Structural basis of amino acid surveillance by higher-order tRNA-mRNA interactions

Shuang Li1,9, Zhaoming Su2,9, Jean Lehmann3, Vassiliki Stamatopoulou4, Nikoleta Giarimoglou4, Frances E. Henderson1, Lixin Fan5, Grigore D. Pintilie2, Kaiming Zhang2, Muyuan Chen6, Steven J. Ludtke6, Yun-Xing Wang5,7, Constantinos Stathopoulos4, Wah Chiu2,8* and Jinwei Zhang2,*

Amino acid availability in Gram-positive bacteria is monitored by T-box riboswitches. T-boxes directly bind tRNAs, assess their aminoacylation state, and regulate the transcription or translation of downstream genes to maintain nutritional homeostasis. Here, we report cocrystal and cryo-EM structures of Geobacillus kaustophilus and Bacillus subtilis T-box-tRNA complexes, detailing their multivalent, exquisitely selective interactions. The T-box forms a U-shaped molecular vise that clamps the tRNA, capturing its 3′ end using an elaborate ‘discriminator’ structure, and interrogates its aminoacylation state using a steric filter fashioned from a wobble base pair. In the absence of aminoacylation, T-boxes clutch tRNAs and form a continuously stacked central spine, permitting transcriptional readthrough or translation initiation. A modeled aminoacyl disrupts tRNA-T-box stacking, severing the central spine and blocking gene expression. Our data establish a universal mechanism of amino acid sensing on tRNAs and gene regulation by T-box riboswitches and exemplify how higher-order RNA-RNA interactions achieve multivalency and specificity.

Amino acids are among the most ubiquitous and essential nutrients for all known cellular life. Diverse mechanisms of amino acid sensing and regulation have evolved to appropriately supply the translation machinery and other cellular processes such as cell wall synthesis1. Free amino acids can be sensed by direct binding to protein sensors, including Sestrins2, CASTORs3, and TRAP4, or RNA sensors, as exemplified by the glycine, lysine and glutamine riboswitches5,6. Alternatively, amino acids can also be sensed in conjugated forms, especially in the form of aminoacyl tRNAs, which are the immediate substrates of protein synthesis. Sensing amino acids on the tRNAs can take advantage of the well-established coupling specificities between tRNAs and their aminoacyl-tRNA synthetases7–9. A number of cellular proteinaceous systems sense aminoacylation on tRNAs to detect nutrient limitation, such as translating ribosomes10 and ribosome-associated ppGpp synthase RelA in bacteria11,12 and α kinase GCN2 in eukaryotes13. Interestingly, mRNA elements have also been proposed to sense aminoacylation on tRNAs, such as the T-box riboswitches that are widespread in Gram-positive bacteria. However, direct sensing of tRNA aminoacylation by another RNA without any proteins, including the transcribing RNA polymerase (RNAP), has not been demonstrated14.

More than one thousand T-box riboswitches have been identified, including in most key human pathogens, such as Bacillus, Clostridium, and Staphylococci15–18. They control essential genes in amino acid biosynthesis, transport and tRNA aminoacylation19,16,19. T-boxes are attractive antimicrobial targets that exhibit drastically reduced antibiotic resistance20, presumably because their universal core structure allows the same compounds to simultaneously target multiple T-boxes, and thus, manifold metabolic pathways. T-boxes share a phylogenetically conserved architecture of two or three domains (Fig. 1a). An obligate 5′ stem I domain initially selects a cognate tRNA through base-pairing interactions to the tRNA anticodon and platform-stacking interactions to the tRNA elbow21–23 (Fig. 1a). An intervening stem II domain of unknown structure and function plays important yet uncharacterized roles in most T-boxes24–26. Curiously, the stem II domain is absent from the most-studied glycine-specific glyQ T-boxes. Next, a 3′ antiterminator or antisequestrator domain supposedly probes the docked tRNA 3′ end to sense aminoacylation and directs RNA conformational switching that governs transcription attenuation27,28 or translation initiation29,30. However, the nature of the T-box mRNA-tRNA interactions, especially those involving the tRNA 3′ end, is poorly understood, due to the lack of high-resolution structural information. Two recent low-resolution studies employing small-angle X-ray scattering (SAXS) led to divergent structural models and do not provide insights into the mechanism of amino acid sensing29,30. The sensory and regulatory functions of the T-boxes are anchored by a presumed ability to sense aminoacylation on a tRNA directly. This complex function is generally accomplished on the ribosome by large, specialized protein sensors, such as RelA31,32 and Gcn2 (ref. 13). Recently, a new class of T-box riboswitches were identified in Actinobacteria that regulate the initiation of translation instead of transcription termination33. These translational T-boxes harbor a...
more divergent stem I domain and a similar 3’ amino acid-responsive domain that either sequesters or exposes the Shine–Dalgarno sequence to control translation initiation26,27. The conserved secondary structure of the ‘antisquestrator’ domain is nearly identical to that of the transcriptional antiterminator, suggesting a unified mechanism of amino acid sensing and conformational switching27.

To investigate how the T-box riboswitches sense and respond to tRNA aminoacylation, we first define a minimal T-box region that selectively binds an uncharged tRNA, which we term the ‘T-box discriminator’. We then elucidate how the T-box discriminator, merely 66 nucleotides (nts) in size, directly senses tRNA aminoacylation by solving its 2.66-Å resolution cocry-EM structure bound to an uncharged tRNA. Furthermore, we also determine a 4.9-Å resolution cocrystal structure bound to charged tRNAs, despite exhaustive trials (Extended Data Fig. 2). This finding suggested that additional T-box sequences required for tRNA binding may be missing. Located immediately upstream of the antiterminator is a variable stem III and removed it from the sequence alignment, we found that the T-box 3′ region to interact with the tRNA 3′ end12. Located 3′ to stem I is a single-stranded linker, the highly variable stem III of unknown function, and an antiterminator domain consisting of two helices (A1 and A2) flanking a 7-nucleotide (nt) bulge15,16 (Fig. 1a and Extended Data Fig. 1). Surprisingly, the antiterminator alone, long assumed to be sufficient for tRNA 3′-end binding and aminoacylation sensing, did not appreciably bind tRNAs, despite exhaustive trials (Extended Data Fig. 2). This finding suggested that additional T-box sequences required for tRNA binding may be missing. Located immediately upstream of the antiterminator is a stem III hairpin that is ubiquitously present in all known T-boxes, but the length and sequences of its stem and loop are highly variable (Extended Data Fig. 1). Stem III was proposed to be a part of the flexible linker, to provide a RNAP pause site10 or to stack underneath helix A1 to stabilize the antiterminator31. When we folded the variable stem III and removed it from the sequence alignment, we discovered that the sequences that flank stem III are actually highly conserved, implying functional importance. Both flanks of stem III are purine rich and conform to a motif of 5′-RRRxG-stem III-AA-3′, where ‘x’ denotes any nucleotide (Fig. 1b,c and Extended Data Fig. 1). Remarkably, appending stem III and its flanking sequences to the binding-incompetent antiterminator conferred robust tRNA binding (Extended Data Fig. 2). We further show that the conserved purines are required for tRNA binding, as their deletions or

**Results**

**A compact T-box discriminator domain directly senses tRNA aminoacylation.** The T-box stem I initially docks a cognate tRNA through contacts to its anticodon and elbow and sets the stage for

---

**Fig. 1 | Overall structure of the T-box discriminator–tRNA complex.** a, gLQ T-box riboswitch–tRNA complex. Dotted lines denote stacking interactions between the tRNA elbow and the distal T loops of stem I. The molecular beacon harboring a 5′-Cy5 fluorophore (F) and 3′-dabcyl quencher (Q) to measure transcription readthrough is rendered in black. b, Sequence and secondary structure of the cocry-EM structure of G. kaustophilus gLQ T-box discriminator and tRNA90. Noncanonical base pairs are denoted by Leontis–Westhof symbols. Lines with embedded arrowheads denote chain connectivity. Green and gray dotted lines denote the purine minor groove and other hydrogen bonding interactions, respectively. The tRNA (shaded) is conventionally numbered, and gray spheres denote the purine minor groove. c, Sequence logo of stem III and its flanking regions. d, Representative ITC analysis of discriminator binding by uncharged (black) and 3′-phosphate (red) tRNAs. The error bars are statistical estimates of the uncertainties of individual injection heats calculated using NITPIC (Methods). The Kᵦ value is the mean and s.d., n = 3 biologically independent samples. e, f, Front and rear views of the overall cocrystal structure. tA76 is highlighted in bright green. Parallel lines connected to orthogonal lines denote stacking. Residues that are not modeled are indicated as gray spheres.
substitutions with uridines abolished tRNA binding. Crucially, this binding is selective for uncharged tRNA, as both binding and transcription readthrough were inversely correlated with the molecular volume of the tRNA 3′ end, supporting a proposed steric occlusion mechanism14 (Extended Data Fig. 2). Using isothermal titration calorimetry (ITC), we further showed that the mere addition of a phosphate group (approximately the same size as glycine) to the uncharged tRNAs and discriminates against charged tRNAs, and we end completely abolished binding (Fig. 1d). We conclude that the phosphate group (approximately the same size as glycine) to the uncharged tRNAs and discriminates against charged tRNAs, and we thus named the region the ‘T-box discriminator’.

Overall structure of the T-box discriminator–tRNA complex. Next, we elucidated how the T-box discriminator, merely 66 nts in size, senses tRNA aminoacylation without proteins. We co crystallized the G. kaustophilus glyQ T-box discriminator in complex with its cognate uncharged tRNA15 and solved the structure at 2.66-Å resolution (Table 1, Extended Data Fig. 3 and Supplementary Movie 1). The discriminator consists of three RNA helices (stem III, helix A1 and helix A2) and three single-stranded regions (both flanks of stem III and the T-box bulge; Fig. 1c,f). The combined stability of helices A1 and A2 dictates the transcriptional outcome, because the 3′ strands of both helices also form the 5′ strand of the terminator hairpin. This arrangement is similar to most small molecule–binding riboswitches, wherein the stability of a proximal ‘P1’ helix determines the genetic outcome15–16. The difference is that with conventional riboswitches, small-molecule metabolites directly bind and modulate P1 stability, whereas in T-boxes, it is the binding of a macromolecule—a tRNA—that controls the stability of the antiterminator. Our co crystall structure revealed Watson–Crick base pairing between the tRNA 3′-UCCA terminus and the 5′-AGGU tetranucleotide of the T-box bulge (Fig. 1b,e,f), consistent with previous covariation, genetic and biochemical analyses15,16,24. Crucially, this intermolecular helix stacks coaxially with the tRNA acceptor stem and helix A1 of the antiterminator (Fig. 1f), thus allowing the tRNA to stabilize the antiterminator, driving transcription readthrough. This functionally critical tRNA-antiterminator stacking is further supported by the nearly complete fluorescence quenching of a 2-aminopurine positioned at the tRNA 3′-end and of a pyrrolocytosine located on helix A1 on the T-box side14. Helix A2 of the discriminator is positioned at an ~80° angle from the tRNA-helix A1 RNA stack, such that its terminal base pair abuts the tRNA 3′-OH and is primarily responsible for sensing tRNA aminoacylation (Fig. 1f, red base pair). Finally, the base of stem III is attached to the proximal region (as opposed to the distal loop region) of helix A2, such that its flanking conserved purines closely track the extended minor groove and make key stabilizing interactions (Fig. 1e).

A tandem adenosine-minor groove ‘latch’ reinforces the tRNA-discriminator interface. The 5′ end of the T-box discriminator features two conserved adenosines (A128 and A129) and tracks the minor groove of the tRNA–antiterminator duplex (Fig. 2a–d). Both adenosines make hydrogen bonds across the groove, thus stabilizing the duplex (Fig. 2b,c). A128 and A129 are not only covalently linked but also stack with each other (Fig. 2a). Because each adenosine interacts with one side of the tRNA–T-box stacking interface, the A128-A129 stacking interaction acts as a latch or staple to reinforce the functionally crucial stacking (Fig. 2d). Thus, this pair of adenosines provide both lateral and axial stabilization to the tRNA-mRNA interface. Furthermore, A129 packs its nucleobase against the ribose of G161 to lend additional stability (Fig. 2a). A similar nucleobase-ribose packing interaction was observed between a cytosine in stem I (C64) and tRNA U20 near the elbow16. Substituting either A128 or A129 with a uridine drastically reduced tRNA-mediated readthrough (Fig. 2h and Extended Data Fig. 4).

In stark contrast, replacing four immediately preceding adenosines (A124–127) not involved in interactions with uridines had no effect (Fig. 2h). Tandem, stacked adenosines are frequently observed to laterally stabilize RNA helices and heteroduplexes across their minor grooves17,18. One prominent example is the highly conserved ribosomal A1492-A1493 dinucleotide in the A site, which, in conjunction with G530, latches the codon-anticodon duplex and triggers 30S closure to ensure decoding fidelity17,18 (Extended Data Fig. 5). The unusual placement of the A128-A129 latch across an important tRNA-mRNA interface locally bolsters RNA-RNA interactions to rheostat gene expression (Fig. 2d).

**Table 1 | Data collection and refinement statistics.**

|                     | SAD Phasing          | Model Complex (PDB 6PMO) | Native |
|---------------------|----------------------|--------------------------|--------|
| **Data collection** |                      |                          |        |
| Space group         | C222                  | C222                     | C222   |
| Cell dimensions     |                      |                          |        |
| a, b, c (Å)         | 95.01, 118.41, 134.09 | 44.85, 139.91, 135.99    | 45.05, 139.55, 134.97 |
| Wavelength (Å)      | 0.9218               | 1.0000                   | 0.9793 |
| Resolution (Å)*     | 74.10–4.53           | 35.0–2.66                 | 67.49–2.84 |
| Rmerge (%)†         | 21.3 (90)            | 6.2 (115)                 | 7.6 (80) |
| /σ(%)               | 24.0 (12.0)          | 16.9 (1.4)                | 11.9 (1.8) |
| CC1/2               | 0.995 (0.993)        | 0.998 (0.698)             | 1.000 (0.465) |
| CC*                 | 0.999 (0.998)        | 1.000 (0.907)             | 1.000 (0.797) |
| Completeness (%)‡   | 100.0 (100.0)        | 99.5 (97.6)               | 99.8 (100.0) |
| Redundancy≤          | 56.8 (58.5)          | 6.0 (5.9)                 | 6.1 (6.0) |
| **Refinement**      |                      |                          |        |
| Resolution (Å)*     | 35.0–2.66             | 69.78–2.84                | 2.94–2.84 |
| No. reflections‡    | 12724 (1215)         | 10437 (1032)              |        |
| Rmerge / Rwork (%)  | 22.7 (37.9)          | 25.3 (38.9)               | 271 (52.4) |
| No. atoms           | 2,797                | 2,880                     |        |
| RNA                 | 2,762                | 2,848                     |        |
| Ions                | 17                   | 32                        |        |
| Water               | 18                   | 0                         |        |
| Mean β factors (Å)  | 89.2                 | 98.3                      |        |
| RNA                 | 88.9                 | 97.5                      |        |
| Mg and Ir ions      | 176.6                | 170.1                     |        |
| Water               | 56                   | –                         |        |
| R.m.s. deviations   |                      |                           |        |
| Bond lengths (Å)    | 0.004                | 0.003                     |        |
| Bond angles (°)     | 0.77                 | 1.50                      |        |
| Maximum likelihood coordinate precision (Å) | 0.47 | 0.48 |        |

*Values in parentheses are for the highest-resolution shell. †Values in parentheses are for the cross-validation set. The SAD phasing dataset contains merged data from nine crystals of the wild-type T-box sequence. The single datasets of the model complex and native crystals were from crystals of the A156C::U189G variant T-box, which exhibited improved diffraction limits.
A long-range pseudohelix dictates the T-box response. Immediately 3′ to the tandem A-minor latch, the RNA backbone bends to pair U131 with C166 of the T-box bulge, forming an unusual cis pyrimidine-pyrimidine pair. G130 intercalates between A164 and the U131-C166 pair and pushes out C165, forming a long-range, parallel-stranded pseudohelix (Fig. 2e–g). To cap this three-layered pseudohelix, A164 flips its nucleobase inward in a rare (~4% occurrence in functional RNAs) energetically unfavorable syn conformation, whereas the regular anti conformation would disengage it from the pseudohelix (Fig. 2f). To offset the energetic penalty of the unusual syn conformation (~1.2 kcal mol⁻¹ at 35 °C)⁴⁰, numerous hydrogen bonds and stacking with G130 hold the spring-loaded A164 in place (Fig. 2e,f). This in turn allows the A164-C165 dinucleotide to act as a lateral latch (Fig. 2d) to stabilize the RNA–T-box duplex via five hydrogen bonds to its minor groove (Fig. 2f). Thus, the inward flipping of A164 and outward flipping of C165 allow the dinucleotide to make extensive nucleobase-specific contacts across the groove for stabilization. This explains the strict sequence conservation of the T-box bulge¹⁵,¹⁶ and highlights the frequent use of compact syn nucleotides in critical functional centers of diverse RNAs³⁹.

Because the pseudohelix stacks against helix A2 at its base (with the G167•U185 pair; Fig. 2e,g), the stability of the pseudohelix impacts that of its neighboring helix A2, which in turn

---

**Fig. 2 | The A-minor latch and pseudohelix stabilize tRNA-T-box interactions.** a, The A128-A129 latch (yellow) reinforces tRNA (green)-helix A1 (blue) stacking interface. b,c, Hydrogen-bonding patterns in a, d, Cartoon illustrating how the A128-A129 axial latch and A164-C165 lateral latch reinforce the tRNA–T-box duplex. e, A pseudohelix stacks with and modulates helix A2 stability. f, A164 and C165 interact across the minor groove of the T-box–tRNA duplex. g, Cartoon diagram of the pseudohelix structure (green shaded triangle). Stacking interactions are indicated by green connectors. h, In vitro transcription readthrough analysis of wild-type and mutant T-boxes (mean and s.d., n = 3 biologically independent samples). *A131U is where the B. subtilis (A) and G. kaustophilus (U) sequences naturally diverge.
controls the level of transcription readthrough. As a result, the purine content of the pseudohelix (that is, the number of purines within the region, which stack more stably than pyrimidines due to their larger, bicyclic nucleobases) strongly modulates the T-box response. This effect is so dominant that it overrides the effect of tRNA stabilization. Specifically, replacing A164 or G130 with a uridine abrogated tRNA-mediated readthrough, and substituting C165 or C166 with a guanosine rendered the T-box constitutively active (Fig. 2h and Extended Data Fig. 4). The latter likely resulted from overstabilization of helix A2 from excessive stacking between the purines in the pseudohelix, obviating the need for tRNA stabilization (Fig. 2g). Thus, the pseudohelix is a key regulatory element through which external signals such as regulatory proteins, metabolites, or antibiotics could bind and dictate the T-box response, consistent with the strong modulatory effects of ribosome inhibitors on T-boxes41–43.

Structural basis of tRNA aminoacylation sensing by the T-box discriminator. Our cocrystal structure reveals that the compact RNA domain accomplishes this by capturing and burying the tRNA 3′-end deep inside the discriminator core and juxtaposing it with a steric barrier intolerant of an aminoacyl group (Fig. 3 and Extended Data Fig. 6).

The tRNA 3′-end (tA76) is paired with U160, stacks against tC75 on the tRNA side and C186 of helix A1 on the T-box side and further engages two hydrogen bonds via its 2′-OH and phosphate oxygen (Figs. 2b and 3a). These interactions collectively bury 500 Å² of solvent-accessible interface (Extended Data Fig. 6) and position the tRNA 3′-OH merely 3.2 Å from a G167•U185 wobble pair at the base of helix A2 (Fig. 3a,b). This base pair acts as the primary steric barrier against 3′-aminoacyl tRNAs. Binding by a 3′-aminoacyl tRNA would place its universal amino moiety in direct steric
clash with the nucleobase of U185 (Fig. 3b). Interestingly, the G•U wobble shifts U185 outward into the major groove, ~2 Å closer to the tRNA 3'-OH than a Watson–Crick pair (Extended Data Fig. 7), potentially amplifying the steric conflict. As a result, the T-box discriminator has constructed a steric pocket even more encompassing and restrictive than the ribosome–RelA complex11,12, which positions the β5 strand of its TGS domain ~5 Å away from the tRNA 3′-OH to exert similar steric selection (Extended Data Fig. 7). Unlike the ribosome, which only receives 3′-aminoacyl tRNAs preselected by EF-Tu44, the T-box functions independently of EF-Tu14 and encounters both 2′- and 3′-aminoacyl tRNAs that undergo rapid regioisomerization (~5 s−1)45. Reassuringly, a 2′-aminoacyl group would clash with the tandem A-minor latch or the pseudohelix, both of which are essential for antiterminator stabilization (Fig. 3c,d). Indeed, interface analysis revealed that the 2′- and 3′-OH of tA76 are completely buried by the discriminator (Extended Data Fig. 6). Thus, a singular discriminator structure tightly clutches tA76 and sterically rejects both 2′- and 3′-aminoacyl tRNAs. We conclude that steric sensing of tRNA aminoacylation is a conserved mechanism to detect nutrient limitation and can be effectively implemented by either a protein or an RNA device. The G167•U185 wobble pair, which acts as the steric filter, occurs in 58% of the 390 transcription-regulating T-boxes analyzed. Other naturally occurring configurations include A-U (25%) and G-C (17%) but essentially no others. Curiously, all of the 39 translation-regulating T-boxes in Actinobacteria exclusively contain the
The G•U wobble pair starched. The clear preference for wobble pairs may stem from their unique geometric and energetic characteristics. The G•U wobble pair underwinds the helix by 14°, significantly enhancing nucleobase overlap and stacking with the penultimate C-G pair, reminiscent of the tRNA T-stem loop 49•U65 wobble pair (Extended Data Fig. 8). Despite the clear G•U preference, Watson–Crick pairs adequately support uncharged tRNA-mediated readthrough and rejection of tRNAs that carry a 3′ phosphate or an extra cytosine in vitro (Extended Data Fig. 3e). In vivo substitutions with Watson–Crick pairs, however, led to a two- to three-fold reduction in downstream gene expression, corroborating the enhanced stacking stability from the G•U wobble pair (Extended Data Fig. 4). Notably, the reversed U•G wobble pair caused a significant defect (Extended Data Fig. 3e). Because a U-A pair is nearly fully functional (Fig. 3e), the G•U wobble pair caused a significant defect from U-A pairs (Fig. 3e).

Perplexingly, their substitution with isosteric C-G/G-C pairs produced a dramatic 57-fold reduction of gene expression in vivo (Extended Data Fig. 8). Our structure revealed that the pairs provide a sequence-specific docking site for the base of stem III via three consecutive sets of purine-minor groove interactions from G133 and A153 and attendant nucleobase-ribose packing interactions (Fig. 4a–c). Consequently, substitutions of G133 or A153 strongly impacted transcription readthrough (Fig. 4f). Thus, instead of the previously proposed coaxial stacking with the base of helix A1 (ref. 30), the terminal pair of stem III and its flanking purines form an extended latch along the minor groove and act to fasten and stiffen the entire discriminator (Figs. 1e and 4a,b). Comparison with an NMR structure of an isolated antiterminator (discriminator without stem III and its flanking purines) revealed that drastic rearrangements are required to bind tRNA: the extensively stacked nucleotides in the free T-box bulge must unstack and unravel to accommodate the tRNA 3′ end (Fig. 4g). Thus, the compact T-box discriminator captures the universal tRNA 3′-NCCA end, employing a cooperative network of energetically coupled contacts that include four intermolecular base pairs, two instances of intermolecular coaxial stacking, six sets of purine-minor groove interactions and three sugar-nucleobase packing interactions (Fig. 1b, e). These multimodal RNA-RNA contacts create a stiff, snug pocket that fashions a steric barrier from a wobble pair to select for uncharged tRNAs.

Unexpected structural role of T-box stem III and flanking purines. Immediately adjacent to the central G167•U185 pair are two nearly invariant G-C/C-G pairs (Extended Data Fig. 1). A full-length T-box mRNA forms a molecular clamp vise. Our cocystal structure reveals the detailed interactions between the tRNA 3′ region and the discriminator, yet it remains unknown
how stem I and discriminator domains coordinate their binding to opposite flanks of tRNAs in a complete T-box. Multiple conflicting structural models have been proposed based on SAXS envelopes and modeling, and it was proposed that stem I must release the tRNA elbow contact in order for the antiterminator to bind the tRNA 3′ end.

Cryo-EM single-particle analysis (SPA) is now routinely used to resolve large protein and protein-stabilized ribonucleoprotein structures at near-atomic resolutions. However, despite enhanced contrast compared to proteins, imaging protein-free structured RNAs to resolutions that provide biochemical insights (better than 4–5 Å) remains a significant technical challenge. This is likely due to RNAs inherent structural flexibility, frequent conformational heterogeneity and the relatively small sizes of most known RNA domains. These challenges are exemplified by a recent cryo-EM analysis of a 30 kDa HIV-1 dimerization initiation site (DIS) RNA at 9 Å resolution, in which low signal-to-noise ratio and RNA flexibility or dynamics hampered further resolution improvement.

In order to distinguish among competing T-box structural models and extend the applicability of single-particle analysis to medium-sized RNA-only structures, we subjected a full-length B. subtilis glyQS T-box–tRNAαGly complex to cryo-EM SPA. The RNA complexes embedded in vitrified thin ice were unambiguously recognized with identified top and side views (Fig. 5a). Two-dimensional class averages revealed salient high-resolution features, and the final reconstruction achieved 4.9–Å overall resolution, despite its unusually small size and flexible structure (78 kDa; Fig. 5b, Table 2 and Extended Data Fig. 9). The central tRNA encased by the peripheral T-box was better resolved to ~4.0 Å resolution, according to the local-resolution map (Extended Data Fig. 10), in which we started to see resolved density of the ribose and phosphate backbone (Fig. 5d). The discriminator-tRNA and stem I-tRNA cocrystal structures were rigidly fitted into the cryo-EM map, resulting in a full-length T-box–tRNA model (Fig. 5d and Supplementary Video 2). The model was flexibly fitted and optimized to the cryo-EM map, with a resolution of 2.1 Å and 1.9 Å from each crystal structure. The resulting cryo-EM model captured all three principal tRNA-mRNA contacts observed in the two cocrystal structures, suggesting that all three contacts are engaged concurrently in the full complex in the readthrough configuration (Fig. 5d and Supplementary Video 2).

The full-length B. subtilis cryo-EM structure revealed a stem I trajectory distinct from and intermediate between those of the stem I-tRNA cocrystal structures from Oceanobacillus iheyensis and G. kaustophilus (Fig. 6a). To adapt to structural divergences among the three stem I domains, the same tRNAαGly flexes about its t26 hinge up to 22° to maintain nearly identical contacts with its anticodon and elbow (Fig. 5d). When the stem I distal interdigitated T-loops are superimposed, the tRNA elbow to which the T-loops bind overlay among the three structures (Fig. 6b), suggesting remarkable conservation of the interface across three species despite the lack of base-pairing interactions. Indeed, mutation or deletion of the interdigitated T-loops was reported to slow tRNA binding and to drastically reduce complex stability. Thus, the tripartite T-box interactions with the tRNA anticodon, elbow, and 3′ end persist concurrently in the same complex. We did not observe the previously proposed 90° pivoting motion that would disengage the stem I distal region from the tRNA elbow, which may occur in transient intermediates not detectable in our analysis. On the 3′ side, while the tRNA and discriminator each superimpose well with the cryo-EM structure (RMSDs of 1.5 and 1.6 Å, respectively), the tRNA body is rotated 8° between the cocrystal and cryo-EM structures, enabled by apparent flexibility of the tRNA acceptor stem (Fig. 6c). To confirm the overall architectures of both T-box complexes in solution, we performed SAXS analysis of both complexes and observed scattering profiles highly congruent with the cocrystal and cryo-EM structures (χ2 ~1.7 and 0.7, respectively; Fig. 6f,g).

The simultaneous binding of stem I and discriminator regions to both flanks of tRNA create a molecular vise that clamps the top half of the uncharged tRNA (Fig. 6d,e). Remarkably, contiguous coaxial stacking traverses three intermolecular interfaces, forming a central spine of 31 or 32 layers of stacked pairs (Fig. 6d,e). This spine provides the requisite stability—in conjunction with the kinetic advantage of being transcribed first—that preclude the formation of a transcription terminator or Shine–Dalgarno sequestra tor. In amino acid abundance, the presence of an esterified amino acid on the tRNA 3′ end interrupts the tRNA-helix A1 stacking, which severs the central spine and prevents the aminoacyl-tRNA from stabilizing the antiterminator. Without tRNA stabilization, the thermodynamically more stable terminator or sequestra tor forms and shuts off downstream amino acid genes, completing a negative feedback loop.

Discussion

Together, our cocrystal, cryo-EM, SAXS, and mutational analyses delineate how the T-box mRNA envelopes a cognate uncharged tRNA through sequential, tripartite interactions to construct a remarkably stable complex (kass ~0.0002 s−1) to permit conditional gene expression (Fig. 7). Our analyses distinguish among a number of proposed structural models and establish how mRNA encapsulates a tRNA, sterically probes its aminoaoylation status, and switches conformations based on this readout. Contoured by
**Fig. 6** The T-box central spine and structural comparisons. 

**a.** Superposition of the stem I distal interdigitated T-loops (boxed) in the cryo-EM structure (blue) with those in the stem I–tRNA cocrystal structures from *O. ihenensis* (orange)\(^2\) and *G. kaustophilus* (cyan)\(^2\). **b.** Comparison of the stem I-tRNA stacking interface among the three structures. Numbering is based on the *B. subtilis* cryo-EM structure (blue). **c.** Superposition of the discriminator structures in the cocrystal (magenta) and cryo-EM (blue and green) structures reveals an 8° rotation of the tRNA acceptor stem. **d.** Secondary structure diagram of the full-length *B. subtilis* glyQS T-box-tRNAGly\(^{22}\) complex. Disordered residues are shown in gray. **e.** Coaxial stacking at three tRNA-T-box interfaces align and assemble a ~31- or 32-layered central spine that stabilizes the antiterminator conformation. **f, g.** Overlay of back-calculated SAXS scattering curves (red lines) computed from the cocrystal structure of the T-box discriminator–tRNA complex (f) or cryo-EM structure of the full-length T-box–tRNA complex (g) with CRYSOL\(^{61}\) with experimental scattering profiles in solution (blue circles; mean and s.d. of 30 measurements of one sample; Methods). \(\chi^2\) values of the fits are indicated.
Fig. 7 | Mechanistic model of a cotranscriptionally acting T-box riboswitch. A transcriptional T-box initially uses its stem I to select a cognate tRNA regardless of its 3’ aminocacylation state, based primarily on anticodon-specifier complementarity. $k_a$ and $k_r$ for the cognate tRNA are derived from single-molecule fluorescence measurements\(^{20,41,42}\). The anticodon-specifier interaction engages first and is reinforced by the stacking interaction at the tRNA elbow (red stick). Subsequent elongation of the T-box transcript exposes stem III and the 5’ half of the antiterminator, which then forms four base pairs with the 3’-NCCA end of the docked tRNA. Finally, the 3’ strand of the antiterminator, transcribed last, will attempt to anneal with its 5’ counterpart. In starvation, the annealing is successful, forming a 32-layered, stacked central spine. This spine stabilizes the antiterminator conformer and allows RNAP to traverse this region to transcribe downstream genes. In amino acid abundance, the presence of an esterified amino acid on the tRNA blocks annealing of the 3’ strand of the antiterminator due to a steric conflict with the G167-U185 wobble pair, leading to terminator formation and termination of gene expression, completing a negative feedback loop.

Multivalent tRNA binding by the modular T-box mRNA may further extend to the enigmatic stem II domain absent from the glyQ T-boxes under study but present in most non-glycyl T-boxes\(^{15,16}\). Secondary structure analyses indicated an elongated stem II hairpin comprising likely also operate in several types of RNA and ribonucleoprotein condensates or granules, including those containing pathogenic repeat expansion RNAs\(^{56–58}\). In these systems, limited sequence complementarity (for example, between short CAG or CUG near-palindromic repeats) first seeds local RNA-RNA interactions near its elbow region, based on a cross-linking analysis\(^{26,27}\). However, detailed structural analysis is required to clarify the structure and function of the stem II domain.

Similar distributive binding mechanisms and extensive utilization of stacking and tertiary contacts to supplement limited sequence complementarity likely also operate in several types of RNA and ribonucleoprotein condensates or granules, including those containing pathogenic repeat expansion RNAs. In these systems, limited sequence complementarity (for example, between short CAG or CUG near-palindromic repeats) first seeds local RNA-RNA interactions. Subsequently, multivalency drives their self-assembly into RNA networks or foci. Similarly, the type of multivalent interactions employing intra- and intermolecular stacking and extensive tertiary contacts found in the T-box–tRNA complexes may also
drive the assembly of long noncoding RNA domains or viral RNA genomes.19

Online content
Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information, details of author contributions and competing interests, and statements of data and code availability are available at https://doi.org/10.1038/s41594-019-0326-7.

Received: 24 June 2019; Accepted: 7 October 2019; Published online: 18 November 2019

References
1. Chantranupong, L., Wolfson, R. L. & Sabatini, D. M. Nutrient-sensing mechanisms across evolution. Cell 161, 67–83 (2015).
2. Saxton, R. A. et al. Structural basis for leucine sensing by the Sestrin2-mTORC1 pathway. Science 351, 53–58 (2016).
3. Chantranupong, L. et al. The CASTOR proteins are arginine sensors for the mTORC1 Pathway. Cell 165, 153–164 (2016).
4. Antson, A. A. et al. Structure of the trp RNA-binding attenuation protein, TRAP, bound to RNA. Nature 401, 235–242 (1999).
5. Serguson, A. & Patel, D. J. Amino acid recognition and gene regulation by riboswitches. Biochim. Biophys. Acta 1799, 592–611 (2009).
6. Roth, A. & Breaker, R. R. The structural and functional diversity of metabolite-binding riboswitches. Annu. Rev. Biochem. 78, 305–334 (2009).
7. Pang, Y. L., Porur, K. & Martinis, S. A. tRNA synthetase: tRNA aminoaoylation and beyond. Wiley Interdiscip. Rev. RNA 5, 461–480 (2014).
8. Bullwinkle, T. J. & Ibaa, M. Emergence and evolution. Top Curr. Chem. 344, 43–87 (2014).
9. Bunning, P. J. & Musier-Forsyth, K. Transfer RNA recognition by aminoacyl-RNA synthetases. Biopolymers 52, 1–28 (1999).
10. Yanofsky, C. Attenuation in the control of expression of bacterial operons. Nature 299, 751–758 (1981).
11. Loveland, A. B. et al. Ribosome*ReA structures reveal the mechanism of stringent response activation. elife 5, e17029 (2016).
12. Brown, A., Fernandez, I. S., Gordiyenko, Y. & Ramakrishnan, V. Ribosome-dependent activation of stringent control. Nature 534, 277–280 (2016).
13. Dong, J., Qiu, H., Garcia-Barrio, M., Anderson, J. & Hinnebusch, A. G. Uncharged tRNA activates GCN2 by displacing the protein kinase moiety from a bipartite tRNA-binding domain. Mol. Cell 6, 269–279 (2000).
14. Zhang, J. & Ferré-D’Amare, A. R. Direct evaluation of tRNA aminoaoylation status by the T-box riboswitch using tRNA-mRNA stacking and steric readout. Mol. Cell 55, 148–155 (2014).
15. Viteschak, A. G., Mironov, A. A., Lyubetsky, V. A. & Gelfand, M. S. Comparative genomic analysis of T-box regulatory systems in bacteria. RNA 14, 717–735 (2008).
16. Gutierrez-Preciado, A., Henkin, T. M., Grundy, F. J., Yanofsky, C. & Merino, E. Biochemical features and functional implications of the RNA-based T-box regulatory mechanism. Microbiol. Mol. Biol. Rev. 73, 63–61 (2009).
17. Saad, N. Y. et al. Two-codon T-box riboswitch binding two tRNAs. Proc. Natl Acad. Sci. USA 110, 12756–12761 (2013).
18. Apostolidi, M. et al. A g(C) T-box riboswitch with species-specific structural features responding to both proteinogenic and nonproteinogenic tRNA_Gly. RNA 30, 967–983 (2014).
19. Liu, J., Qiu, H., Garcia-Barrio, M., Anderson, J. & Hinnebusch, A. G. Structural determinants for geometry and information decoding of tRNA by T-box leader RNA. Structure 21, 2025–2032 (2013).
20. Grundy, F. J. & Henkin, T. M. RNA as a positive regulator of transcription antitermination in vitro. Cell 74, 475–482 (1993).
21. Grundy, F. J., Winkler, W. C. & Henkin, T. M. tRNA-mediated transcription antitermination in vitro: codon-anticodon pairing independent of the ribosome. Proc. Natl Acad. Sci. USA 99, 11121–11126 (2002).
55. Klein, D. J., Schmeing, T. M., Moore, P. B. & Steitz, T. A. The kink-turn: a new RNA secondary structure motif. *EMBO J.* **20**, 4214–4221 (2001).

56. Van Treeck, B. & Parker, R. Emerging roles for intermolecular RNA-RNA interactions in RNP assemblies. *Cell* **174**, 791–802 (2018).

57. Langdon, E. M. et al. mRNA structure determines specificity of a polyQ-driven phase separation. *Science* **360**, 922–927 (2018).

58. Jain, A. & Vale, R. D. RNA phase transitions in repeat expansion disorders. *Nature* **546**, 243–247 (2017).

59. Dethoff, E. A. et al. Pervasive tertiary structure in the dengue virus RNA genome. *Proc. Natl Acad. Sci. USA* **115**, 11513–11518 (2018).

60. Leontis, N. B. & Westhof, E. Geometric nomenclature and classification of RNA base pairs. *RNA* **7**, 499–512 (2001).

61. Franke, D. et al. ATSAS 2.8: a comprehensive data analysis suite for small-angle scattering from macromolecular solutions. *J. Appl. Crystallogr.* **50**, 1212–1225 (2017).

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

This is a U.S. government work and not under copyright protection in the U.S.; foreign copyright protection may apply. 2019
Methods

Conservation analysis of the T-box discriminator region. T-box sequences were retrieved from the Rfam database. The search motif used to localize the T-box discriminator in each sequence was 5′-UGGNANC-3′, where N denotes any nucleotide (Extended Data Fig. 1). Sequences of 49 residues were aligned, and duplicates were removed. All remaining sequences (N = 799) were folded with RNAfold while forcing the three residues upstream and the three residues downstream of the 5′-UGGNANC-3′ motif to base pair with downstream residues and prohibiting the residues within the T-box motif and the last eight residues of the sequences from base pairing. Only the lowest free-energy folds that satisfy these requirements were selected (N = 390). Because the helix A2 is not conserved, sequence alignment had to be established separately on either side of helix A2. Aligned sequences were used to generate sequence logos with Weblogo (Extended Data Fig. 1). To determine a conservative conservation pattern on the 5′ side of highly variable stems III, only sequences with the highly conserved adenine doublet (AA) in tRNA152-153 (six residues upstream of the 5′-UGGNANC-3′ motif) were selected from the Rfam database. The size of these sequences was limited to 38 nucleotides to limit folding artifacts with the upstream segment. Duplicates were then removed. The remaining sequences (N = 373) were folded with RNAfold while enforcing base pairing to 4 residues upstream of the adenine doublet. After the first round of folding, additional constraints were applied to select folded sequences in order to prevent alternate folds from occurring on the 5′ side of the sequence. Sequences that did not meet these requirements were discarded. To determine the base conservation pattern, all folded sequences were subsequently aligned to the 5′ end of stem III (whose stem length ranges from about 25 to about 20 base pairs). Because the overall conservation pattern shows that the functionally important G130 is not always a guanosine (upper panel, Extended Data Fig. 1), the conservation pattern was split for the sequences bearing a G at that position (middle panel, Extended Data Fig. 1; n = 205 sequences) and for the sequences that do not have a G (lower panel, Extended Data Fig. 1; n = 46 sequences).

RNA preparations. B. subtilis tRNA Sprut was used for in vitro termination-readthrough assays was transcribed in vitro by T7 RNA polymerase via ‘GMP priming’, where a 20 mM GMP and 2 mM ATP mixture was used to install predominantly (~91%) 5′ monophosphate. The RNAs were purified by electrophoresis on 10% polyacrylamide (29:1 acrylamide/bisacrylamide), 8 M urea TBE (Tris-borate-EDTA) gels, electroeluted, washed once with 1 M KCl, desalted via ultrafiltration, and stored at −80 °C before use. tRNA variants carrying 3′-azido- and 3′-N-pentenoyl-glycine modifications were prepared using the Klenow fragment and flexizyme. The 3′-phosphate tRNA was generated by in vitro transcription with T7 RNA polymerase, then cleaved by a transacting Varkud satellite ribozyme. The cleaved RNA was treated with 10 mM HCl at room temperature for 30 min to open the 2′,3′-cyclic phosphate to yield a 3′ phosphate. The HCl-treated tRNA was washed three times with diethyl pyrocarbonate (DEPC)-treated H2O and stored at −80 °C before use. To prepare the T-box discriminator-tRNA complex for co-crystallization and SAXS analysis, wild-type crystal (crystal 1–9; Table 1), or A156C or A156G mutant (crystal 10–11; stabilized helix A1) G. kaustophilus tRNA was co-transcribed in vitro by T7 RNA polymerase using PCR products as templates as described. The complex was isolated from excess tRNAs by size exclusion chromatography in a buffer consisting of 10 mM Tris-HCl, pH 7.4, 100 mM KCl and 20 mM MgCl2. Fractions containing the complex were pooled and concentrated to 8 mg/ml.

To prepare the full-length T-box riboswitch–tRNA complex for cryo-EM and SAXS analyses, B. subtilis tRNA Sprut control and B. subtilis tRNA Sprut were co-transcribed in vitro by T7 RNA polymerase, using PCR products as templates. The RNA complex was purified first using anion exchange chromatography by Mono Q column, then by size-exclusion chromatography on a Superdex 200 Increase column in a buffer consisting of 10 mM Tris-HCl (pH 7.4), 100 mM KCl and 20 mM MgCl2. Peak fractions containing the T-box–tRNA complex were pooled and concentrated to 8 mg/ml.

Electrophoretic mobility shift assay. G. kaustophilus GlyT3 T-box discriminator RNA (residues 126–191) was co-transcribed with excess B. subtilis (or G. kaustophilus, identical sequences) tRNA Sprut in vitro by T7 RNA polymerase using PCR products as templates. The complex was isolated from excess tRNAs by size exclusion chromatography in a buffer consisting of 10 mM Tris-HCl, pH 7.4, 100 mM KCl and 20 mM MgCl2. Fractions containing the complex were pooled and concentrated to 8 mg/ml.

Isothermal titration calorimetry. ITC was performed essentially as described in ref. 1. To suppress oligomerization, tRNA Sprut was heated to 90 °C in water for 3 min and snap-cooled to 0 °C over 2 min in microwells. For ITC experiments were equilibrated in a buffer comprised of 100 mM KCl, 25 mM Tris-HCl, pH 7.4, 10 mM MgCl2 via ultrafiltration. ITC experiments were performed in triplicates at 20 °C with 25 mM T-box discriminator RNA in the cell and 250 μM uncharged or 3′-phosphate tRNA Sprut in the syringe, using a MicroCal iTC200 microcalorimeter (GE). Data processing was carried out using unbiased integration software NITPIC2, then by SEDPHAT. The titration shown in Fig. 1a was performed three times using biologically independent samples. For both 3′-phosphate- and uncharged tRNA Sprut (M = 0.66 μM, ΔH = −2016 ± 60 kJ mol−1, −ΔG = 15.1 ± 0.8 kcal mol−1). These values are mean, and the errors are s.d.; n = 3 biologically independent samples. The average binding-incompetent fraction was ~10%. For three independent titrations using 3′-phosphate tRNA Sprut, the curves cannot be reliably fit, with an estimated Kd > 50 μM and negligible ΔH (Fig. 1d).

In vitro transcription termination-readthrough assay using a fluorescent reporter. Transcription termination-readthrough assay was performed essentially as described previously, with modifications. A hairpin molecular beacon was designed to specifically hybridize to the B. subtilis glyQ T-box downstream transcript to measure T-box-controlled gene expression (Fig. 1a). It is of the sequence 5′-CSC-GGAC-3′ (crystal 10–11; stabilized helix A1) where upper- and lowercase letters denote deoxyribonucleotides and 2′-O-methyl ribonucleotides, respectively. The 2′-O-methyl modifications were used to enhance hybridization with the target transcript. Multi-round transcription termination-readthrough assays were performed in a transcription buffer composed of 20 mM Tris-HCl, pH 7.4, 4 mM MgCl2, 1 mM spermidine, and 5 mM spermidine. Each 36 μl reaction also contained 10 mM PCR-derived linear DNA template, 0.05 U/μl E. coli RNA polymerase holozyme (New England Biolabs), 50 mM molecular beacon, and 250 nM snap-copolyed RNA, when specified. Reactions were initiated by addition of nucleotide triphosphates (NTPs) to provide 10 mM ATP, 10 mM GTP, and 50 mM GTP. The reaction was followed by fluorescence measurements, selected reactions were analyzed on a 10% urea-PAGE and visualized with SYBR Gold to validate fluorescence-based measurements (Extended Data Fig. 4). Due to hydrolytic instability of the 3′-aminoacyl tRNA, slightly alkaline conditions (pH 7.4 or higher) in transcription readthrough assays, which exacerbates tRNA deacylation, and extended incubation time (3 h or longer), it was not practical to use 3′-aminoacyl RNA. Instead, a 3′-phosphate tRNA was used as a stable mimic of 3′-aminoacyl tRNA, as the 3′-phosphate is comparable in size to 3′-glycyl and, tRNAs bearing the two modifications were found to elicit nearly identical transcriptional responses. We also used another negative control tRNA termed the ExIC tRNA, in which an extra cytosine was appended to the tRNA 3′-CCA end. This tRNA strongly inhibits transcription, as c77 competes with C186 for pairing with G159, destabilizing helix A1 and preventing tRNA-helix A1 stacking.
visible tRNA and T-box residues were built, the partial model was located in a high-resolution Ir-SAD dataset collected from cocrystals of the A156C-U189G mutant discriminator–tRNA complex (2.66 Å resolution) using Phaser©, which allowed identification and completion of the tRNA structure. The S1199G substitution replaced a non-conserved A-P pair in helix A1 with a C-G pair, which resulted in altered unit cell parameters and extended the diffraction limit from 3.4 Å to 2.66 Å (Table 1). Model building was performed in Coot©, as guided by iteratively generated MR-SAD maps using an evolving model, corrected by ERRASER©, and refined using Phenix.Refine®. The refined structure without Ir atoms was located in a native 2.84 Å data set using Phaser (Translational Function Z (TFZ) score = 36.9; log-likelihood gain (LLG) = 2,681) and was also refined. The Ir-containing (complex model) and native structures are nearly identical, with a r.m.s.d. of 0.2 Å. Refinement statistics are summarized in Table 1.

**Reporting Summary.** Further information on research design is available in the Nature Research Reporting Summary linked to this article.

**Data availability**

Atomic coordinates and structure factor amplitudes for the T-box discriminator in complex with tRNA© have been deposited at the Protein Data Bank (PDB) under accession code 6P6O. Cryo-EM structure of the full-length T-box–tRNA complex and map have been deposited to Electron Microscopy Data Bank under EM-DATA-201416 and PDB 6P6O. All other data are available upon reasonable request.

**References**

62. Kalvari, I. et al. Rham 13:0 to a genome-centric resource for assigning riboswitch families. *Nucl. Acids Res.* 46, D335–D342 (2018).

63. Lorenz, R. et al. ViennaRNA Package 2.0. *Algorithms Mol. Biol.* 6, 26 (2011).

64. Crooks, G. E., Hon, G., Chandonia, J. M. & Brenner, S. E. WebLogo: a sequence logo generator. *Genome Res.* 14, 1188–1190 (2004).

65. Sampson, J. R. & Uhlenbeck, O. C. Biochemical and physical characterization of an unmodified yeast phenylalanine transfer RNA transcribed in vitro. *Proc. Natl. Acad. Sci. USA* 85, 1033–1037 (1988).

66. Zhang, J. & Ferre-D’Amare, A. R. A flexible, scalable method for preparation of homogeneous aminocylated tRNAs. *Methods Enzymol.* 549, 105–113 (2014).

67. Ferré-D’Amare, A. R. & Doudna, J. A. Use of cis- and trans-ribozymes to remove 5’ and 3’ heterogeneities from milligrams of in vitro-transcribed RNA. *Nucleic Acids Res.* 24, 977–978 (1996).

68. Xiao, H., Murakami, H., Saga, H. & Ferré-D’Amare, A. R. Structural basis of specific tRNA aminoclylation by a small in vitro selected ribozyme. *Nature* 454, 358–361 (2008).

69. Wood, I. V. et al. Crystal structure of an adenovirus virus-associated RNA. *Nat. Commun.* 10, 2871 (2019).

70. Keller, S. et al. High-precision isothermal titration calorimetry with automated peak-shape analysis. *Anal. Chem.* 84, 5066–5073 (2012).

71. Zhao, H., Piszczek, G. & Schuck, P. SEDPHAT–a platform for global ITC analysis and global multi-method analysis of molecular interactions. *Methods* 77, 137–148 (2015).

72. Zeng, C., Zhou, S., Bergmeier, S. C. & Hines, J. V. Factors that influence T box riboswitch efficacy and tRNA affinity. *Biorg. Med. Chem.* 23, 5702–5708 (2015).

73. Yousef, M. R., Grundy, F. J. & Henkin, T. M. tRNA requirements for glyQ antitermination: a new twist on tRNA. *RNA* 9, 1148–1156 (2003).

74. Forster, A. et al. Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography. *Acta Crystallogr.* D 69, 1617–1632 (2013).

75. Adams, P. et al. PHENIX: a comprehensive python-based system for macromolecular structure solution. *Acta Crystallogr.* D 66, 213–221 (2010).

76. McCoy, A. et al. Phaser crystallographic software. *J. Appl. Cryst.* 40, 658–674 (2007).

77. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of coot. *Acta Crystallogr.* D 66, 486–501 (2010).

78. Chou, F. C., Sripakdeeveong, P., Dibrov, S. M., Hermann, T. & Das, R. Correction of angle errors in RNA crystallography through rigorous structure prediction. *Nat. Methods* 10, 74–76 (2013).

79. Chen, M. et al. Convolutional neural networks for automated annotation of cellular cryo-electron tomograms. *Nat. Methods* 14, 983–985 (2017).

80. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Mol. Biol.* 410, 80–93 (2011).

81. Pettersen, E. E. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612 (2004).

82. Afonine, P. V. et al. Real-space refinement in PHENIX for cryo-EM and crystallography. *Acta Crystallogr.* D 74, 531–544 (2018).
83. Wang, J. et al. A method for helical RNA global structure determination in solution using small-angle x-ray scattering and NMR measurements. J. Mol. Biol. 393, 717–734 (2009).

84. Fang, X. et al. An unusual topological structure of the HIV-1 Rev response element. Cell 155, 594–605 (2013).

Acknowledgements
We thank I. Botos for computational support, J.R. Hogg, S. Ranganathan, G. Piszczek, D. Wu, J.C. Lee and M. Watson for support in fluorescence analyses; Y. He and N. Tjandra for fermentation support; R. Levine and D.-Y. Lee for MS support; W. Zhang and J.W. Soostak for a gift of Ir(III) Hexammine; M. Apostolidi for β-galactosidase assay protocols; CAS-Shanghai Science Research Center High-End User Project for preliminary cryo-EM data collection; and S.K. Buchanan, A. Ferré-D’Amaré, N. Baird, M. Lau, K. Saddula, C. Bou Nader, and J. M. Gordon for insightful discussions. This work was supported by the intramural research programs of NIDDK and NCI, an NIH DDIR Innovation Award to J.Z. and Y.-X.W., NIH U54GM103297 and U54AI150470 (Center for HIV RNA Studies, CRNA), P41GM103832, R01GM079429 and S10OD021600 to W.C., and Fonation Santé E515 and project “INSPIRED” (MIS/5002550) implemented under the Action “Reinforcement of the Research and Innovation Infrastructure”, funded by the Operational Programme “Competitiveness, Entrepreneurship and Innovation” (NSRF 2014-2020) and co-financed by Greece and the European Union (European Regional Development Fund), and a Fulbright Scholarship to C.S. V. S. is supported by an IKY-Siemens 2017 Excellence Postdoctoral Fellowship.

Author contributions
S.L. and J.Z. designed experiments. S.L. prepared RNA samples, cocrysalts, and samples for cryo-EM and SAXS, with the help of F.E.H, and performed in vitro assays. S.L. and J.Z. collected X-ray diffraction data, solved and refined the crystal structure, and analyzed SAXS data. Z.S. collected cryo-EM data, Z.S., G.D.P., K.Z., M.C., S.J.L. and W.C. performed cryo-EM data processing and modeling. J.L. performed phylogenetic analyses. V.S., N.G. and C.S. carried out in vivo experiments. L.F. and Y.-X.W. collected and processed SAXS data. All authors contributed to the preparation of the manuscript.

Competing interests
Authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41594-019-0326-7.

Correspondence and requests for materials should be addressed to W.C. or J.Z.

Peer review information Anke Sparmann was the primary editor on this article and managed its editorial process and peer review in collaboration with the rest of the editorial team.

Reprints and permissions information is available at www.nature.com/reprints.
Extended Data Fig. 1 | Secondary structures and conservation analyses of glycyl T-box riboswitches. a, b, Secondary structures of *G. kaustophilus* glyQ and *B. subtilis* glyQS T-box riboswitch. Glycine-specific T-boxes lack the stem II and stem IIA/B pseudoknot structures. Conserved nucleotides are highlighted, based on previous reports supplemented by new phylogenetic analysis (Fig. 1; Methods). Previous sequence annotations of the *G. kaustophilus* glyQ T-box had omitted a 5′ ssRNA leader that precedes stem I in all validated T-boxes, which is now restored. The probable transcription start site, 17 nts upstream of stem I, was identified using prokaryotic promoter prediction algorithms. Nucleotide numbering is thus offset by +17 relative to previous reports.

c, Sequence conservation of the T-box discriminator region based on *G. kaustophilus* glyQ T-box. The split patterns show that the intercalating G130 is at the center of a 5′-AR(U/A)-3′ motif (middle). This motif is shifted 1 nt to the left when there is no G in position 130 (bottom). In this case, a moderately conserved G is predominant in position 129, and two pyrimidines are present in positions 130 and 131. Assuming that G129 is the intercalating nucleotide equivalent to G130 in the middle panel, one of these pyrimidines (nt 130 or 131) may adopt an extrahelical conformation to account for the motif shift.
Extended Data Fig. 2 | Mutational analysis of T-box discriminator-tRNA interactions. a, Secondary structures of wild-type, mutant, and truncated T-box discriminators. Deletions are indicated by red boxes. b, Electrophoretic mobility shift assay (EMSA) analysis of the constructs shown in a, showing the requirement of stem III and flanking purines for tRNA binding. The antiterminator (discriminator without stem III and its flanking purines; Δ3 mutant) is prone to dimerization. c, tRNA variants used that carry various 3′ chemical modifications. Only the terminal tA76 is shown. d, EMSA analysis of constructs in c, showing that binding is selective for uncharged tRNA. e, Quantitation of d and comparison with previously reported in vitro transcription readthrough data of the same tRNA variants. The values and error bars represent mean and s.d., n = 3 biologically independent samples.
Extended Data Fig. 3 | Representative X-ray crystallographic electron density maps. a, Composite simulated anneal-omit 2F_0 − F_2 electron density calculated using the final model (1.0 s.d.) superimposed with the final refined model. b–d, Portions of the map showing tRNA-T-box discriminator coaxial stacking (b), encapsulation of tRNA 3′-end by the discriminator (c), and long-range interactions between stem III 5′ flanking purines and the T-box bulge (d). Note the density fusion as a result of nucleobase-ribose packing interactions between A129 and G161.
Extended Data Fig. 4 | In vitro transcription termination-readthrough assay and in vivo β-gal assay. a, Representative raw data of in vitro transcription termination-readthrough assay using wild-type *B. subtilis* T-box riboswitch. The rates of fluorescence increase between 34 and 180 min (segments with trendlines) report the production of readthrough transcripts. b, Quantitation of data in a. Rates of fluorescence increase (slopes) were subsequently normalized to that of the reference in the presence of NTP but absence of tRNA (green data points) and reported in Figs. 2h, 3e and 4f. c, Validation of fluorescence-based readthrough assay in a and b with subsequent, conventional gel-based analysis of the same samples. Addition of the uncharged tRNA led to significantly increased transcription readthrough. d, Scheme of in vivo gene expression assay using the *G. kaustophilus glyQ* T-box riboswitch transcriptionally fused with *lacZ*. e, Relative β-gal activity of wild-type and mutant T-boxes under glycine-replete and glycine-starvation conditions, normalized to wild-type T-box-containing strain grown in minimal media supplemented with glycine. The values and error bars represent mean and s.d., *n* = 3 biologically independent samples.
Extended Data Fig. 5 | Comparison of the T-box tandem A-minor latch with the A1492-A1493-G530 latch in the ribosome A site. a, The A128-A129 latch reinforces the functionally important tRNA (green)-helix A1 (blue) stacking interface. A128 and A129 form a continuous adenosine stack. b and c, The stacked A128 and A129 engage extensive hydrogen bonds with the minor groove, reinforce tRNA-T-box base-pairing and stacking, and “staple” the two RNAs together. d, In the ribosome A site, A1492 and A1493 similarly reinforce the intermolecular codon-anticodon duplex via tandem, stacked A-minor interactions in conjunction with G530. e and f, Hydrogen-bond patterns in the ribosome A site resemble those in the T-box (b and c).
Extended Data Fig. 6 | Intermolecular interface of the T-box discriminator-tRNA complex. 
a. Solvent-accessible surface colored according to area buried from light blue or white (no burial) to red (>25 Å² per atom). b. Open-book view of the binding interface. The lower inset shows the extensive burial of tRNA tA76, particularly its Watson–Crick edge (N6-N1-C2) and both 2′-OH and 3′-OH. c, d. Solvent-accessible surface area buried per residue for tRNA (c) and discriminator (d).
Extended Data Fig. 7 | Effect of the G•U wobble pair on tRNA aminoacylation sensing and comparison of steric sieves in the T-box and ribosome-RelA complex. a, A modeled tRNA 3′-glycyl moiety strongly clashes with the U185 nucleobase of the G•U wobble pair. b, Modeled Watson–Crick pair (C185, white) still clashes with the tRNA 3′-glycyl moiety, albeit to a lesser extent than the G•U wobble pair (a). c,d, Comparison of the steric sieves in the T-box (c) and RelA–ribosome complex (d). Solid green lines indicate inter-atomic distances in Å. The RelA–ribosome complex structure is based on PDB 5IQR.
Extended Data Fig. 8 | A conserved G•U wobble pair enhances stacking with its neighboring base pair both in the T-box discriminator and in the tRNA T-loop. a, Through local helix underwinding, helix A2 terminal G•U wobble pair produces exceptionally large nucleobase overlap areas and enhances stacking with the penultimate C-G pair. b, Reduced nucleobase overlap areas between a modeled G-C pair and the penultimate C-G pair. c, The G•U wobble pair is reminiscent of the conserved G49•U65 wobble pair found in the tRNA\(^{23S}\) T-loop in the same complex. d, For comparison, the penultimate C-G pair stacks with its neighboring G-C pair with less than half of the total overlap area (5.5 Å\(^2\) versus 13.9 Å\(^2\)). Overlap areas (in Å\(^2\)) between stacked nucleobases were calculated with 3DNA.
Extended Data Fig. 9 | Cryo-EM single particle analysis (SPA) workflow of full-length *B. subtilis* T-box−tRNA complex. **a**, 3D classification yielded two major classes (black boxes) that were combined for auto refinement. **b**, Final reconstruction. **c**, FSC curve showing 4.9-Å resolution at 0.143 cut-off. **d**, Euler angle distribution of the final reconstruction.
Extended Data Fig. 10 | Relion 3D classification and local resolution of the full-length *B. subtilis* T-box—tRNA complex. 3D classification of the complex converged to three maps. Superposition of the tRNA density in three maps revealed motions of the T-box relative to tRNA as indicated by the arrows. This flexibility of the T-box RNA is the major limitation that prevented cryo-EM reconstruction from achieving higher resolutions. Resmap analysis shows that the tRNA was better resolved at ~4-Å resolution (upper right).
Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

- Confirm the exact sample size (n) for each experimental group/condition, given as a discrete number and unit of measurement.
- A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly.
- The statistical test(s) used AND whether they are one- or two-sided. Only common tests should be described solely by name; describe more complex techniques in the Methods section.
- A description of all covariates tested.
- A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons.
- A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) and variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals).
- For null hypothesis testing, the test statistic (e.g. F, t, r) with confidence intervals, effect sizes, degrees of freedom and P value noted. Give P values as exact values whenever suitable.
- For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings.
- For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes.
- Estimates of effect sizes (e.g. Cohen’s d, Pearson’s r), indicating how they were calculated.

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection

In-house data collection software was provided by SER-CAT beamline ID-22 at the Advanced Photon Source (APS), Argonne National Laboratory (ANL), and by SLAC National Accelerator Laboratory.

Data analysis

The X-ray diffraction data were indexed, integrated, and scaled by XDS. Multi-crystal datasets were analyzed and merged with BLENDeR. The substructure was identified by HKL2map, ShelDX, and Phenix.Hys. De novo phasing was performed using Phenix.AutoSol. Model building was performed using Coot. Refinement was performed using Phenix.REFINE and ERRASER. Cryo-EM data were processed using MotionCor2, corrected by CTFFIND4. Particle picking and classifications were performed using EMAN2.2 and refinement was performed using Relion2. Modeling building was performed with Coot and refined with Phenix. SAXS data were analyzed using the ATSAS suite. No custom algorithms or softwares were used.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

Atomic coordinates and structure factor amplitudes for the T-box discriminator in complex with tRNA-Gly has been deposited at the Protein Data Bank under the accession code 6PML. Cryo-EM structure of the full-length T-box-trNA complex and map have been deposited to Electron Microscopy Data Bank under the accession code EMD-20416, and PDB code 6POM. All other data generated or analyzed during this study are included in this article (and its supplementary information files).
Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☑️ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size

One RNA crystal sample per X-ray dataset; one solution sample for each Cryo-EM dataset and SAXS dataset; In vitro and in vivo assays were performed in independent triplicates using distinct samples

Data exclusions n/a

Replication n/a

Randomization n/a

Blinding n/a

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a □ Antibodies □ Eukaryotic cell lines □ Palaeontology □ Animals and other organisms □ Human research participants □ Clinical data

Methods

n/a □ Involved in the study □ ChIP-seq □ Flow cytometry □ MRI-based neuroimaging