

Two-codon T-box riboswitch binding two tRNAs

Nizar Y. Saad^{a,b}, Vassiliki Stamatopoulou^{c,1}, Mélanie Brayé^b, Denis Drinas^c, Constantinos Stathopoulos^{c,2}, and Hubert Dominique Becker^{a,2}

^aUnité Mixte de Recherche 7156 Génétique Moléculaire, Génomique, Microbiologie, Centre National de la Recherche Scientifique, Université de Strasbourg, 67084 Strasbourg, France; ^bUnité Propre de Recherche Architecture et Réactivité de l'ARN, Centre National de la Recherche Scientifique, Institut de Biologie Moléculaire et Cellulaire, Université de Strasbourg, F-67084 Strasbourg, France; and ^cDepartment of Biochemistry, School of Medicine, University of Patras, 26504 Patras, Greece

Edited by Paul Schimmel, Skaggs Institute for Chemical Biology, La Jolla, CA, and approved June 13, 2013 (received for review March 5, 2013)

T-box riboswitches control transcription of downstream genes through the tRNA-binding formation of terminator or antiterminator structures. Previously reported T-boxes were described as single-specificity riboswitches that can bind specific tRNA anticodons through codon–anticodon interactions with the nucleotide triplet of their specifier loop (SL). However, the possibility that T-boxes might exhibit specificity beyond a single tRNA had been overlooked. In *Clostridium acetobutylicum*, the T-box that regulates the operon for the essential tRNA-dependent transamidation pathway harbors a SL with two potential overlapping codon positions for tRNA^{Asn} and tRNA^{Glu}. To test its specificity, we performed extensive mutagenic, biochemical, and chemical probing analyses. Surprisingly, both tRNAs can efficiently bind the SL in vitro and in vivo. The dual specificity of the T-box is allowed by a single base shift on the SL from one overlapping codon to the next. This feature allows the riboswitch to sense two tRNAs and balance the biosynthesis of two amino acids. Detailed genomic comparisons support our observations and suggest that “flexible” T-box riboswitches are widespread among bacteria, and, moreover, their specificity is dictated by the metabolic interconnection of the pathways under control. Taken together, our results support the notion of a genome-dependent codon ambiguity of the SLs. Furthermore, the existence of two overlapping codons imposes a unique example of tRNA-dependent regulation at the transcriptional level.

tRNA specificity | antitermination | metabolic networks

T-box riboswitches are important regulators of gene transcription found mainly in Gram-positive bacteria (1, 2). They change conformation upon binding a cognate tRNA molecule, thus controlling the transcription of downstream genes (3–7). Transcriptional read-through is allowed when the 3'-end of the nascent T-box transcript switches from the terminator to the antiterminator conformation upon binding of the cognate tRNA. Transcription is therefore directed by the concentration-dependent binding of charged and uncharged tRNAs to the T-box, with the level of charging serving as a sensor of the intracellular concentration of the cognate amino acid. By using this feedback mechanism, they control bacterial growth, in response to nutritional stress signals, via regulation of transcription of genes encoding mainly for amino acid biosynthesis enzymes, amino acid transporters, and aminoacyl-tRNA synthetases (8–11). T-boxes bind specific, cognate tRNA molecules mainly at two conserved sites: the anticodon sequence, which forms base pairs with the codon sequence of the specifier loop (SL); and the tRNA's NCCA 3' accepting end, which pairs with the UGGN sequence found in the antiterminator bulge (12, 13). Recently, a two-checkpoint mechanism has been proposed to serve as a molecular ruler, ensuring appropriate tRNA accommodation through interaction of the apical loop with the D and T loop of the tRNA during the progress of transcription (14).

As the major specificity determinant of a T-box is the unique codon-like nucleotide triplet of the SL, which recognizes only one tRNA anticodon sequence, all T-boxes were considered thus far to be of single specificity, and the possibility of a broader specificity of T-boxes was overlooked (10). However, a closer look at all the available T-box sequences suggests the existence

of T-boxes harboring SLs with two putative codon regions (Tables S1 and S2). In addition, the exact structural requirements dictating the tRNA:T-box interactions remain mainly elusive, despite many attempts based on mutational analyses of the tRNA or the T-box (13, 15). Base replacements in the SL or in the anticodon sequence responsible for specificity lead to a complete or partial decrease in the capacity to bind uncharged tRNA, and subsequently in the ability to trigger antitermination. On the contrary, codon sequence alterations in the SL reassigned the specificity of the T-box, allowing the binding of a new tRNA.

Previous work from our group characterized a T-box (termed NT-box for asparagine) that regulates the transcription of a four-gene operon encoding the enzymes of the entire tRNA-dependent transamidation pathway (the heterotrimeric tRNA-dependent amidotransferase GatCAB and the nondiscriminating AspRS) in *Clostridium acetobutylicum* (*Cac*) (Fig. 1) (16). This NT-box contains an SL that interacts with the GUU anticodon sequence of the *Cac* tRNA^{Asn}. Herein, we provide in vitro and in vivo evidence demonstrating that the NT-box is an example of a class of T-boxes containing a single SL harboring two distinct codons. We have determined the secondary structure of the NT-box, and we show that the SL domain contains an additional guanine inserted between the conserved GAA triplet and the codon sequence. This additional guanine is responsible for an apparent “ambiguity” in codon recognition, which allows the riboswitch to alternately bind two different tRNAs, tRNA^{Asn} and tRNA^{Glu}. When we reduced the size of the SL and changed the position of the specifier codon sequence, we confirmed that the SL remains adequately “flexible” for such interactions. By using a β -gal reporter assay, we also confirmed in vivo that tRNA^{Asn} and tRNA^{Glu} can efficiently mediate transcription antitermination. Finally, we show that this SL codon ambiguity is more widespread than previously thought. Strikingly, bioinformatics analysis suggests that many T-boxes of different bacterial origins harbor putative overlapping or multiple codon SLs that cluster depending on the interconnection of the metabolic pathways that they control. This work provides an example of a T-box that can regulate the synthesis of two amino acids by interacting with two tRNAs involved in different metabolic networks. Our data suggest a revision of current notions of T-box specificity and regulation, and raise questions on the evolution of tRNA-mediated transcription regulation.

Author contributions: N.Y.S., V.S., C.S., and H.D.B. designed research; N.Y.S., V.S., M.B., C.S., and H.D.B. performed research; N.Y.S. and D.D. contributed new reagents/analytical tools; N.Y.S., V.S., C.S., and H.D.B. analyzed data; and N.Y.S., V.S., D.D., C.S., and H.D.B. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

¹Present address: Laboratoire du Métabolisme de l'ARN FRS/FNRS, Institut de Biologie et de Médecine Moléculaires, Center for Microscopy and Molecular Imaging, Université Libre de Bruxelles, 12 B-6041 Charleroi, Belgium.

²To whom correspondence may be addressed. E-mail: h.becker@unistra.fr or cstath@med.upatras.gr.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1304307110/-DCSupplemental.

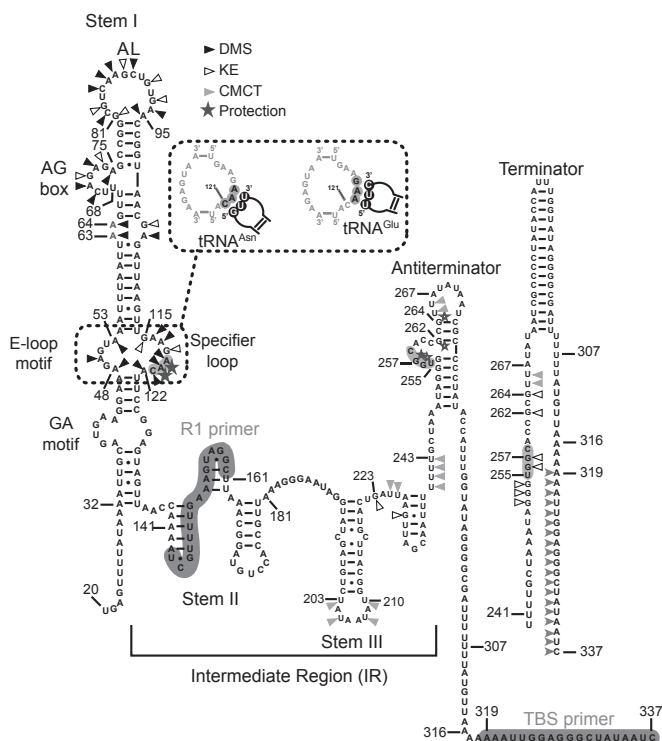


Fig. 1. Secondary structure of the *Cac* T-box (i.e., NT-box). Stem I contains the conserved GA motif and AG box, as well as the SL that displays the GAA and the AAC codons, encoding respectively for Glu and Asn. On the SL, the base pairing of the anticodon sequence of tRNA^{Asn} and tRNA^{Glu} is shown. The 3' end conformations of the terminator and antiterminator structure are shown in comparison. Chemical probing was performed as described in Fig. S1 by using DMS, KE, and CMCT as chemical modification agents. The nucleotides that were accessible for chemical modification are marked using triangles. The protection (star) against modification of specific nucleotides in the SL and the T-box bulge occurs when the tRNA binds the T-box.

Results

Determination of Specific Structural Features of the NT-Box:tRNA^{Asn} Complex After Footprinting Analysis. The secondary structure of the *Cac* NT-box was determined through extensive footprinting analysis of the stem I domain and the terminator/antiterminator domain with the use of specific primers (*Materials and Methods*). The NT-box stem I consists in the apical loop (81–95 nt) and the SL that binds tRNA^{Asn} (Fig. 1 and Fig. S1A, lanes 2 and 10) and contains two smaller bulges at positions 63 to 64 and 68 to 75 (also known as AG box; Fig. S1B). The tRNA^{Asn} GUU anticodon interacts with the AAC codon sequence of SL (positions 119–121; Fig. 1). This interaction is (A₁₁₉ and A₁₂₀) protected from DMS modification (Fig. S1A). Our primer extension analysis also allowed probing of the terminator and the antiterminator conformations in the presence of tRNA^{Asn} (Fig. S1B, lane 9). Two important guanosines at positions 262 and 264 do not seem to participate in the formation of the terminator, but most likely remain unpaired in the absence of the cognate tRNA^{Asn} (Fig. 1 and Fig. S1B, lanes 9–11). However, binding of tRNA^{Asn} protects the two guanosines, which now become inaccessible to modification, presumably because of their participation in base-pairing. This interaction results in the stabilization of the apical stem of the antiterminator (Fig. 1). Analysis of the antiterminator secondary structure showed that it contains two loops, which are located in the positions 255 to 261 and 265 to 273 (Fig. S1B), and tRNA^{Asn} binds to the antiterminator bulge at positions 255 to 261. Finally, in the presence of *Cac* tRNA^{Asn}, stems I and III retain their overall structure (Fig. 1 and Fig. S1A and B).

Identification of Two Codons in the SL of NT-Box. Detailed *in silico* analysis of the SLs from all available T-box sequences suggested clustering according to the annotated function of the gene under control and its corresponding metabolic pathway (17). Initially, we identified SLs with putative multiple codons (PMCs) encoding metabolically related amino acids (Fig. 2 and Table S1). In a next step, we clustered the PMCs according to the function of the downstream genes under regulation. It is evident that, with few exceptions, all the identified genes implicated in amino acid metabolism and transport were under the control of T-boxes bearing PMCs (Fig. 2 and Table S1). Interestingly, all the genes involved in aminoacyl-tRNA synthesis were preceded by T-boxes having a single codon SL, with only three noteworthy exceptions: *asnC* T-box (T-box40), NT-box (T-box41), and *hisSaspS* T-box (T-box43 and T-box46; Fig. 2 and Table S2). The first two T-boxes possess identical SLs with two codons, specifying two metabolically related amino acids (Glu and Asn). We found that *asnC* that encodes a transcriptional regulator belonging to the feast/famine family of regulatory proteins actually encodes an asparaginyl-tRNA synthetase, which is known to be specific for tRNA^{Asn} (18). As *asnC* T-box also contains a Glu codon, it seems to be related to the transamidation pathway of tRNA-dependent asparagine biosynthesis using glutamine as an amide group donor (19). In agreement with a previous study (8), we also showed that the SL of *hisSaspS* T-box (T-box43 and T-box46) contains two overlapping codons (GAC for Asp and CAC for His). Therefore, the *hisSaspS* T-box could sense at least two different amino acids to control—through a yet elusive mechanism—the expression of two aminoacyl-tRNA synthetases (HisRS and AspRS). In a similar mechanism, some T-boxes found upstream of genes or operons involved in amino acid metabolism and bearing potential two-codon SLs could control the balance between two different metabolic pathways by sensing the level of two different amino acids (Table S1). Interestingly, we also noticed that some of the SLs contained more than two PMCs: T-box3 and T-box15 (three PMCs) and T-box46 (four PMCs), and it seems that the number of PMCs increases along with the size increase of SL. In conclusion, our *in silico* analysis clearly suggests that many T-boxes (including the *Cac* NT-box) may respond to more than a single tRNA ligand.

Presence of a Two-Codon SL in *Clostridium acetobutylicum* NT-Box.

To test the hypothesis that emerged from the computational analysis, we performed mutagenesis on the nucleotide positions inside the SL based on our previous studies (16). Our goal was to clarify the specificity of the NT-box and to assign roles of the SL and other important structural domains of the NT-box (Fig. S2A). In addition, we wanted to evaluate the importance of the SL length and capability for codon-anticodon pairing. Therefore, we mutated the NT-box SL and examined the T-box:tRNA complex formation by using EMSA analysis and size-exclusion chromatography (Fig. S3). In addition, we produced two important mutants: NT-box^{Δ(UGG)} and NT-box^(GTBox_IR) (Fig. S2A). The first was necessary to explore whether we could suppress the complex formation after the removal of the highly conserved UGG sequence in the antiterminator bulge. The second mutant [NT-box^(GTBox_IR)] was used to verify whether the change in length of the intergenic region (IR) found between the specifier and the terminator/antiterminator domains affects the complex formation. The long IR of the NT-box was replaced by a shorter one present in *Bacillus subtilis* glyQS T-box (20). Our predictions were verified when we observed that the NT-box^(GTBox_IR) was able to bind *Cac* tRNA^{Asn} (Fig. S2B). The formed complex migrated as a single band and had an apparent K_d of 7 μ M.

As expected from previous *in vivo* experiments (14), removal of the UGG sequence in NT-box^{Δ(UGG)} resulted in an inability to bind *Cac* tRNA^{Asn} (Fig. S2C). In addition to EMSA experiments, we used size-exclusion chromatography to verify the T-box:tRNA complex formation, obtaining similar results (Fig. S3). Although the NT-box^(GTBox_IR) showed a single conformation on

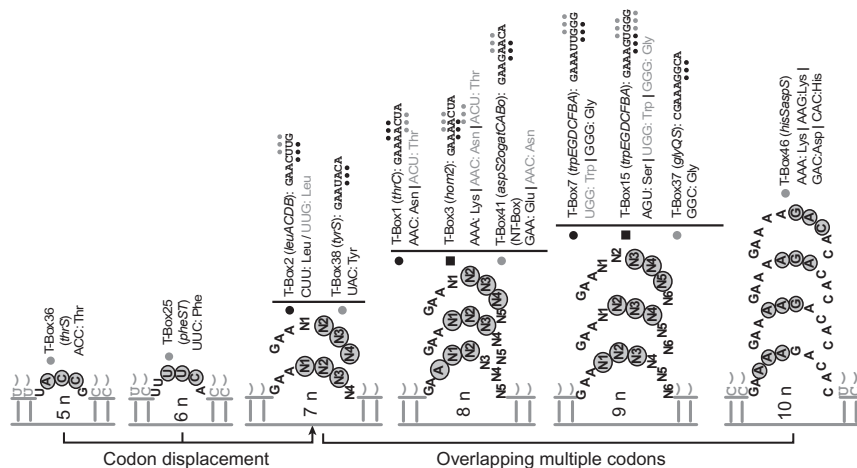


Fig. 2. The SL sequences of T-boxes from different bacterial species. The PMCs (nucleotide triplet) are underlined or overlined with dots, and their identity is indicated for each SL. The number of nucleotides in each SL is also indicated (detailed in Tables S1 and S2). The listed SLs correspond to T-boxes found upstream of genes or operons implicated in amino acid metabolism (black square), transport (grey circle), and in aminoacyl-tRNA biosynthesis (black circle). The codon displacement represents the changing position of the codon toward the 5' or the 3' end of the SL when comparing many T-box SLs.

EMSA, the size-exclusion chromatography profile showed two peaks, suggesting the presence of two conformations. The reason for the late elution of the NT-box^(GTB_{box}IR) compared with the elution of NT-box^{WT} (Fig. S3) could be attributed to the more compact structure of the NT-box^(GTB_{box}IR), which has a shorter IR and lacks stem II. Finally, we confirmed the inability of the *Cac* tRNA^{Asn} to bind NT-box^{Δ(UGG)} by using the same methodology (Fig. S3). To determine the SL-anticodon recognition requirements, we generated a series of mutants (Fig. 3A). In NT-box^(AAC::AAU), the AAC codon was replaced by AAU to check whether wobble pairing takes place during anticodon:specifier codon interaction. This mutant was unable to bind *Cac* tRNA^{Asn}, showing that the anticodon-SL interaction does not tolerate wobble pairing at least in vitro (Fig. S2D). When we used the mutant NT-box^{ΔG} lacking the aforementioned additional guanosine, we observed only a moderate effect (0.2 μM; Fig. S2E). Additionally, we created NT-box^{ΔGW} (W for wobble), which combines the two previous mutations. Interestingly, in this mutant, wobble pairing was allowed, although binding was weak ($K_d > 10 \mu\text{M}$; Fig. S2E). To test the impact of changing the location of the specifier codon in the SL, we generated the NT-box⁽ⁿ⁻¹⁾ and the NT-box⁽ⁿ⁻²⁾ mutants in which the AAC specifier

codon was shifted respectively by one or two nucleotides toward the 5'-end of the loop. In the n-2 mutant, the first A of AAC codon corresponds to the last A of the GAA sequence. As shown in Fig. 3, these two mutants were able to bind *Cac* tRNA^{Asn}. Therefore, we subsequently produced two other mutants: NT-box^{(n-1)ΔA} and NT-box^{(n-2)ΔA}. The first is similar to NT-box⁽ⁿ⁻¹⁾ except that the first A of the AAC specifier codon was removed. In this mutant, the codon is located at the same position as in the NT-box⁽ⁿ⁻²⁾ mutant, except that NT-box^{(n-1)ΔA} has a 7-nt SL. Likewise, NT-box^{(n-2)ΔA} corresponds to NT-box⁽ⁿ⁻²⁾ but is missing the first adenosine of the AAC specifier codon. In this mutant, the two adenines of the AAC codon are located immediately downstream of the G belonging to the GAA sequence. Both mutants can bind *Cac* tRNA^{Asn} as efficiently as NT-box⁽ⁿ⁻¹⁾ and NT-box⁽ⁿ⁻²⁾ mutants. Our data clearly show that the SL can be shortened by 1 nt without significantly altering tRNA^{Asn} binding. Moreover, we show that the AAC codon can be shifted as much as 3 nt toward the GAA sequence without abolishing binding of tRNA^{Asn}. Taken together, our data indicate that T-boxes can exhibit enhanced flexibility in the SL codon sequence. This flexibility could facilitate accommodation of extra tRNAs, and, by doing so, could regulate transcription of the downstream genes in response to several intracellular amino acid concentrations.

Our previous observations prompted us to look for a putative extra interaction of the flexible SL with additional tRNA ligands. The most evident possibility was the binding of the tRNA^{Glu(UUC)} isoacceptor that could interact with the GAA overlapping codon of NT-box SL (Fig. S4). Initially, we tested binding of the NT-box^{WT} with the cytoplasmic tRNA^{Glu} from *Saccharomyces cerevisiae* (*Sce* tRNA^{Glu}). This heterologous tRNA transcript was used because of its high homology to *Cac* tRNA^{Glu} (Fig. S5). Indeed, *Sce* tRNA^{Glu} was able to bind the NT-box^{WT} and the minimum concentration at which a complex started to be detected was 0.2 μM. Our data clearly show that *Cac* NT-box^{WT} can accommodate two different tRNA species to regulate the transcription of the downstream NT-box operon. This result suggests that the flexibility of the *Cac* NT-box SL probably allows two codon-anticodon pairings to occur, enabling tRNA^{Asn} and tRNA^{Glu} to regulate tRNA-dependent synthesis of Asn in response to both Asn and Glu availability.

NT-Box Antitermination Can Be Directed In Vivo both by tRNA^{Asn} and tRNA^{Glu}. To investigate the impact of the tRNA^{Asn}-binding strength on the NT-box mutants as well as the ability of the NT-box^{WT} to recognize different tRNA anticodons in the cellular environment, we used a modified *Escherichia coli* plasmid-based antitermination system (16) (Materials and Methods). By using specific strains, we placed the transcription of *Cac* tRNA^{Asp}, tRNA^{Ser}, *Sce* tRNA^{Asp(C36U)}, and tRNA^{Glu} genes under the

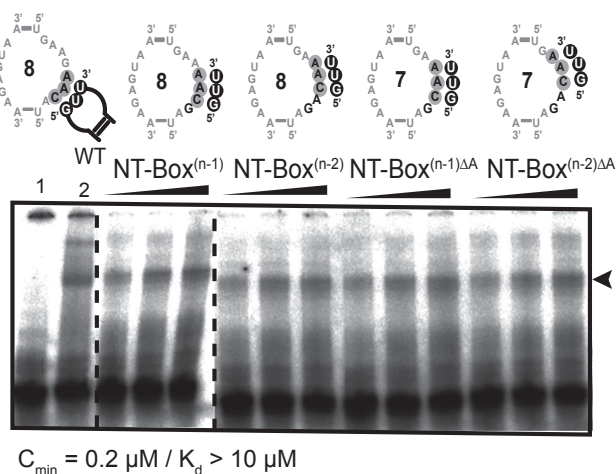


Fig. 3. Specific binding of *Cac* tRNA^{Asn} to NT-box⁽ⁿ⁻¹⁾, NT-box⁽ⁿ⁻²⁾, NT-box^{(n-1)ΔA}, and NT-box^{(n-2)ΔA} as monitored by EMSA. A constant concentration of the 5' [³²P]-labeled *Cac* tRNA^{Asn} was incubated in the presence of increasing concentrations (0.2, 2, and 5 μM) of NT-box variants. The shifted bands are indicated by an arrowhead. Lane 1, *Cac* tRNA^{Asn}; lane 2, NT-box^{WT}. *Cac* tRNA^{Asn}. The number of nucleotides in the SL is indicated in the center of the loop.

control of a T7 promoter (Fig. 4A). As expected, *Cac* tRNA^{Asp} and tRNA^{Ser} failed to induce antitermination of the NT-box^{WT} (Fig. 4B). On the contrary, *Cac* tRNA^{Asn} induced antitermination of the NT-box^{WT} up to 500 Miller Units (MU). Interestingly, *Sc* tRNA^{Glu} reached the same antitermination efficiency, confirming that, in vivo, the GAA codon present in the NT-box^{WT} is fully functional and can be used to regulate transcription of the downstream genes. The *Sc* tRNA^{Asp(C36U)} in which the Asp GUC anticodon was changed into an Asn GUU codon (Fig. S5) also generated antitermination, but with slightly decreased efficiency (250 MU). As expected, the NT-box SL mutants displayed a range of antitermination efficiencies related to their tRNA^{Asn} binding ability, as monitored by EMSA (Fig. 4B). The highest antitermination effect was observed for the NT-box^{ΔG} [compared with NT-box^{WT}, NT-box^{ΔGW}, NT-box⁽ⁿ⁻¹⁾, and NT-box⁽ⁿ⁻²⁾], whereas NT-box^{(n-1)ΔA} retained more than 70% of the NT-box^{WT} antitermination capacity. These results confirm that, in the cellular context, codon recognition can take place as much as 2 nt upstream of the canonical position of the SL codon. As a result, the last A of the GAA sequence can be used for codon-anticodon interaction. In addition, it further confirms that the flexibility in anticodon recognition renders the GAA codon (Glu) fully functional in vivo. On the contrary, NT-box^{(n-2)ΔA} exhibited a weak antitermination activity (50 MU), suggesting that, although the second A of the GAA triplet can be used for base-pairing with tRNA^{Asn} anticodon (Fig. 3), this interaction is too weak to trigger efficient antitermination. This is probably because of steric hindrance resulting from the proximity to the stem located above the SL.

In good agreement with the EMSA results, *Cac* tRNA^{Asn} failed to induce antitermination of the NT-box^(AAC::AAU), confirming that, in vivo, the codon-anticodon pairing does not tolerate wobble pairing in the 8-nt-long SL context of the NT-box. Moreover, the NT-box^(GTBox-IR) retained 90% of NT-box^{WT} antitermination, showing that, in vivo, the length of the IR is not critical for tRNA-T-box binding. Finally, when, by primer extension, we tested the tRNA^{Asn} anticodon binding toward the 5' end of the NT-box⁽ⁿ⁻²⁾ SL (Fig. 4C, lanes 8 and 9), it became evident that binding of tRNA^{Asn} protected the two essential bases, A117 and A118, from DMS modification.

Discussion

Identification and Characterization of a T-Box Riboswitch Harboring a Two-Codon SL. The in vitro and in vivo experiments described herein provide insights concerning the specificity of the NT-box and shed light on the adjustability of codon-anticodon pairing between T-boxes and their tRNA ligands. Our study provides an example of a T-box riboswitch bearing a two-codon SL that is of surprising flexibility. This translates to an overarching flexibility of the organism attained by controlling a single T-box riboswitch using two metabolically related amino acids (Fig. 5). The accommodation of the tRNA's anticodon in the SL can change

through a precise shift by 1 nt toward the 5' extremity, and this key feature switches the specificity of the T-box.

In addition to our in vitro data, our in vivo studies verified that the *Cac* NT-box has dual specificity and can efficiently use two overlapping codons to pair with two different tRNA molecules to trigger antitermination. The high flexibility in anticodon pairing is, however, restricted by the fact that the wobble pairing within the Asn codon is not allowed. This flexibility is therefore less well tolerated in the 3' direction of the SL, at least for the NT-box, suggesting that the structural constraints in codon-anticodon pairing are stronger at the wobble position of the specifier codon.

When we shortened the NT-box length by removing the extra guanosine, wobble pairing was tolerated well (Fig. S2). This observation suggests an essential role for this extra base in the overall SL conformation, probably through structural pressure. This observation is in agreement with recent findings on the structure of the specifier domain in the *tyrS* T-box of *B. subtilis* (12), where the nucleotides of the E-loop motif, bulging at the opposite side of the SL, favor structural stacking on the specifier sequence nucleotides, allowing their orientation toward the minor groove in the specifier domain (12). The fact that the NT-box^{ΔG} was able to induce antitermination beyond any doubt and even better than the NT-box^{WT} suggests that the presence of the inserted guanosine confers selectivity pressure for tRNA binding and antitermination, which is raised when the extra base is absent.

A recent study showed that binding of tRNA to the T-box can also involve interactions between specific nucleotides of the stem I apical loop with the D- and T- loop nucleotides of the tRNA (14). These interactions were proposed to serve as a specific molecular ruler that ensures effective binding of the correct tRNA. The alignment of the apical loops of all *Cac* T-boxes and all *Cac* tRNAs (Fig. S6) showed that the nucleotides involved in such interactions are conserved only in the tRNAs. If we take into account the different lengths of the T-box apical loops together with the replacement of critical nucleotides mediating this apical loop-tRNA interaction, it would be far too speculative to draw any conclusions on whether this second checkpoint also applies to the NT-box-tRNA^{Asn} or tRNA^{Glu} binding. Such a possibility cannot be excluded, but would require further extensive studies.

The T-Box Riboswitch with Two-Codon SL Senses the Balance Between Two Metabolically Related Amino Acids. Our results show that the NT-box presents two equally accessible codons of equal antitermination efficiency, corresponding to the metabolically related amino acids asparagine and glutamate. Glutamate is the precursor of glutamine that provides the most efficient amide group donor for GatCAB amidotransferases, which, in turn, catalyze the tRNA-dependent synthesis of asparagine (Fig. 5). In addition, as the *Cac* GatCAB AdT is dual-specific, Glu can alternatively also be used for the tRNA-dependent synthesis of Gln (16). Our data suggest that the presence of two codons in the NT-box possibly reflects the environmental adaptation developed by

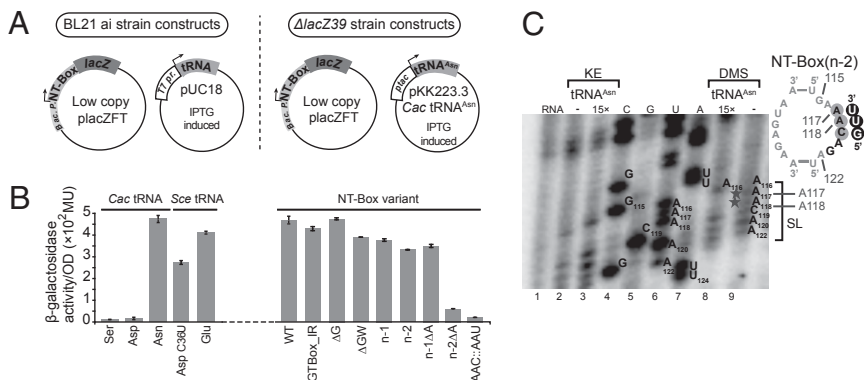


Fig. 4. NT-box-mediated β -gal activity assay. (A) The NT-box^{WT} was cloned into the placZFT plasmid and transformed into a BL21 ai strain. The NT-box^{mutant} and NT-box^{WT} constructs were cloned into the placZFT plasmid and transformed into a Δ lacZ39 *E. coli* strain. (B) The bars represent β -gal activity \pm SD, in Miller Units. (C) Primer extension analysis of the SL in the NT-box⁽ⁿ⁻²⁾ (Materials and Methods). Lane 1 corresponds to the untreated NT-box transcript. Lanes 4–7 correspond to the sequencing reactions.

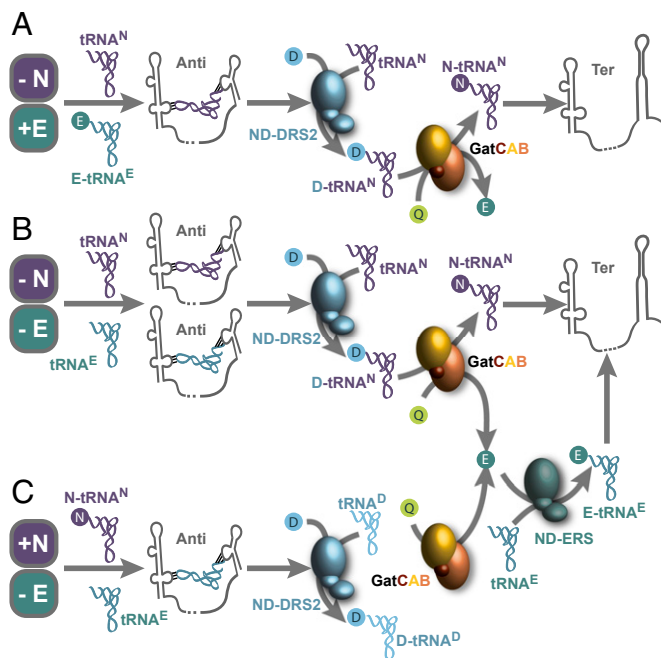


Fig. 5. Two metabolically related, albeit distinct, pathways regulated by a dual-specificity T-box riboswitch. (A) When Asn (marked as “N”) is limited but in the presence of Glu (marked as “E”) (–N/+E), the binding of tRNA^{Asn} to the NT-box triggers the expression of the transamidation pathway (ND-DRS2 and GatCAB) for the tRNA-dependent synthesis of Asn. (B) When both Asn and Glu are limited (–N/–E), the binding of tRNA^{Asn} and tRNA^{Glu} to the NT-box triggers the expression of the transamidation pathway and fosters the glutaminase activity of the tRNA-dependent amidotransferase (GatCAB) for the tRNA-independent synthesis of Glu. Consequently, Glu becomes available for Glu-tRNA^{Glu} synthesis through the activity of the non-discriminating glutamyl-tRNA synthetase (ND-ERS). (C) In the case of limited Glu and in the presence of Asn (+N/–E), most of the tRNA^{Asn} is charged with Asn, and therefore the expressed GatCAB would be used for glutaminase activity. The produced Glu is then used for Glu-tRNA^{Glu} synthesis. The switch from the terminator (Ter) to the antiterminator (Anti) conformation is triggered when Asn and/or Glu are not available. Finally, the activation of the transamidation and glutaminase activity increases the charging of tRNA^{Asn} and/or tRNA^{Glu}, which triggers the switch to the terminator conformation. Q, glutamine or Gln.

Cac to sense glutamate and asparagine levels and respond by transcribing the tRNA-transamidation pathway through which biosynthesis of both amino acids is interconnected.

When the cellular concentration of Asn decreases (Fig. 5A), uncharged tRNA^{Asn} triggers NT-box antitermination and allows transcription and expression of the transamidation pathway enzymes. Transamidation would then allow the Gln-hydrolyzing synthesis of Asn-tRNA^{Asn} and, more importantly, the biosynthesis of asparagine itself, as *Cac* lacks asparagine synthetase activity (16). Because *Cac* encodes the asparaginyl-tRNA synthetase, Asn-tRNA^{Asn} levels are nevertheless coupled to the concentration of free Asn, making the charging level an effective sensor.

Similarly, when Glu levels become low and Asn is available (Fig. 5C), NT-box can respond to the existing uncharged tRNA^{Glu} levels and promote the utilization of GatCAB as a deaminase to produce Glu from Gln. At the same time, the non-discriminating AspRS2 would primarily serve to generate Asp-tRNA^{Asp}. It has been shown that the GatCAB enzyme can function as a glutaminase and that this activity is not necessarily coupled to transamidation of the mischarged Asp-tRNA^{Asn} or Glu-tRNA^{Gln} (21). The ability of *Cac* to restore the Gln levels consumed to generate Glu would not be affected by the use of GatCAB solely as glutaminase. It is known that Gln biosynthesis in *Cac* can be tRNA-independent because a Gln synthetase is present (22).

If, exceptionally, both Glu and Asn are lacking (Fig. 5B), this would be alleviated by the tRNA-dependent, glutamine-hydrolyzing synthesis of Asn. The presence of metabolically related amino acid codons in T-boxes enables a coordination of the expression of the related enzymes with the availability of their related substrates. This feature of dual-specific T-boxes raises questions concerning the evolution of bacterial mechanisms that sense amino acid levels in the environment and regulate their metabolism in the cell. That protein synthesis and amino acid synthesis should be directly coupled to the availability of these substrates is evident, and these results and accompanying analyses describe a potent regulatory link between them. It would therefore be of great importance to extend this investigation to other T-boxes bearing putative PMCs and to investigate whether those PMCs serve as essential regulators.

Multiple-Codon T-Boxes Are Widespread Sensors and Regulators of Interconnected Metabolic Networks.

The most intriguing finding of our *in silico* analysis is the observation that SLs can be clustered in two major groups depending on the function of downstream genes regulated by these riboswitches. T-boxes with single specificity (Table S2) appear to regulate genes encoding mainly aminoacyl-tRNA synthetases. On the contrary, T-boxes with at least two-codon SLs seem to regulate operons that encode metabolically related proteins or encode enzymes that lead to the biosynthesis of metabolically related amino acids (Table S1 and Fig. S7). An exception would be the *Cac* NT-box, which regulates a four-gene operon encoding two enzymes primarily involved in amino acid-tRNA synthesis, and, as such, should be controlled by a single-codon T-box. However, the presence of the tRNA-dependent transamidation pathway does not only compensate for the absence of asparaginyl-tRNA synthetase but also for the absence of asparagine synthetase (16, 23, 24), and thus can also be considered as an amino acid-forming pathway. This intriguing observation suggests that, at least initially, this pathway served as a tRNA-dependent amino acid synthesizing pathway rather than a tRNA charging route. Moreover, the apparent link between the activity of enzymes regulated by a T-box and the number of SL codons displayed suggest that a balancing act must exist between biosynthesis and sensing of amino acids during protein synthesis. This accurate regulation ensures that the organism will be able to manage the available amino acid levels according to fluctuating needs, thus allowing viability even at low amino acid concentrations. The NT-box is a striking example of a molecular switch from a primordial era when tRNA-dependent amino acid and protein biosynthesis as well as tRNA-dependent transcriptional regulation were controlled by the decoding of multiple codon-anticodon pairs, in effect making the genetic code the ultimate regulator of these pathways.

Materials and Methods

In Silico Analysis. Modeling of the T-box secondary structures and the identification of the T-box SLs were done by using the S2S program. The T-box sequences were taken from the Rfam seed (<http://rfam.sanger.ac.uk/> accession no. RF00230) (25). The SLs were compared and the PMCs were identified based on the function of the downstream gene and the corresponding metabolic pathway.

Cloning of T-Boxes, tRNA Genes, and In Vitro Transcription. The WT NT-box (NT-box^{WT}) sequence was PCR-amplified and cloned as previously described (16). The NT-box mutants as well as the *Cac* tRNA^{Asn(GUU)} (*Cac* tRNA^{Asn}), tRNA^{Asp(GUC)} (*Cac* tRNA^{Asp}), and tRNA^{Ser(UGA)} (*Cac* tRNA^{Ser}) were ordered from GenScript and Integrated DNA Technologies. For *in vitro* experiments, the T7 RNAP promoter was integrated in all the NT-box mutant constructs that were cloned into the pUC57 plasmid. The tRNAs were cloned into pIDTSMART plasmids, except for the tRNA^{Asn} gene, which was cloned downstream a *ptac* promoter into the pKK223.3 plasmid. The two tRNAs from *S. cerevisiae* (*Sce* tRNA^{Asp(C36U)} and tRNA^{Glu(UUC)} (*Sce* tRNA^{Glu})) were cloned into the pUC18 plasmid (26). For β -gal assays, the WT and mutant NT-box constructs were cloned, along with their endogenous promoter, upstream of the *lacZ* gene in the placZFT plasmid (27). The NT-box promoter is typically identical to those of *E. coli*. *In vitro* T7 transcripts of the WT and mutated NT-box and tRNAs were obtained as previously described (16, 28, 29).

The NT-box transcript was treated as previously described (16) before binding assays.

Chemical Probing in Vitro of WT and n-2 Mutant NT-Boxes. Modification by DMS, kethoxal (KE) or carbodiimidemetho-*p*-toluenesulfonate (CMCT) was done as previously described (30), with modifications (*SI Materials and Methods*). The chemical modifications were analyzed by primer extension. Two oligonucleotides were used: R1 (5'-GCCTACTTTCAAACGA-3') that hybridizes at nucleotides 144 to 160 and TBS (5'-GATTATAGCCCTCAAATTTT-3') that hybridizes at nucleotides 319 to 337 (positions shown in Fig. 1). The PE experiments were done as previously described (31). DMS modified the base-pairing faces of unpaired A and C, KE modified unpaired G, and CMCT modified, to a higher extent, the exposed bases of U.

EMSA and Size-Exclusion Chromatography. For EMSA, [γ - 32 P]ATP was used for tRNA labeling at their 5'-end. The labeling reaction was performed according to standard protocols. The size-exclusion chromatography assay was performed by using an analytical Superdex 200 column and ÄKTA purifier HPLC (Amersham Biosciences). The two binding assays were performed as previously described (16).

β -Gal Assay. Two *E. coli* strains were used: the *E. coli* M5154 *lacZ* strain [F- Δ *lacZ39*, λ -, *trpA49*(Am), *recA11*, *relA1*, *rpsL150*(strR), *spoT1*] and the

BL21 ai [F- *ompT hsdS_B*(r_B- m_B-) *gal dcm araB::T7RNAP-tetA*] strain that contains a plasmid encoding the T7 RNAP. The growth conditions, as well as the β -gal measurements, were performed as previously described (16, 32). A full description is provided in *SI Materials and Methods*. *E. coli* strains were grown on Luria-Bertani medium supplemented when necessary with 100 mg/L of ampicillin, streptomycin, clarithromycin, and tetracycline. *E. coli* transformation, maxi- and minipreparations of dsDNA, DNA manipulations, and agarose-gel electrophoresis were conducted by using standard procedures.

ACKNOWLEDGMENTS. The authors thank Dr. Jonathan Huot for critical reading of the manuscript and suggestions. This work was supported by the University of Strasbourg, National Program Investissement d'Avenir (LabEx Mitocross), Centre National de la Recherche Scientifique, Agence Nationale pour la Recherche Grant ANR-09-BLAN-0091-02, Association pour la Recherche sur le Cancer, Japan Science and Technology Agency-Centre National de la Recherche Scientifique Strategic Japanese-French Cooperative Program, and University of Patras K. Karatheodoris Grant D164 (to C.S.). Part of this work was implemented under the ARISTEA Action (Hellenic General Secretariat of Research and Technology) of the Operational Program Education and Lifelong Learning cofunded by the European Social Fund and by National Resources Grant D608 (to C.S.). V.S. was the recipient of a Federation of European Biochemical Societies (FEBS) Short-Term Fellowship; and N.Y.S. was the recipient of a fellowship from the Ministère de l'Éducation Nationale, de la Recherche et de la Technologie-Bourse de la Présidence de l'Université and a FEBS Summer Fellowship.

- Grundy FJ, Hodil SE, Rollins SM, Henkin TM (1997) Specificity of tRNA-mRNA interactions in *Bacillus subtilis* tyrS antitermination. *J Bacteriol* 179(8):2587-2594.
- Wels M, Groot Kormelink T, Kleerebezem M, Siezen RJ, Francke C (2008) An *in silico* analysis of T-box regulated genes and T-box evolution in prokaryotes, with emphasis on prediction of substrate specificity of transporters. *BMC Genomics* 9:330.
- Henkin TM, Grundy FJ (2002) In vitro transcription assay for T-Box antitermination system. US Patent 7413856.
- Putzer H, Condon C, Brechemier-Baey D, Brito R, Grunberg-Manago M (2002) Transfer RNA-mediated antitermination *in vitro*. *Nucleic Acids Res* 30(14):3026-3033.
- Mandal M, et al. (2004) A glycine-dependent riboswitch that uses cooperative binding to control gene expression. *Science* 306(5694):275-279.
- Serganov A, Huang L, Patel DJ (2008) Structural insights into amino acid binding and gene control by a lysine riboswitch. *Nature* 455(7217):1263-1267.
- Serganov A, Patel DJ (2009) Amino acid recognition and gene regulation by riboswitches. *Biochim Biophys Acta* 1789(9-10):592-611.
- Gutiérrez-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(1):36-61.
- 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(4):1299-1306.
- 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(17):11121-11126.
- Nelson AR, Henkin TM, Agris PF (2006) tRNA regulation of gene expression: Interactions of an mRNA 5'-UTR with a regulatory tRNA. *RNA* 12(7):1254-1261.
- 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(10):3388-3398.
- Grundy FJ, Moir TR, Haldeman MT, Henkin TM (2002) Sequence requirements for terminators and antiterminators in the T box transcription antitermination system: disparity between conservation and functional requirements. *Nucleic Acids Res* 30(7):1646-1655.
- Grigg JC, et al. (2013) T box RNA decodes both the information content and geometry of tRNA to affect gene expression. *Proc Natl Acad Sci USA* 110(18):7240-7245.
- Grundy FJ, Collins JA, Rollins SM, Henkin TM (2000) tRNA determinants for transcription antitermination of the *Bacillus subtilis* tyrS gene. *RNA* 6(8):1131-1141.
- Saad NY, et al. (2012) Riboswitch (T-box)-mediated control of tRNA-dependent amidation in *Clostridium acetobutylicum* rationalizes gene and pathway redundancy for asparagine and asparaginyln-tRNA^{Asn} synthesis. *J Biol Chem* 287(24):20382-20394.
- Jossinet F, Westhof E (2005) Sequence to Structure (S2S): Display, manipulate and interconnect RNA data from sequence to structure. *Bioinformatics* 21(15):3320-3321.
- Tumbula-Hansen D, Feng L, Toogood H, Stetter KO, Söll D (2002) Evolutionary divergence of the archaeal aspartyl-tRNA synthetases into discriminating and non-discriminating forms. *J Biol Chem* 277(40):37184-37190.
- Sheppard K, Akochy PM, Salazar JC, Söll D (2007) The *Helicobacter pylori* amidotransferase GatCAB is equally efficient in glutamine-dependent transamidation of Asp-tRNA^{Asn} and Glu-tRNA^{Gln}. *J Biol Chem* 282(16):11866-11873.
- Yousef MR, Grundy FJ, Henkin TM (2005) Structural transitions induced by the interaction between tRNA(Gly) and the *Bacillus subtilis* glyQS T box leader RNA. *J Mol Biol* 349(2):273-287.
- Horiuchi KY, et al. (2001) Mechanistic studies of reaction coupling in Glu-tRNA^{Gln} amidotransferase. *Biochemistry* 40(21):6450-6457.
- Fierro-Monti IP, Reid SJ, Woods DR (1992) Differential expression of a *Clostridium acetobutylicum* antisense RNA: Implications for regulation of glutamine synthetase. *J Bacteriol* 174(23):7642-7647.
- Becker HD, Kern D (1998) *Thermus thermophilus*: A link in evolution of the tRNA-dependent amino acid amidation pathways. *Proc Natl Acad Sci USA* 95(22):12832-12837.
- Curnow AW, Tumbula DL, Pelaschier JT, Min B, Söll D (1998) Glutamyl-tRNA(Gln) amidotransferase in *Deinococcus radiodurans* may be confined to asparagine biosynthesis. *Proc Natl Acad Sci USA* 95(22):12838-12843.
- Gardner PP, et al. (2011) Rfam: Wikipedia, clans and the "decimal" release. *Nucleic Acids Res* 39(database issue):D141-D145.
- Becker HD, Giegé R, Kern D (1996) Identity of prokaryotic and eukaryotic tRNA(Asp) for aminoacylation by aspartyl-tRNA synthetase from *Thermus thermophilus*. *Biochemistry* 35(23):7447-7458.
- Feustel L, Nakotte S, Dürre P (2004) Characterization and development of two reporter gene systems for *Clostridium acetobutylicum*. *Appl Environ Microbiol* 70(2):798-803.
- Milligan JF, Groebe DR, Witherell GW, Uhlenbeck OC (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res* 15(21):8783-8798.
- Milligan JF, Uhlenbeck OC (1989) Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol* 180:51-62.
- Luo D, Condon C, Grunberg-Manago M, Putzer H (1998) *In vitro* and *in vivo* secondary structure probing of the thrS leader in *Bacillus subtilis*. *Nucleic Acids Res* 26(23):5379-5387.
- Stamatopoulou V, Toumpeki C, Tzakos A, Vourekas A, Drains D (2010) Domain architecture of the DRpp29 protein and its interaction with the RNA subunit of *Dicystostelium discoideum* RNase P. *Biochemistry* 49(50):10714-10727.
- Henkin TM (2009) Riboswitches: Methods and protocols. *Springer Protocols: in Molecular Biology*, ed Serганov A (Humana, Totowa, NJ), Vol 540, pp 281-290.