the wild-type tRNA-leader RNA interaction for all systems promoted higher antitermination compared to the non-cognate tRNAs, consistent with the results from *tyrS*.

In *C. acetobutylicum*, the *aspS2o-gatCABo* operon is regulated by a T box riboswitch with overlapping Specifier Sequences that bind both tRNA^{Asn} and tRNA^{Glu} (20). The indirect sensing of asparagine and glutamate levels permits appropriate expression of a non-discriminating AspRS and the GatCAB complex of the tRNA-dependent transamidation pathway, which connects the metabolism of these amino acids. In silico analysis suggests that T box genes or operons involved in the metabolism of multiple amino acids are more likely to contain a larger Specifier Loop capable of binding multiple physiologically relevant tRNA species. The ability of such T box RNAs to respond to the charging levels of multiple tRNAs may be a mechanism to fine-tune the expression of genes involved in complex metabolic networks.

The base pairing of the tRNA anticodon with the Specifier Sequence resembles a codon-anticodon interaction, but it occurs independently of the ribosome. The tRNA-mRNA interaction is strictly monitored during translation, which requires base pairing at the first two positions with tolerance for mismatches and non-Watson-Crick interactions at the third (wobble) position. Phylogenetic analysis of T box family genes revealed a bias for C at the third position of the Specifier Sequence, which does not correlate with codon bias or abundance of tRNA isoacceptors (1, 9). This observation suggested that the constraints that govern the Specifier Sequence-anticodon interaction differ from those for the anticodon-codon interaction in the context of the ribosome. The rules governing the Specifier Sequence-anticodon interaction were tested in vitro using the *B. subtilis* *glyQS* system and tRNA^{Gly}, which natively does not contain modifications in the anticodon loop (21). Mismatches of all possible combinations were introduced at every position of the base pairing in both the tRNA and Specifier Sequence, and the effects on tRNA binding and antitermination were determined. In general, mismatches are least well tolerated at the second position, and mismatches are better tolerated at the first and third positions. Nearly all mismatches are better tolerated for tRNA binding than for antitermination, indicating that antitermination has more stringent requirements for the Specifier Sequence-anticodon pairing. Overall, the binding affinity is related to the predicted stability of the interaction between the Specifier Sequence and anticodon, which differs from the mechanisms used by the ribosome to ensure accurate translation.

Antiterminator-tRNA Acceptor Arm Interaction

In addition to the base pairing between the Specifier Sequence and anticodon, another base pairing interaction occurs between the acceptor arm of uncharged tRNA and the antiterminator bulge. This base pairing interaction, which stabilizes the antiterminator helix and prevents formation of the terminator helix, is essential for gene expression. Pairing between these sequences was proposed initially due to the complementarity and covariation between residues at the 5' end of the antiterminator bulge (5'-UGGN-3') and the acceptor

end of tRNA (5'-NCCA-3') (4). The specificity of this interaction was first shown in *tyrS*, where mutational analysis of the tRNA discriminator base and antiterminator bulge revealed that base pairing is required for maximal expression and induction. Additional studies provided biochemical evidence for the specificity of this interaction in vitro and indicated that tRNA tertiary structure is also important for this interaction (22).

The specificity of the antiterminator-acceptor arm interaction is achieved by base pairing, but the interaction is stabilized primarily by coaxial stacking (23). The tRNA acceptor end-antiterminator bulge pairing stacks with the antiterminator helix below the bulge, which extends the coaxial stacking of the tRNA T stem and acceptor stem. In this way, uncharged tRNA becomes an integral part of the antiterminator domain to prevent formation of the terminator helix. Most charged tRNAs in the cell are bound to EF-Tu, but the antiterminator discriminates against charged tRNA in the absence of EF-Tu by directly sensing the presence of the amino acid at the 3' end (23). Even the smallest amino acid, glycine (57 Da), destabilizes the antiterminator-acceptor arm interaction through steric rejection and causes maximal termination. The competition between charged and uncharged tRNAs has been studied in vitro using a charged tRNA mimic that contains an additional C residue at the 3' end (EXIC), which prevents stabilization of the antiterminator (5). Both charged and uncharged tRNAs interact with the Specifier Sequence in Stem I, so charged tRNA acts as a competitive inhibitor to prevent binding of uncharged tRNA. These studies revealed that tRNA-EXIC can interact with the leader RNA until the antiterminator helix is fully transcribed, at which point the transcriptional fate of the gene is determined. This ability to monitor both charged and uncharged tRNA indicates that the T box riboswitch senses the tRNA charging ratio rather than the absolute levels of uncharged tRNA, which is appropriate for regulation of genes the products of which are involved in conversion of uncharged tRNA to charged tRNA, allowing both induction by the substrate and feedback repression by the product of the pathway.

The T box antiterminator domain was the first RNA shown to perform several functions previously observed only in proteins, particularly aaRSs. With as few as 30 nucleotides, this RNA recognizes the tRNA discriminator base as a specificity determinant and directly evaluates tRNA charging status. Using these functions as inputs, the antiterminator then acts as a transcriptional switch to control gene expression, where it exploits the structure of the tRNA ligand to promote its own stabilization. The variety of functions accomplished through this RNA-RNA interaction in the absence of proteins highlights the versatility of RNA, and adds to the growing list of biochemical functions of this macromolecule.

Stem I Platform-tRNA Elbow Interaction

For many years, the two sites of base pairing between T box leader RNA and tRNA were the only known intermolecular contacts. The conservation of secondary structure and sequence motifs in T box RNAs suggested that there were likely to be additional leader RNA-tRNA interactions. In particular, the highly conserved residues within the Stem I terminal region were known to be essential for antitermination, but their mechanistic role was not understood (14). Crystallization of the *glyQS* Stem I terminal region revealed a structure that was predicted to interact with tRNA (24), and this interaction was confirmed by co-

crystallization of the Stem I domain with tRNA (10, 11) (Figure 3). The conserved sequence motifs in the AG bulge and Stem I terminal loop interact to form a docking platform that contacts the tRNA D-/T-loop elbow region. The tRNA elbow is a structural feature of the canonical tRNA L-shape, and is formed by tertiary interactions between residues in the D-loop and T-loop.

The Stem I domain monitors both the anticodon sequence and the tertiary structure of the cognate tRNA, and these two interactions comprise the majority of the tRNA binding affinity for *glyQS* leader RNA (11). The helical region between the Specifier Loop and the terminal region tracks along the tRNA body from the anticodon stem to the elbow, and the Stem I domain has been proposed to act as a molecular ruler for tRNA (25). Extending the length of Stem I by a full helical turn (+11 bp) eliminates tRNA binding, which can be restored by extending the tRNA anticodon stem by the same length, suggesting that both Stem I contacts are important for binding (10). The order of interaction for the two Stem I binding sites remains unknown, but the Stem I platform is transcribed first and may interact with the pool of cellular tRNAs to facilitate Specifier Sequence-anticodon pairing. Alternatively, tRNA may bind the Specifier Sequence first, followed by anchoring and stabilization by the Stem I platform. Currently, all structural data for the T box Stem I platform are from glycyl T box genes (e.g., *glyQS*), which contain a unique set of residues that comprise that platform that are not present in the majority of T box family genes. These glycyl-specific residues may contribute to tRNA^{Gly} specificity, either by selecting for cognate tRNA^{Gly} or discriminating against non-cognate tRNAs. The role of the platform in tRNA specificity is currently under investigation, and the structural differences between the consensus and glycyl Stem I platforms remain unknown.

The interaction between the tRNA elbow and T box Stem I platform is structurally analogous to similar interactions that occur in the ribosome and RNase P (26). Both the ribosome and RNase P are essential in all three domains of life, and their RNA-tRNA elbow interactions are crucial to their cellular functions. In 23S rRNA, an interlocking T-loop module in the L1 stalk stacks with the tRNA elbow during the P/E site transition (27). This interaction is essential for directing the translocation and ejection of deacylated tRNA during the elongation cycle. RNase P, a ribonucleoprotein required for tRNA 5' end maturation (Chapter by Mohanty and Kushner), contains an interlocking T-loop module in the J11/12-J12/11 domain (28). The interlocking T-loop motif is defined primarily by structure and not sequence, so although these three distinct RNAs all utilize this motif to bind the tRNA elbow, they do not share a common evolutionary origin or a conserved sequence. Furthermore, the docking platform in the ribosome and RNase P approaches the elbow from the tRNA D-loop side, while the T box Stem I approaches from the tRNA T-loop side. This indicates that each double T-loop-tRNA elbow interaction arose independently through a striking example of convergent evolution.

Stem II and Pseudoknot Interactions with tRNA

Most in vitro studies of T box regulation have relied on the *glyQS* system, which lacks the conserved Stem II and Stem IIA/B pseudoknot elements found in the majority of T box genes. In *tyrS*, the S-turn in Stem II and the pseudoknot are required for maximal

antitermination in vivo (14), but the exact roles of these elements are not understood. A role for the region between Stem I and Stem III in positioning of the tRNA acceptor end for interaction with the antiterminator was suggested by studies with *B. subtilis glyQS* in which in vitro antitermination was disrupted by extension of the tRNA acceptor arm by a half-turn of the helix, but restored by extension by a full turn (+11 bp) (29). Introduction of a second full turn (+22 bp) was not tolerated. As this region is a single-stranded linker in *glyQS* genes, it appears that it allows some flexibility in its measurement of tRNA length, but positioning of the tRNA in the correct orientation for interaction with the antiterminator is required.

The translational *ileS* systems from Actinobacteria permit the biochemical analysis of these domains and their role in tRNA-dependent regulation (8). The Stem II S-turn and the pseudoknot of *Mycobacterium smegmatis ileS* contribute substantially to tRNA binding affinity (30). Furthermore, SHAPE and tRNA crosslinking experiments provide evidence that the Stem II S-turn contacts the tRNA T-arm and the pseudoknot contacts the D-loop. This is the first biochemical evidence that shows these elements contribute to the molecular recognition of the tRNA ligand. While this *ileS* RNA contains the structural elements lacking in *glyQS*, the Stem I domain belongs to the ultrashort class, which lacks the Stem I platform and contains the Specifier Sequence within the terminal loop. It remains to be discovered if these same interactions occur in T box RNAs that contain the canonical Stem I, Stem II S-turn, and pseudoknot elements together, although their conservation and sensitivity to mutation is consistent with the hypothesis that they are utilized in RNAs in which they are present.

STRUCTURAL ANALYSES OF T BOX LEADER RNAS

Antiterminator Domain

The structures of riboswitch RNAs are essential to their biological function, but high-resolution structural analysis can be particularly challenging if such RNAs are inherently flexible. To address this, riboswitches are often divided into discrete structural domains and studied in isolation. The first structural study of the T box riboswitch was the NMR analysis of a model antiterminator RNA based on the *tyrS* sequence (31). This RNA is composed of two short helices (A1 and A2) separated by a 7 nt bulge that contains the 5'-UGGA-3' motif that pairs with the acceptor arm of uncharged tRNA. The remaining three nucleotides of the bulge (5'-ACC-3') are highly conserved throughout the T box family, and mutations in this sequence disrupt *tyrS* antitermination in vivo (14). The NMR analysis revealed a structural basis for the observed conservation; the residues in the 3' portion of the bulge participate in a stacking interaction that introduces an -80° kink between the helices. In contrast to the relatively rigid 3' bulge residues, the bases that pair with the tRNA acceptor arm are more flexible and are likely to sample different conformations to achieve an induced fit with tRNA. Studies of fluorescently-labeled antiterminator RNAs indicate that the bulge structure changes in the presence of Mg^{2+} (32). This structural change is necessary for proper base pairing with the tRNA acceptor arm, and additional fluorescence changes upon tRNA binding support the induced fit model.

Specifier Loop and Kink-turn Motif

The canonical Stem I domain contains several RNA structural motifs, including an S-turn motif within the Specifier Loop and a kink-turn motif at the base of Stem I. The first structural information for the S-turn was from NMR studies of the lower portion of *tyrS* Stem I (33). Within the Specifier Loop, the S-turn is directly above the Specifier Sequence, and the conserved 5'-AGUA-3' motif on the 5' side of Stem I forms non-canonical interactions with the 5'-GAA-3' motif on the opposite side of the loop. These interactions generate a distortion in the phosphate backbone that causes the Specifier Sequence nucleotides to rotate into the minor groove. A rotation toward the minor groove exposes more of the Watson-Crick face, such that the S-turn motif assists in the presentation of the Specifier Sequence for tRNA binding. Structural analyses of this region in complex with tRNA revealed that the conserved purine 3' of the Specifier Sequence stabilizes the interaction by stacking with the leader RNA-anticodon duplex (10, 11). Additionally, the structural perturbation caused by the S-turn allows the Specifier Loop to accommodate large post-transcriptional modifications that are common at position 37 in the anticodon loop (34).

The kink-turn element is found just below the Specifier Loop, and is comprised of flanking 5'-GA-3' residues on both sides of an asymmetric internal loop that introduces a kink in the helical axis (12). This structural motif is found in many RNAs, including other riboswitches, 23 S rRNA, and RNase P, where it is typically a binding site for an L7Ae family protein (35) (Chapter by Meyer). The T box mechanism does not appear to require this RNA-protein interaction, as *tyrS* expression is not affected in strains in which L7Ae homologs are inactivated by mutation (Grundy & Henkin, unpublished), and other leader RNAs function in vitro without additional protein factors (6–8). The kink-turn in T box leader RNAs is predicted to determine the relative orientation of the Stem I domain with the downstream RNA, which would facilitate simultaneous tRNA binding with Stem I and the antiterminator bulge. NMR analysis of the kink-turn in *tyrS* revealed the sheared G-A pairs that are characteristic of this motif, but the RNA is in an extended conformation rather than the canonical kinked conformation (36). The crystal structure of *glyQS* Stem I in complex with tRNA contains the kink-turn, but the kinked conformation is stabilized by the L7Ae homolog used to promote co-crystallization (11). The structure of this motif is likely to be influenced by the context of the upstream and downstream RNA (37), so further studies of this structural element in the context of the full length leader RNA are necessary to validate its role in the T box mechanism.

Stem I Platform and Hinge

The crystal structures of *glyQS* Stem I from *Geobacillus kaustophilus* (10, 24) and *Oceanobacillus iheyensis* (11) revealed key details of Stem I architecture. The most striking observation was the formation of a platform at the terminal region of Stem I that interacts with the tRNA elbow (Figure 3). This platform is formed by the interaction of two highly conserved sequence motifs, one of which (5'-AGAGA-3'; 5'-AGCGA-3' in *glyQS*) is within the AG bulge, and the other (5'-GSUGNRA-3'; S=G or C; R=A or G) is within the terminal loop. The structure formed by this interaction alters the direction of base stacking by 90° relative to the Stem I helix to allow coaxial stacking with the tRNA. The third residue in the AG bulge motif (C in most glycyl T box genes) forms a cis-Watson-Crick interaction with a

glycyl-specific G at the 3' end of the terminal loop, and an additional Hoogsteen interaction from a residue (G or A) in the terminal loop forms a base triple. This triplet is the outermost platform layer of a total of five base layers that stack with tRNA. In the majority of T box family genes, the residues that comprise the base triple that contacts the tRNA elbow are either conserved as other residues or are absent. Further studies are needed to investigate the structural differences between glycyl and non-glycyl T box Stem I platforms.

The helical region between the AG bulge and Specifier Loop is always disrupted by either a bulge or internal loop that is not conserved in sequence. The disruption to the helix introduces structural flexibility that acts as a hinge to bend the flanking helices and form a cradle-shaped Stem I domain. In the *G. kaustophilus glyQS* Stem I, the hinge is an asymmetric internal loop with 3 nucleotides opposite a single nucleotide (10), which is similar to the predicted *B. subtilis glyQS* hinge. In contrast, the hinge in *O. iheyensis* is comprised of only 2 bulged nucleotides on the 5' side (11). The additional unpaired nucleotides on the 5' side of both Stem I structures causes a length discrepancy that kinks the distal portion of Stem I toward the tRNA elbow. This hinge element is likely to be crucial for Stem I to properly orient the Specifier Loop and terminal platform for optimal interaction with tRNA, and mutations to this region decrease tRNA binding (11).

Leader RNA-tRNA Complexes

The high resolution crystal structures have provided extensive insight into the Stem I-tRNA interactions. Comparisons of these complexes with the structures of both RNAs in the unbound state reveals the changes that occur upon binding. The anticodon loop of the free tRNA^{Gly} adopts an extended conformation (38), while the bound tRNA^{Gly} anticodon loop is more ordered to interact properly with the Specifier Sequence (10, 11). Stem I also induces tRNA bending at the 26–44 pair at the junction of the anticodon stem and the D stem. This junction is inherently flexible, and its flexibility is important for tRNA bending during translation, particularly in the P/E hybrid state where it is bent up to 37° (39). Flexibility in tRNA structure was shown to be important for both *glyQS* and *alaS* antitermination in vitro (7, L. C. Liu, F. J. Grundy and T. M. Henkin, personal communication), consistent with the structural studies. These comparisons reveal that there are both localized and global structural changes in tRNA that occur upon interaction with Stem I, and there are likely to be additional changes when more of the leader RNA is present.

The Specifier Loop also undergoes changes upon tRNA binding, such that the Specifier Sequence residues rotate farther out from the minor groove to expose the Watson-Crick face (11). This rotation also affects the position of the conserved 3' A residue, allowing it to stack with the Specifier Sequence-anticodon pairing. In the *G. kaustophilus* Specifier Loop, the S-turn element is separated from the Specifier Sequence by two nonconserved nucleotides, each of which forms a hydrogen bond with nucleotides in the anticodon loop (10). These additional nucleotides may contribute to tRNA specificity in the T box genes that contain them, or they may generate overlapping Specifier Sequences that allow response to multiple tRNA species (20, 40). The structural changes that occur in both tRNA and Stem I in addition to the flexible hinges in these RNAs support a mutually induced fit model of binding.

The crystal structure of the Stem I-tRNA complex revealed new aspects of the structural basis for tRNA recognition, but it lacks important downstream RNA elements such as Stem III and the antiterminator. Recent studies report low-resolution structural models of the full complex in solution using small-angle x-ray scattering (SAXS). SAXS-derived molecular envelopes of the *B. subtilis* *glyQS* leader RNA (up to the antiterminator) in the absence of tRNA suggest an elongated structure (41). Molecular envelopes were also determined for mutants containing extensions or deletions in helical regions to determine the positions of Stem I, Stem III, and the antiterminator. The leader RNA-tRNA complex resembles a flattened disc shape, which suggests that all RNA elements lie in the same plane (42). Modeling the RNAs to fit this envelope suggests that Stem III stacks with the A1 helix of the antiterminator. The additional stabilization of the antiterminator conformation by Stem III would explain why Stem III is highly conserved but variable in sequence and length.

The *O. iheyensis* *glyQS* leader RNA-tRNA complex (41) forms a different structure compared to the *B. subtilis* complex (42). The *O. iheyensis* complex is more linear and elongated, and molecular modeling indicates that the Stem I platform-elbow interaction does not occur in the antiterminator conformation. The interaction is restored when the tRNA is unable to pair with the antiterminator (by removal of the 5'-UCCA-3' end), which led to a proposed capture and release mechanism by the Stem I platform. This mechanism is not supported by the structural model of the *B. subtilis* complex, where the Stem I platform and antiterminator bulge interact with tRNA simultaneously. Additional studies are needed to investigate the equilibrium binding state of the leader RNA-tRNA complex to resolve the apparent structural discrepancies.

THE T BOX REGULATORY SYSTEM AS AN ANTIBIOTIC TARGET

Bacterial riboswitch RNAs, including the T box riboswitch, have drawn interest as a potential target for the development of novel antibiotic compounds (43). Many T box family genes, such as aaRS genes, are essential and disruption of the regulation of these genes can reduce competitive fitness or be lethal. Many human pathogens in Firmicutes contain multiple T box regulated genes; for example, *B. anthracis* is predicted to contain a T box regulon of 39 transcriptional units, 21 of which include aaRS genes (9). The presence of multiple essential targets in combination with the default “off” state of the T box riboswitch would help to prevent mutations that confer antibiotic resistance. Mutations in the T box leader RNA that prevent interaction with the antibiotic compound may also prevent interaction with tRNA, resulting in loss of expression. The identification and characterization of candidate compounds that target the T box regulatory system is ongoing (44). The primary target in these studies is a model antiterminator RNA based on the *B. subtilis* *tyrS* sequence (22). The antiterminator is highly conserved and is essential for gene expression. The goal of these studies is to identify compounds that interact with the antiterminator in a way that prevents it from binding the tRNA acceptor end, but does not excessively stabilize the antiterminator independent of tRNA binding. Several classes of chemical compounds have been identified that bind the antiterminator RNA with structural specificity and low micromolar affinity (45, 46) and disrupt binding of the tRNA acceptor end (47).

Several commonly used antibiotics that are known to interact with structured RNA have been shown to interact with T box leader RNA and affect regulation (48). Many antibiotics that inhibit protein synthesis by binding to ribosomal RNA can also bind to other non-coding RNAs with potentially antagonistic or synergistic antibiotic effects (49). Some antibiotics, such as tigecycline, increase tRNA-dependent antitermination, while others, such as linezolid, decrease tRNA-dependent antitermination (48). Molecular modeling of the binding of the antibiotics to the RNAs suggests that they interact with crucial RNA sequences such as the tRNA elbow, anticodon, Specifier Loop, and Stem I terminal region. Additional studies of T box riboswitch structure and function will aid in the development of antibiotic compounds that specifically target this widespread regulatory mechanism.

CONCLUSIONS

For many bacterial species, riboswitches are one of the primary mechanisms to control gene expression in response to changing environmental conditions. The T box riboswitch is a widespread mechanism that regulates the expression of amino acid-related genes, most commonly aaRS genes, by directly monitoring tRNA aminoacylation. The phylogenetic distribution of this system highlights its evolutionary adaptability to monitor tRNAs of all canonical amino acid classes (9). It remains unknown why T box regulons vary so much among different species, but it is likely related to the different ecological niches that require optimal responses to changes in the environment and nutrient availability.

With a typical length of less than 300 nucleotides, the T box leader RNA recognizes the sequence and structure of tRNA and determines its aminoacylation status to control premature transcription termination or translation initiation. A combination of genetic, biochemical, and structural techniques has revealed many fundamental aspects of the leader RNA-tRNA interactions. These interactions include anticodon recognition and stacking with the tRNA elbow, which are similar to interactions with tRNA in other cellular RNAs, including the ribosome and RNase P. Furthermore, the T box RNAs utilize tRNA recognition determinants analogous to those often utilized by aaRS enzymes. Despite the significant progress in understanding this regulatory system since its discovery, many aspects remain to be investigated. In particular, it is evident that variations in RNA structure yield alternate modes of tRNA recognition for which structural information is currently lacking. Further studies of both canonical and variant T box RNAs and their interactions with tRNA will contribute to our understanding of modes of RNA-RNA recognition.

ACKNOWLEDGEMENTS

This work was supported by National Institutes of Health Institute of General Medical Sciences grant R01 GM047823. We thank members of the Henkin lab past and present for helpful discussions.

REFERENCES

1. Grundy FJ, Henkin TM. 1993 tRNA as a positive regulator of transcription antitermination in *B. subtilis*. *Cell* 74:475–482. [PubMed: 8348614]
2. Giegé R, Jühling F, Pütz J, Stadler P, Sauter C, Florentz C. 2012 Structure of transfer RNAs: Similarity and variability. *Wiley Interdiscip Rev RNA* 3:37–61. [PubMed: 21957054]

3. Henkin TM, Glass BL, Grundy FJ. 1992 Analysis of the *Bacillus subtilis tyrS* gene: conservation of a regulatory sequence in multiple tRNA synthetase genes. *J Bacteriol* 174:1299–1306. [PubMed: 1735721]
4. Grundy FJ, Rollins SM, Henkin TM. 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. [PubMed: 8045882]
5. Grundy FJ, Yousef MR, Henkin TM. 2005 Monitoring uncharged tRNA during transcription of the *Bacillus subtilis glyQS* gene. *J Mol Biol* 346:73–81. [PubMed: 15663928]
6. Grundy FJ, Winkler WC, Henkin TM. 2002 tRNA-mediated transcription antitermination *in vitro*: codon-anticodon pairing independent of the ribosome. *Proc Natl Acad Sci USA* 99:11121–11126. [PubMed: 12165569]
7. Liu LC, Grundy FJ, Henkin TM. 2015 Non-conserved residues in *Clostridium acetobutylicum* tRNA^{Ala} contribute to tRNA tuning for efficient antitermination of the *alaS* T box riboswitch. *Life* 5:1567–1582. [PubMed: 26426057]
8. Sherwood AV, Grundy FJ, Henkin TM. 2015 T box riboswitches in Actinobacteria: translational regulation via novel tRNA interactions. *Proc Natl Acad Sci USA* 112:1113–1118. [PubMed: 25583497]
9. Gutierrez-Preciado A, Henkin TM, Grundy FJ, Yanofsky C, Merino E. 2009 Biochemical features and functional implications of the RNA-based T-box regulatory mechanism. *Microbiol Mol Biol Rev* 73:36–61. [PubMed: 19258532]
10. Grigg JC, Ke A. 2013 Structural determinants for geometry and information decoding of tRNA by T box leader RNA. *Cell* 21:2025–2032.
11. Zhang J, Ferré-D' Amaré AR. 2013 Co-crystal structure of a T-box riboswitch stem I domain in complex with its cognate tRNA. *Nature* 500:363–366. [PubMed: 23892783]
12. Winkler WC, Grundy FJ, Murphy BA, Henkin TM. 2001 The GA motif: an RNA element common to bacterial antitermination systems, rRNA, and eukaryotic RNAs. *RNA* 7:1165–1172. [PubMed: 11497434]
13. Klein DJ, Schmeing TM, Moore PB, Steitz TA. 2001 The kink-turn: A new RNA secondary structure motif. *EMBO J* 20:4214–4221. [PubMed: 11483524]
14. Rollins SM, Grundy FJ, Henkin TM. 1997 Analysis of *cis*-acting sequence and structural elements required for antitermination of the *Bacillus subtilis tyrS* gene. *Mol Microbiol* 25:411–421. [PubMed: 9282752]
15. Vitreschak A, Mironov A. 2008 Comparative genomic analysis of T-box regulatory systems in bacteria. *RNA* 14:717–735. [PubMed: 18359782]
16. Grundy FJ, Hodil SE, Rollins SM, Henkin TM. 1997 Specificity of tRNA-mRNA interactions in *Bacillus subtilis tyrS* antitermination. *J Bacteriol* 179:2587–2594. [PubMed: 9098057]
17. Marta PT, Ladner RD, Grandoni JA. 1996 A CUC triplet confers leucine-dependent regulation of the *Bacillus subtilis ilv-leu* operon. *J Bacteriol* 178:2150–2153. [PubMed: 8606198]
18. Luo D, Leautey J, Grunberg-Manago M, Putzer H. 1997 Structure and regulation of expression of the *Bacillus subtilis* valyl-tRNA synthetase gene. *J Bacteriol* 179:2472–2478. [PubMed: 9098041]
19. Brill J, Hoffmann T, Putzer H, Bremer E. 2011 T-box-mediated control of the anabolic proline biosynthetic genes of *Bacillus subtilis*. *Microbiology* 157:977–987. [PubMed: 21233158]
20. Saad NY, Stamatopoulou V, Braye M, Drains D, Stathopoulos C, Becker HD. 2013 Two-codon T-box riboswitch binding two tRNAs. *Proc Natl Acad Sci USA* 110:12756–12761. [PubMed: 23858450]
21. Caserta E, Liu LC, Grundy FJ, Henkin TM. 2015 Codon-anticodon recognition in the *Bacillus subtilis glyQST* box riboswitch: RNA-dependent codon selection outside the ribosome. *J Biol Chem* 290:23336–23347. [PubMed: 26229106]
22. Gerdeman MS, Henkin TM, Hines JV. 2002 *In vitro* structure-function studies of the *Bacillus subtilis tyrS* mRNA antiterminator: evidence for factor-independent tRNA acceptor stem binding specificity. *Nucleic Acids Res* 30:1065–1072. [PubMed: 11842119]
23. Zhang J, Ferré-D' Amaré AR. 2014 Direct evaluation of tRNA aminoacylation status by the T-box riboswitch using tRNA-mRNA stacking and steric readout. *Mol Cell* 55:148–155. [PubMed: 24954903]

24. Grigg JC, Chen Y, Grundy FJ, Henkin TM, Pollack L, Ke A. 2013 T box RNA decodes both the information content and geometry of tRNA to affect gene expression. *Proc Natl Acad Sci USA* 110:7240–7245. [PubMed: 23589841]
25. Zhang J, Ferré D' Amaré AR. 2016 The tRNA elbow in structure, recognition and evolution. *Life* 6:3.
26. Lehmann J, Jossinet F, Gautheret D. 2013 A universal RNA structural motif docking the elbow of tRNA in the ribosome, RNase P and T-box leaders. *Nucleic Acids Res* 41:5494–5502. [PubMed: 23580544]
27. Trabuco LG, Schreiner E, Eargle J, Cornish P, Ha T, Luthey-Schulten Z, Schulten K. 2010 The role of L1 stalk-tRNA interaction in the ribosome elongation cycle. *J Mol Biol*, 402:741–760. [PubMed: 20691699]
28. Reiter NJ, Osterman A, Torres-Larios A, Swinger KK, Pan T, Mondragón A. 2010 Structure of a bacterial ribonuclease P holoenzyme in complex with tRNA. *Nature* 468:784–789. [PubMed: 21076397]
29. Yousef MR, Grundy FJ, Henkin TM. 2003 tRNA requirements for *glyQS* antitermination: a new twist on tRNA. *RNA* 9:1148–1156. [PubMed: 12923262]
30. Sherwood AV, Frandsen JK, Grundy FJ, Henkin TM. 2018 New tRNA contacts facilitate ligand binding in a *Mycobacterium smegmatis* T box riboswitch. *Proc Natl Acad Sci USA* 115:3894–3899. [PubMed: 29581302]
31. Gerdeman MS, Henkin TM, Hines JV. 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. [PubMed: 12547201]
32. Means JA, Simson CM, Zhou S, Rachford A, Rack JJ, Hines JV. 2009 Fluorescence probing of T box antiterminator RNA: insights into riboswitch discernment of the tRNA discriminator base. *Biochem Biophys. Res Commun* 389:616–621.
33. Wang J, Henkin TM, Nikonowicz EP. 2010 NMR structure and dynamics of the Specifier Loop domain from the *Bacillus subtilis* *tyrS* T box leader RNA. *Nucleic Acids Res* 38:3388–3398. [PubMed: 20110252]
34. Chang AT, Nikonowicz EP. 2013 Solution NMR determination of hydrogen bonding and base pairing between the *glyQS* T box riboswitch specifier domain and the anticodon loop of tRNA^{Gly}. *FEBS Letters* 587:3495–3499. [PubMed: 24036450]
35. Schroeder KT, McPhee SA, Ouellet J, Lilley DMJ. 2010 A structural database for k-turn motifs in RNA. *RNA* 16:1463–1468. [PubMed: 20562215]
36. Wang J, Nikonowicz EP. 2011 Solution structure of the K-turn and Specifier Loop domains from the *Bacillus subtilis* *tyrS* T-box leader RNA. *J Mol Biol* 408:99–117. [PubMed: 21333656]
37. Lilley DMJ. 2014 The K-turn motif in riboswitches and other RNA species. *Biochim Biophys Acta* 1839:995–1004. [PubMed: 24798078]
38. Chang AT, Nikonowicz EP. 2012 Solution nuclear magnetic resonance analyses of the anticodon arms of proteinogenic and nonproteinogenic tRNA^{Gly}. *Biochemistry* 51:3662–3674. [PubMed: 22468768]
39. Dunkle JA, Wang L, Feldman MB, Pulk A, Chen VB, Kapral GJ, Cate JHD. 2011 Structures of the bacterial ribosome in classical and hybrid states of tRNA binding. *Science* 332:981–984. [PubMed: 21596992]
40. Zhang J, Ferré-D' Amaré AR. 2015 Structure and mechanism of the T-box riboswitches. *Wiley Interdiscip Rev RNA* 6:419–433. [PubMed: 25959893]
41. Fang X, Michnicka M, Zhang Y, Wang YX, Nikonowicz EP. 2017 Capture and release of tRNA by the T-loop receptor in the function of the T-Box riboswitch. *Biochemistry* 56:3549–3558. [PubMed: 28621923]
42. Chetnani B, Mondragón A. 2017 Molecular envelope and atomic model of an antiterminated *glyQS* T-box regulator in complex with tRNA^{Gly}. *Nucleic Acids Res* 45:8079–8090. [PubMed: 28531275]
43. Thomas JR, Hergenrother PJ. 2008 Targeting RNA with small molecules. *Chemical Reviews* 108:1171–1224. [PubMed: 18361529]

44. Zhou S, Means JA, Acquaah-Harrison G, Bergmeier SC, Hines JV. 2012 Characterization of a 1,4-disubstituted 1,2,3-triazole binding to T box antiterminator RNA. *Bioorg Med Chem* 20:1298–1302. [PubMed: 22230198]
45. Means JA, Hines JV. 2005 Fluorescence resonance energy transfer studies of aminoglycoside binding to a T box antiterminator RNA. *Bioorg Med Chem Lett* 15:2169–2172. [PubMed: 15808490]
46. Orac CM, Zhou S, Means JA, Boehm D, Bergmeier SC, Hines JV. 2011 Synthesis and stereospecificity of 4,5-disubstituted oxazolidinone ligands binding to T-box riboswitch RNA. *J Med Chem* 54:6786–6795. [PubMed: 21812425]
47. Zhou S, Acquaah-Harrison G, Bergmeier SC, Hines JV. 2011 Anisotropy studies of tRNA-T box antiterminator RNA complex in the presence of 1,4-disubstituted 1,2,3-triazoles. *Bioorg Med Chem Lett* 21:7059–7063. [PubMed: 22001091]
48. Stamatopoulou V, Apostolidi M, Li S, Lamprinou K, Papakyriakou A, Zhang J, Stathopoulos C. 2017 Direct modulation of T-box riboswitch-controlled transcription by protein synthesis inhibitors. *Nucleic Acids Res* 45:10242–10258. [PubMed: 28973457]
49. Dar D, Shamir M, Mellin JR, Koutero M, Stern-Ginossar N, Cossart P, Sorek R. 2016 Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. *Science* 352:aad9822. [PubMed: 27120414]

Figure 1. The T box mechanism.

A. tRNA interacts with Stem I of the T box leader RNA at two locations. The tRNA anticodon base pairs with the Specifier Sequence (green), and the tRNA elbow stacks with the Stem I platform (orange), which is formed by interactions between conserved sequence motifs. The presence of an amino acid (AA) at the 3' end of a charged tRNA blocks the base-pairing interaction with a bulge in the antiterminator helix. The terminator helix forms and transcription terminates, which turns gene expression off. **B.** Uncharged tRNA also interacts with the Specifier Sequence and Stem I platform, and the acceptor arm base pairs

with a bulge in the antiterminator helix (cyan). The stabilization of the antiterminator prevents formation of the competing terminator helix, and RNA polymerase continues to transcribe the downstream coding sequence.

Figure 2. *B. subtilis* *tyrS* leader RNA and tRNA^{Tyr} secondary structure.

A. The *tyrS* sequence is numbered from the transcriptional start site to the end of the transcriptional terminator element. Conserved structural domains are labeled, including Stems I, II, IIA/B, III, and mutually exclusive terminator and antiterminator helices. The Stem IIB pseudoknot interaction is shown in magenta. The Specifier Sequence in the Specifier Loop and residues in the antiterminator bulge that base pair with tRNA are shown in green and cyan, respectively. The orange sequences in the AG bulge and Stem I terminal loop interact to form the Stem I platform that contacts the tRNA elbow. The red and blue labeled sequences interact to form the antiterminator element shown above the terminator conformation. The antiterminator is composed of helices A1 and A2. **B.** Cloverleaf structure of tRNA^{Tyr}. The anticodon sequence is shown in green, and the nucleotides in the acceptor arm that base pair with the *tyrS* antiterminator bulge are shown in cyan. The orange residues in the D-Loop and T-Loop (G19 and C56) interact to form the outermost tertiary interaction of the elbow and stack with the Stem I platform.

Figure 3. Co-crystal structure of the *glyQS* Stem I-tRNA complex.

The *Geobacillus kaustophilus glyQS* Stem I (gray) bound to tRNA^{Gly} (purple). The Specifier Sequence-anticodon interaction is shown in green, and the Stem I platform that stacks with the tRNA elbow is shown in orange. Adapted from (10) (PDB ID: 4MGN).