Riboswitch (T-Box)-mediated control of tRNA-dependent amidation in *Clostridium acetobutylicum* rationalizes gene and pathway redundancy for asparagine and asparaginyl-tRNA<sup>Asn</sup> synthesis*

Nizar Y. Saad<sup>2</sup>, Bettina Schiel<sup>3</sup>, Mélanie Brayé<sup>2</sup>, John T. Heap<sup>4</sup>, Nigel P. Minton<sup>4</sup>, Peter Dürre<sup>3</sup>, and Hubert Dominique Becker<sup>1,2</sup>

<sup>1</sup>UMR ‘Génétique Moléculaire, Génomique, Microbiologie’, CNRS, Université de Strasbourg, 21 rue René Descartes, 67084 Strasbourg, France.

<sup>2</sup>UPR ‘Architecture et Réactivité de l’ARN’, Université de Strasbourg, CNRS, Institut de Biologie Moléculaire et Cellulaire, 15, Rue René Descartes, F-67084 Strasbourg Cedex, France.

<sup>3</sup>Institut für Mikrobiologie und Biotechnologie, Universität Ulm, 89069 Ulm, Germany,

<sup>4</sup>Clostridia Research Group, Centre for Biomolecular Sciences, School of Molecular Medical Sciences, Nottingham Digestive Diseases Centre, NIHR Biomedical Research Unit, University of Nottingham, Nottingham NG7 2RD, UK.

*Running title: tRNA-dependent synthesis of Asn regulated by a T-Box riboswitch

To whom correspondence should be addressed: Pr. Hubert Dominique Becker, UMR ‘Génétique Moléculaire, Génomique, Microbiologie’, CNRS, Université de Strasbourg, 21 rue René Descartes, 67084 Strasbourg, France, Tel: 33 (0)3 68 85 14 70; Fax: 33 (0)3 68 85 13 65; E-mail: h.becker@unistra.fr and Dr. Bettina Schiel, Institut für Mikrobiologie und Biotechnologie, Universität Ulm, 89069 Ulm, Germany, Tel: 49 (0)731 50 22715; Fax: 49 (0)731 50 22719; E-mail: bettina.schiel@uni-ulm.de

**Keywords:** T-Box; antitermination; transcription; tRNA; tRNA-dependent synthesis of Asn

**Background:** Asparagine (Asn) and Asn-tRNA<sup>Asn</sup> synthesis in *Clostridium acetobutylicum*.

**Results:** Synthesis of Asn is tRNA-dependent and regulated by a T-Box riboswitch that is functional in a Gram-negative environment.

**Conclusion:** The gene redundancy may be connected to the regulation of Asn tRNA-dependent synthesis.

**Significance:** This study points to the involvement of aminoacyl-tRNA synthetases beyond their canonical function in Asn-tRNA<sup>Asn</sup> synthesis.

**SUMMARY**

Analysis of the Gram-positive *Clostridium acetobutylicum* genome reveals an inexplicable level of redundancy for the genes putatively involved in asparaginyl-tRNA<sup>Asn</sup> synthesis. Besides a duplicated set of *gatCAB* tRNA-dependent amidotransferase genes, there is a triplication of aspartyl-tRNA synthetase (AspRS) genes and a duplication of asparagine synthetase B (AsnB) genes. This genomic landscape leads to the suspicion of the incoherent simultaneous use of the direct and indirect pathways of Asn and Asn-tRNA<sup>Asn</sup> formation. Through a combination of biochemical and genetic approaches, we show that *C. acetobutylicum* forms Asn and Asn-tRNA<sup>Asn</sup> by tRNA-dependent amidation. We demonstrate that an entire transamidation pathway composed of AspRS and one set of *GatCAB* genes is organized as an operon under the control of a tRNA<sup>Asn</sup>-dependent T-Box riboswitch. Finally, our results suggest that this exceptional gene redundancy might be interconnected to control tRNA-dependent Asn synthesis, which in turn might be involved in controlling the metabolic switch from acidogenesis to solventogenesis in *C. acetobutylicum*.

Given that ribosomal-directed protein synthesis rests upon the supply of 20 species of aminoacyl-tRNAs (aa-tRNAs), each organism is, in theory, expected to encode a
tRNA-dependent synthesis of Asn regulated by a T-Box riboswitch

complete and unique set of 20 aminoacyl-tRNA synthetases (aaRS). Each aaRS would be responsible for the attachment of a single amino acid (aa) onto its corresponding tRNA. Yet, only a minority of organisms, especially in prokaryotes, encodes this unique and complete set of aaRSs. The vast majority is either lacking, or containing extra aaRSs (1, 2).

In bacteria, the most frequently missing aaRSs are asparaginyl- (AsnRS) and glutaminyl-tRNA synthetases (GlnRS) (3). Their absence is compensated by an alternate and two-steps route called the transamidation pathway. In this pathway, a mischarged glutamyl-tRNA_Gln (Glu-tRNA_Gln) is formed by a non-discriminating glutamyl-tRNA synthetase (ND-GluRS) and subsequently converted, by amidation of the attached Glu, into glutaminyl-tRNA_Gln (Gln-tRNA_Gln) by a tRNA-dependent amidotransferase (AdT). Similarly, asparaginyl-tRNA_Asn (Asn-tRNA_Asn) is synthesized by an AdT-catalyzed amidation of a mischarged Asp-tRNA_Asn, generated by a non-discriminating aspartyl-tRNA synthetase (AspRS).

All bacterial AdTs isolated so far, are composed of three subunits, GatC, GatA and GatB that are assembled in a heterotrimeric enzyme called GatCAB (3). They are dual-specific, capable of converting both, Glu-tRNA_Gln and Asp-tRNA_Asn into Gln-tRNA_Gln and Asn-tRNA_Asn respectively, at least in vitro (4, 5). In vivo, depending on whether only AsnRS or GlnRS are missing or if both are absent, a single GatCAB AdT will be required to generate either only one amide aa-tRNA species or both. Note that the GatCAB-mediated transamidation pathway is predominantly used by bacteria to generate the amide aa-tRNA species, since over the 1086 bacterial genomes that have been sequenced, 90% contain the *gatC*-A and -B genes.

In all bacteria in which amide aa-tRNA formation has been examined, it was found that both pathways, direct charging of tRNA by AsnRS or GlnRS, and transamidation by a GatCAB AdT are mutually exclusive (6) (Figure 1A and 1B). Hence, the use of a GatCAB AdT to generate Gln-tRNA_Gln precludes that of a GlnRS for the same reaction. Likewise, a GatCAB AdT used to generate Asn-tRNA_Asn will exclude the presence of an AsnRS in the organism. However, if asparagine synthetase (AsnA/B), the metabolic enzyme that generates Asn in a tRNA-independent manner, is missing, the GatCAB AdT will be retained together with AsnRS. In this case, the transamidation pathway is necessary for the synthesis of Asn-tRNA_Asn under Asn starvation conditions, and AsnRS is more efficient than the AdT when Asn is present in the medium (7).

The presence of extra aaRSs of the same specificity is also widespread among all species of bacteria. The number of copies of the same aaRS rarely exceeds two copies (duplicated aaRS) and has been reported for almost all of the twenty aaRS species (reviewed in 4). In all cases, the two copies display various degrees of sequence variations and either do not share the same tRNA specificity or are differentially expressed during physiological or environmental changes (8, 9-11). AspRS and GluRS are the most frequently duplicated aaRS found in bacteria, and duplication of these two aaRS species always correlates with the use of the transamidation pathway to generate one or the two amide aa-tRNAs. The rationale for these duplications is that one AspRS or GluRS will charge the cognate tRNA species (tRNA_Asp or tRNA_Glu respectively), while the other will mischarge the tRNA_Asn or tRNA_Gln species to supply the GatCAB AdT with its mischarged substrate (7, 12).

If one only considers synthesis of amide aa-tRNA species, one would realize the huge variability in the combinations of pathways and enzymes used by the bacterial species. This combinatorial diversity is reflected at the genomic level with an extraordinary variation in the composition of the pool of genes devoted to the synthesis of these two particular aa-tRNAs. However, in *Clostridium acetobutylicum*, the combination of genes putatively involved in Asn-tRNA_Asn synthesis escapes any rationale (Figure 1C). *Clostridium acetobutylicum* (Cac) is a spore-forming, Gram-positive, obligate anaerobe, with a high A-T base content (72%) (13). Like most Gram-positive bacteria, Cac is lacking GlnRS.
tRNA-dependent synthesis of Asn regulated by a T-Box riboswitch

and thus forms Gln-tRNA\textsuperscript{Gln} via the transamidation pathway. It is therefore not surprising that the genome encodes a GatCAB AdT. However, the Cac genome reveals the presence of a duplicated set of gatCAB genes in addition to the genes encoding both AsnRS and two truncated asparagine synthetases (Nt-AS and Ct-AS). More surprising, the AspRS is triplicated with one copy, AspRS1, resembling bacterial AspRSs and the two other copies, AspRS2 and AspRS2o, being typically of archaeal architecture (8) (Figure 1C). With respect to Asn-tRNA\textsuperscript{Asn} formation, this gene redundancy looks completely aberrant, since the enzymes of both the direct pathway of tRNA asparaginylation (asparagine synthetase and AsnRS) and of two transamidation pathways (two archaeal-like AspRSs and two GatCAB AdTs) seem concomitantly present in this species. One possible explanation for this gene redundancy and pathway duplication would be that expression of these enzymes is regulated in response to specific physiological or environmental conditions.

We therefore scanned the 5'- and 3'-untranslated regions (UTRs) flanking the subset of redundant genes involved in Asn-tRNA\textsuperscript{Asn} synthesis, searching for mRNA regulatory elements that might regulate expression of these genes. We found that the genes encoding one of the archaeal-like AspRS2 (aspS2o) and one of the GatCAB AdT (gatCABo) are potentially organized in an operon (Figure 1D), and under the control of 5'-UTR cis-acting non-coding RNA, called T-Box. To distinguish the AspRS2 and the GatCAB that are encoded by this operon, we added at the end of each gene name an “o” that stands for “operon” (AspRS2o and GatCABo).

T-Box is a cis-acting riboswitch that is predominantly found in Gram-positive bacteria (14, 15). It uses uncharged tRNA as a ligand that binds to the riboswitch and triggers transcription antitermination. However, the inability of charged tRNA to bind the T-Box induces transcription termination. T-Boxes have been shown to control transcription of a variety of genes encoding aaRSs, aa-forming enzymes or aa transporters (15). This riboswitch allows many bacterial species to respond to the changing levels of the corresponding aa by adapting the expression of genes transporting or using these aa and thus, to respond to certain stress signals (16, 17). When suffering from certain nutritional stresses, the ratio of uncharged tRNAs versus charged tRNAs increases, allowing these uncharged tRNAs to bind to T-Boxes and act as effector molecules to regulate global gene expression (18, 19). These T-Boxes are usually 200- to 300-nucleotides (nt) long, and include a factor-independent (intrinsic) transcription termination signal that adopts a competing antiterminator conformation upon binding of the uncharged tRNA (20, 21), allowing transcriptional read-through of the downstream gene or set of genes. In this case, the increased level of uncharged tRNA serves as signal, relaying to the transcriptional machinery, a deficiency in either aa-tRNA- or aa-forming enzymes. So far, it is known that the specificity of the T-Box response is depending on a single codon present in the specifier domain of this riboswitch, which by pairing with the anticodon of the cognate tRNA, adapts transcription of genes to the level of a single aa intracellular concentration (14, 16).

In the present report we show the presence, in Cac, of an operon encoding the enzymes of an entire transamidation pathway (aspS2ogatCABo operon) regulated by a T-Box riboswitch. We do not only show that this T-Box is functional \textit{in vitro}, but also \textit{in vivo} in a Gram-negative environment. Finally, by a combination of biochemical and genetic approaches, among which the generation of knock-out strains, we bring answers to the apparent aberrant gene redundancy in Cac Asn-tRNA\textsuperscript{Asn} formation.

EXPERIMENTAL PROCEDURES

\textbf{Materials} - L-asparagine, L-aspartate and L-glutamine were from Merck, hydroxyapatite, DEAE–cellulose DE-52 from Whatman, Mono-Q column (Mono-Q\textsuperscript{TM} 10/100 GL) from Amersham Biosciences. All primers were from Sigma; all enzymes were purchased from Fermentas, except restriction enzymes (New England Biolabs) and T7 RNA
polymerase that was prepared as previously described (22). Cac genomic DNA was from ATCC. Plasmid DNA was prepared using the GenElute™ HP plasmid Maxiprep Kit from Sigma-aldrich®. RNA elution was done using Clontech columns (CHROMA SPIN-30) (Clontech laboratories, Inc.). [32P]UTP, and [32P]γATP were from Hartmann Analytic. Cac genomic DNA was from ATCC. Plasmid DNA was prepared using the GenElute™ HP plasmid Maxiprep Kit from Sigma-aldrich®. RNA elution was done using Clontech columns (CHROMA SPIN-30) (Clontech laboratories, Inc.). [32P]αUTP, and [32P]γATP were from Hartmann Analytic. Cac cells were broken using FastPrep® from MP Biomedicals. Sonication was done using VibraCell from Bioblock Scientific. RNA extraction and purification was done using the RNeasy® Mini Kit from QIAGEN.

Construction of T-Box and tRNA gene constructs - The NT-Box sequence was PCR-amplified using Phusion polymerase, 3.65×10^-6 pmoles (100 ng) of the 4.15 Mb Cac genomic DNA, 0.1 nmole of sense (NT-Box_F) and anti-sense (NT-Box_R) primers (Supplementary Table S1). The 425 bp PCR product contains a T7 RNA Polymerase promoter sequence (TAATACGACTCACTATA) extended by two G residues fused to the +1 position of the aspS2ogatCABo leader. The 3'-end of the NT-Box corresponds to the 46th nucleotide (nt) of the aspS2o open reading frame (ORF). The PCR product was cloned into pUC18 plasmid using the HindIII and BamHI restriction sites flanking the fragment's 5'- and 3'-ends, respectively. Cac tRNAAsn(GUU) gene (Cac tRNAAsn) was synthesized by “IDT” (Integrated DNA Technologies, Inc.) and was cloned into pIDTSMART plasmid. The tRNA Asn gene was flanked at the 5'- end with a transzyme as previously described (23). Overexpressed E. coli tRNAAsp and Thermus thermophilus tRNAAsm(GUU) (Tth tRNAAsm) tRNAAsn were already available (24). The NT-Box_placZFT plasmid (25) was constructed for β-galactosidase activity test and for tRNA-directed antitermination in vitro. The NT-Box sequence along with its endogenous promoter was cloned upstream the lacZ gene using Sall and BamHI restriction sites and transformed into E. coli ER strain (asnA-, asnB-) ordered from the Genetic Stock Center (Yale University, New Haven, CT). The NT-Box promoter is typically identical to those of E. coli.

Construction of in vitro T7 RNA transcripts - In vitro T7 transcripts of the NT-Box were obtained as previously described (26, 27). The NT-Box transcript was then resuspended in binding buffer containing 50 mM Tris-acetate, pH 7.0, 25 mM calcium acetate, 100 mM ammonium-acetate and 5% v/v glycerol prior to binding assays. Transcribed tRNAs, and tRNAs overexpressed in E. coli were obtained as previously described (24, 28).

In vitro tRNA-directed antitermination assay - Halted transcription assay were performed using the E. coli RNA polymerase (RNAP) from usb. The reactions were carried out essentially as previously described (29) with some modifications (see supplemental data).

β-galactosidase activity test - To assay tRNA-dependent antitermination in vivo, the E. coli asparagine auxotroph ER strain was cotransformed with the NT-Box_placZFT and the Cac tRNAAsm_pKK223-3 recombinant plasmids. ER strain transformed with the empty (promoterless) placZFT plasmid served as negative control and ER strain transformed with the placZFT plasmid in which the β-galactosidase gene was under control of the bdhB promoter (25) served as positive control. The growth conditions used for β-galactosidase measurements were performed as described previously (30) using the spectrofluorometer from Glomax® Multi Detection System. All measurements were carried out in triplicate and all experiments were performed at least twice.

Amidation - The reaction mixture contained 100 mM NaHepes (pH 7.2), 12 mM MgCl2, 10 mM ATP, 2 mM NH4Cl, 2 mM L-Gln, 10 μM L-[14C]Asp (207 mCi/mmol, from Amersham), and 100 μg of crude extract from E. coli, T. thermophilus, yeast, or Cac. The reaction was conducted at 42°C as previously described (7), and the labeled products were revealed by the image plate reader (Fujifilm FLA 5100, Fujifilm Corporation).

tRNA-dependent transamidation - tRNA-dependent transamidation reactions were performed at 37°C for 10 min in a 50 μl standard reaction mixture as previously described (5, 7). 0.6 mM of an enzyme
preparation of Cac GatCABo or of \( \text{H. pylori} \) (Hp py) GatCAB were used and 0.5 mM of Cac t\( \text{RNA}^{\text{Asn}} \) transcript or 1.6 \( \mu \text{M} \) of Hp py Glu-\( \text{tRNA}^{\text{Gin}} \) were added to complete the reaction. \([^{14}\text{C}]\) amino acids were visualized on TLC plates (TLC cellulose plates: \( 20 \times 20 \text{ cm}^2 \)) and were revealed by scanning the dried TLC plates with the image plate reader.

**Bacterial growth and preparation of protein and RNA extracts** - Cac ATCC 824 cells were grown under batch culture conditions in minimal MES-buffered medium (31) and under strictly anaerobic conditions at 37°C. When necessary, media were supplemented with asparagine (1 mM), aspartic acid (1 mM), ampicillin (100 \( \mu \text{g/mL} \)), chloramphenicol (30 \( \mu \text{g/mL} \)), clarithromycin (5 \( \mu \text{g/mL} \)), or erythromycin (50 \( \mu \text{g/mL} \)). For knock-out strain preparation, Cac ATCC 824 was grown in liquid and solid CGM complex medium (32) when necessary. Thiamphenicol (5 mg/mL) and clarithromycin (5 \( \mu \text{g/mL} \)) were added for different mutant selection steps. Genomic DNA from Cac was isolated by a variation of the Marmur procedure (33).

**Cloning, expression, and purification of the Cac GatCABo AdT** - We designed the Cac \( \text{gatCABo} \) operon as described previously (35) except that the 3'-end of the \( \text{gatB} \) gene was extended in-frame by the sequence encoding the V5 epitope and 6\*His. The \( \text{gatCABo} \) operon was synthesized by Genscript with codon optimization, subcloned into pET20b (Novagen) between the NdEl and XhoI restriction sites. Cells were grown in Luria-Bertani medium at 37°C until mid-exponential phase and expression was then induced by the addition of 0.5 mM IPTG. 10% w/v of glucose was added for the expression of stress chaperons and the cells were left to grow at 18°C under shaking. Purification of the AdT was carried out as previously described (36) with modifications (see supplemental data).

**RT-PCR operon validation** - RT-PCR was performed as previously described (37) with some modifications. cDNA was generated using purified Cac total RNA as a template and the following reverse primers (Supplementary Table S1): \( \text{aspS2oRevRT}, \text{gatCoRevRT}, \text{gatAoRevRT} \) and \( \text{gatArRevRT} \). cDNA synthesis was done in one cycle: 1 h at 42°C, 15 min at 70°C and cooled at 4°C. PCR amplification followed the cDNA synthesis by using the cDNA and the corresponding primers (Supplementary Table S1).

**Preparation of Cac knock-out strains** - Knock-out strains for \( \text{aspS2} \) (CAC3564), \( \text{aspS1} \) (CAC2269), \( \text{gatBo} \) (CAC2976) and \( \text{Nt-AS} \) (CAC2243) were prepared using the Clostridial ClosTron System (38). Primers were designed by using the Targetron Gene Knock-out System Kit from Sigma-Aldrich (see manual’s description for details). For every integration site, four primers were used, among which the EBS universal primer. Integration control was performed as previously described (38). All knock-outs were checked for pMTL007 loss using thiamphenicol selection. All those who lost thiamphenicol resistance and gained clarithromycin resistance were selected. All cultures were grown in complex CGM or minimal MES medium at 37°C under anaerobic conditions.

**Electrophoresis mobility shift assay** - \([^{32}\text{P}]\gamma\text{ATP} \) was used for 5'-end tRNA labeling that was performed according to standard protocols. Binding between tRNAs and NT-Box was assessed using a PAGE mobility shift assay. The binding assay was performed by mixing 0.35 \( \mu \text{M} \) of 5'-\([^{32}\text{P}]\) labeled tRNAs with 0.078 to 20 \( \mu \text{M} \) NT-Box. After denaturing and renaturing the two RNAs
separately, the NT-Box transcript was treated by adding Mg\(^{2+}\) to the binding mixture before tRNA addition. The radiolabeled tRNA was added at constant concentration to an increasing amount of NT-Box, and the mix was left at room temperature for 25 to 30 min before loading on 6% v/v non denaturing polyacrylamide gel. The gel composition was 6% v/v polyacrylamide in Tris-Borate buffer, 5 mM MgCl\(_2\), 50 mM NaCl and 5% v/v glycerol. The shifted complex and free \[^{32}P\] tRNAs were visualized by scanning with the Fujifilm image plate reader.

**Size-exclusion chromatography assay**

The size-exclusion chromatography assay was performed using an analytical size-exclusion chromatography column and the ÄKTA purifier HPLC (Amersham Biosciences). The binding assay was performed by mixing different concentrations of tRNA\(^{Asn}\) and NT-Box. Before binding, Cac and Tth tRNA\(^{Asn}\) were denatured for 5 min at 90°C and let to fold for 10 min at room temperature. The NT-Box was heated at 70°C for 10 min and mixed immediately, in the EMSA binding buffer, with 25 mM Mg\(^{2+}\) and tRNA\(^{Asn}\). The mix was then incubated at Room temperature for 25 min before loading on column. 20 µL of sample volume containing the Cac tRNA\(^{Asn}\) transcript were loaded. The Tth tRNA\(^{Asn}\) transcript was loaded in a 40 µL sample volume. All size-exclusion chromatographies were performed at 4°C using the binding buffer.

**Analysis of fermentation products by GC**

The concentration of the fermentation products acetone, ethanol, butanol, butyrate, acetate, and 3-hydroxybutanone (acetoin) was determined by gas chromatography as previously described (31).

**RESULTS AND DISCUSSION**

**Cac gene redundancy for Asn/Asp-tRNA synthesis** - Analysis of the genomic content of Cac shows a high redundancy in genes encoding enzymes involved in Asn- and Asp-tRNA formation (Figure 1C). In addition, preliminary analysis of the loci encoding these genes, suggest that both gat\(^{CAB}\) sets of genes would be arranged in an operon except that one operon, asp\(^{S2ogatCABo}\), would also include an additional archaeal-like AspRS2 located upstream the AdT genes (Figure 1D). Our first goal was to confirm the operon organisation of the two gat\(^{CAB}\) genes (gat\(^{CABo}\) and gat\(^{CAB}\)) and to validate the presence of a bigger operon for containing the asp\(^{S2o}\) and the gat\(^{CAB}\).

**Existence of a gat\(^{CAB}\) and a asp\(^{S2ogatCABo}\) operon in Cac** - Using RNA extracts from different growth conditions, we were able to amplify by RT-PCR the RNA sequences located between asp\(^{S2o}\) and gat\(^{Co}\) as well as between gat\(^{C3}\) and gat\(^{A}\) thereby confirming the presence of an asp\(^{S2ogatCABo}\) and a gat\(^{CAB}\) operon (Figure 2). The RT-PCR experiments showed that both operons are transcribed regardless of the metabolic phase used by Cac to process carbohydrates. Indeed, Cac is capable of fermenting a large variety of carbohydrates into acids and solvents. Acids like acetate and butyrate are produced during exponential growth phase, also called the acidogenesis phase, and their accumulation triggers a shift to a solventogenesis phase during which solvents such as butanol are produced (39). Therefore, one possible explanation for this gene redundancy and pathway duplication would be that these enzymes are differentially expressed during acidogenesis and solventogenesis.

**Construction and analysis of Cac strains knocked out for glutamine-dependent amidotransferase (gat) or asp\(^{S}\) genes** - To test the aforementioned hypothesis, we engineered Cac knock-out strains for a subset of these redundant genes. Using the Clostron system, based on the genomic integration of a group II intron in the target genes, we succeeded to target four genes: CAC3564, coding for the archaeal-like AspRS2; CAC2269, coding for the bacterial-type AspRS1; CAC2243, coding for the Nt-AS and CAC2976, coding for the GatBo subunit of the GatCABo AdT (Supplementary Figure S2). Except Nt-AS (Nt-AsnB) mutant strains, all knock-out strains grew on minimal medium in both acidogenesis and solventogenesis (not shown). This observation strongly suggests that the activity of the corresponding enzymes can be compensated by their duplicated homologs,
and that there are indeed redundant genes of Asn/Asp-tRNA synthesis in Cac. Until now, we do not know why the Nt-AS encoding gene would be essential. The most obvious possibility would be that Nt-AS exhibits an essential function beyond Asn synthesis. 

Expression of AspRS2 is related to acid and solvent production - We also grew the knock-out strains on minimal medium in the presence of Asn in order to check whether Asn downregulates the expression of some of these genes. Consequently, it would render the corresponding knock-out strain unable to sustain growth in the presence of this aa. In addition, analyses of acids (acetate and butyrate) and solvents (ethanol, acetone, butanol and 3-hydroxybutanone) production were checked by gas chromatography. The growth (Figure 3) and level of acids and solvents production of the wild-type (WT) Cac (Supplementary Figure S3) were similar regardless of the presence or absence of Asn. However, the WT strain showed changes in the physiological events when shifting from acidogenesis to solventogenesis after 42 hours (Figure 3 and Supplementary Figure S3) as previously observed (40).

Analysis of the growth curves and acid and solvent production of the aspS1 knock-out mutants showed a profile similar to that of the WT strain (Supplementary Figure S4 and S5). However, analysis of the growth curve of the aspS2 knock-out mutant showed significant delay (30 hours) of the shift from acidogenesis to solventogenesis (Figure 3) in the absence of Asn. This growth profile influenced the timing (delay of 30 hours) of acid and solvent production, although their concentrations and levels of production remained comparable to that of the WT (Supplementary Figure S3 and S4). Adding Asn to the culture of the aspS2 knock-out mutant allowed the strain to regain the WT growth profile (Figure 3), as well as the WT acid and solvent production yields (Supplementary Figure S3). These results tend to suggest that there is an Asn-dependent expression of the aspS2 gene and that this regulation is somehow involved in acid and solvent production. These lines of experiments suggest that AspRS2 might be involved in Asn synthesis and that Asn is somehow connected to the switch from acidogenesis to solventogenesis.

The synthesis of asparagine in Cac is tRNA-dependent - In all the bacteria examined so far, both pathways, direct charging of Asn onto tRNA<sub>Asn</sub> by AsnRS and transamidation by a GatCAB AdT, have been shown to be mutually exclusive (6). However, Cac genome encodes both AsnRS and GatCAB AdTs, suggesting that this organism uses both pathways for Asn synthesis. To clear out this issue, we checked whether Cac was able to generate free Asn.

Two truncated asparagine synthetase B-related ORFs, Nt-AS and Ct-AS, were identified in Cac genome. Alignment of the Nt-AS and Ct-AS aa sequence with that of the <i>E. coli</i> AsnB (Supplementary Figure S6) allowed us to verify the presence and conservation of motifs critical for AS activity. The alignment shows that Nt-AS had conserved the glutamine-binding domain and the AMP-generating domain (41), and might therefore be capable of generating Asn. On the contrary, Ct-AS had only conserved the AMP-generating domain and should therefore be unable to catalyze amidation of Asp into Asn. To verify the activity of both Nt-AS and Ct-AS, we first checked their ability to complement, independently or in combination, the <i>E. coli</i> Asn auxotrophic ER strain (asn<sup>A</sup>, asn<sup>B</sup>) (42). Transformation of the ER strain with the plasmid-borne Cac Nt-AS and Ct-AS was verified by PCR and expression of the enzymes in the ER recombinant strains was checked by Western blot (not shown). Figure 4A shows that both the Nt-AS and Ct-AS constructs can complement Asn auxotrophy of the ER strain. However, when the two constructs were co-transformed into the ER strains, they were not able to complement the Asn auxotrophy, suggesting that they are not functional when combined. The discrepancy between the ER complementation assays, using Nt-AS and Ct-AS individually or in combination, is unquestionably puzzling and we have no clear answer for it. A possible explanation for this result would be that the Ct-AS may regulate the expression or activity of the Nt-AS similarly to what has been described for the nitrogen assimilation control
protein (Nac) in *E. coli* (43). In *E. coli*, expression of AsnC and AsnA enzymes is repressed by Nac. In this case, Nac directly represses *asnC* expression, whose product is required for the activation of *asnA* transcription.

To verify whether Cac is able or not to generate free Asn by using the conventional pathway catalyzed by AsnB, we checked the capacity of Cac protein extracts to catalyze tRNA-independent Asn formation. Proteins were extracted from Cac cells grown in minimal medium or minimal medium supplemented with either Asp or Asn. Additionally, growth cultures were either stopped during the acidogenesis or the solventogenesis phase. Figure 4B shows that all extracts were unable to catalyze in vitro amidation of Asp into Asn in conditions in which the asparagine synthetase activity of a *S. cerevisiae* crude extract could be detected.

These lines of experiments confirm that, under conventional physiological states, Cac does not display any detectable tRNA-independent Asn formation activity. This result is in agreement with the absence of complementation of the co-transformed ER strain, as well as with the proteomic and transcriptomic results obtained by Janssen and coworkers (44). In their recent work, they showed no evidence for asparagine synthetase expression in these conditions. However, the oddity regarding asparagine synthetase expression in these conditions, the oddity regarding asparagine synthetase complementation assays has to be further investigated in order to decipher the real function of both AS proteins. Their individual enzymatic activities must be verified as well as their expression profiles.

*No redundancy in Asn-tRNA<sup>Asn</sup> synthesis in Cac* - Given that Cac is unable to generate Asn in a tRNA-independent manner, generation of Asn-tRNA<sup>Asn</sup> is not accomplished by the concomitant use of direct and indirect pathways. However, enzymes of both routes have been kept. It is surprising that AsnRS was retained, despite the fact that the organism is unable to synthesize its aa substrate. This situation was already reported for *T. thermophillus* (7). We therefore hypothesized that, like in the *Thermus* case, when Cac can find and import Asn from its environment, AsnRS would catalyze formation of Asn-tRNA<sup>Asn</sup>, probably because of its higher catalytic efficiency. On the other hand, the transamidation pathway has been logically conserved, being essential when Asn is unavailable. In order to bring evidence to this scenario, we checked the capacity of at least one of the GatCAB AdT (GatCABo) to catalyze in vitro the tRNA-dependent generation of Asn. Figure 5A shows that the purified GatCABo AdT is able to transamidate Cac Asp-tRNA<sup>Asn</sup> transcript. This activity is strictly Gln- and tRNA-dependent.

The *GatCABo* AdT is able to generate both Asn-tRNA<sup>Asn</sup> and Gln-tRNA<sup>Gln</sup> - Since Cac possesses two GatCAB AdTs (GatCABo and GatCABb), one possible explanation for this duplication would be that one AdT would be restricted to Asn-tRNA<sup>Asn</sup> synthesis, and the other to Gln-tRNA<sup>Gln</sup> formation. However, this would be really surprising, since all bacterial GatCAB AdTs have been shown to be dual specific (36). Figure 5B confirms, using pure heterologous *H. pylori* Glu-tRNA<sup>Gln</sup>, that the GatCABo AdT is indeed dual specific, able to form Gln-tRNA<sup>Gln</sup>. Note that none of the studied bacterial GatCAB AdTs exhibited a species specificity for amide tRNAs, because all tRNA<sup>Asn</sup> and tRNA<sup>Gln</sup> display the same tRNA identity elements for the GatCAB AdTs, including Cac tRNAs (36). This result additionally shows that the absence of GlnRS in Cac can be compensated by GatCAB-mediated formation of Gln-tRNA<sup>Gln</sup>. However, this line of evidence further supports the presence of redundant AdTs in Cac, especially when considering the close structural relationship that can be deduced from the phylogeny of both GatCAB GatB subunits (Supplementary Figure S7). We analyzed the phylogeny of GatB, since this subunit is restricted to AdTs. It showed that both GatBo and GatB are in the same bacterial clade, and are not clustered with the archaeal GatB of the monospecific archaeal GatCAB AdTs (45). As a result, both GatCABs - GatCABo and GatCABb - will very likely display the same activities and specificities. This observation is further supported by our knock-out strain analysis, showing that the loss of the GatBo subunit is not lethal.
only explanation for this is the complementation by the remaining GatB subunit (Figure 1C). However, further investigations are needed to verify whether both AdTs have the exact same substrate specificities or whether they are identically expressed and distributed along the two physiological states.

Regulation of aspS2ogatCABo transcription by a tRNAAsn-dependent T-Box riboswitch - The presence of the two GatCAB AdTs would particularly make sense if, for example, one GatCAB would be preferentially expressed during acidogenesis, while the other would be prevalently used during solventogenesis. From the growth analysis of the knock-out strain we generated, we already knew that removal of GatCABo activity by deletion of GatBo can be compensated by the GatCAB AdT. However, this result does not preclude a preferential use of one GatCAB AdT over the other in a metabolic phase-dependent manner. We therefore screened the 5’ and 3’ UTRs flanking the aspS2ogatCABo and gatCAB operons for mRNA regulatory elements that might regulate their expression, and found a putative T-Box located in the 5’UTR of the aspS2ogatCABo operon. RT-PCR-amplification of the RNA sequence located between the T-Box and aspS2o confirmed the presence of a T-Box in the 5’UTR of the aspS2ogatCABo transcript (Figure 2A and 2B).

Figure 6A shows the model of secondary structure of the Cac aspS2ogatCABo leader sequence that displays the T-Box we reconstructed using mfold (46), the known structures of T-Boxes and the Rfam T-Box alignment (47). All idiosyncratic T-Box structural and sequence motifs could be found. The presence of an “AAC” Asn codon in the stem I specifier-loop suggests that the tRNA ligand of this T-Box is tRNAAsn. As a consequence, we named this T-Box, NT-Box (N for asparagine). The sequence also shows the stem II and the stem III, which both form the unconserved intermediate region of the T-Box. The riboswitch ends with a 14 nt-long conserved T-Box sequence or domain able to adopt the mutually exclusive terminator or antiterminator conformations, in response to tRNA binding. Base-pairing between the tRNAAsn and the NT-Box is done as follows: in the 5’-end, the NT-Box 119AAC121 nucleotides, located in the specifier-loop, base-pair with the tRNAAsn anticodon triplet 3GUU36. In the 3’-end, the 255UGGC259 NT-Box antiterminator bulge base-pairs with the tRNAAsn 77GCCA76 3’-end (Figure 6A). Based on previous studies on T-Boxes, the NT-Box may respond to Asn starvation or supply. In principle, upon Asn starvation or limitation, tRNAAsn is mainly uncharged, and therefore capable to interact with the nascent leader RNA, stabilizing the antiterminator conformation, thereby allowing transcription of aspS2ogatCABo (Figure 6B). Since this operon encodes both enzymes necessary to the tRNA-dependent formation of Asn and Asn-tRNAAsn, the transcriptional read-through would allow Asn formation and utilization. When the physiological levels of Asn are restored, tRNAAsn is probably mainly charged with Asn, and therefore unable to stabilize the antiterminator structure of the T-Box domain. This domain, by adopting the more stable terminator structure, may prevent the transcription of the aspS2ogatCABo operon (Figure 6B).

In order to validate this regulatory mechanism and to confirm that the NT-Box is functional, we designed experiments that aimed at confirming the capacity of purified in vitro transcribed NT-Box to specifically recruit and bind Cac tRNAAsn transcript. Formation of the NT-Box•Cac tRNAAsn duplex was assayed using both electrophoretic mobility shift assay (EMSA) and size exclusion chromatography.

EMSA analysis - So far, EMSA studies that analyzed T-Box-tRNA complex formation were only performed using truncated T-Box transcripts, forming either the specifier stem-loop or the T-Box domain (19, 48). However, the structural probing studies that analyzed the conformational switch in response to tRNA binding were done using full-length T-Boxes (49, 50).

Figure 7A shows formation of the NT-Box•Cac tRNAAsn duplex, in which the entire 400 nt-long NT-Box transcript was used. Absence of duplex formation using Cac
tRNA<sup>Asp(GUC)</sup> (Cac tRNA<sup>Asp</sup>) transcript (Figure 7A, lane 2) confirmed the strict tRNA<sup>Asn</sup> ligand specificity of this T-Box. The tRNA specificity exhibited by the NT-Box was remarkably high, since the absence of one of the seven base pairs involved in T-Box•tRNA duplex formation hindered tRNA binding and/or antitermination. The presence of C<sub>36</sub> in tRNA<sup>Asp</sup> (Supplementary Figure S8) instead of U<sub>36</sub>, found in tRNA<sup>Asn</sup>, avoided base-pairing with NT-Box A<sub>119</sub>. Other residues from the T-Box and the tRNA are likely to interact, and some may also be important for specificity either directly or indirectly.

In the EMSA assay, we noticed the presence of two main NT-Box•Cac tRNA<sup>Asn</sup> duplexes. The use of increasing concentrations of NT-Box allowed the determination of the two dissociation constants (K<sub>d</sub>) corresponding to the two forms of NT-Box•Cac tRNA<sup>Asn</sup> duplex. K<sub>d</sub> of 6 to 8 µM and of 10 µM were determined for the upper (black arrow) and lower duplex forms respectively (See supplemental data for K<sub>d</sub> determination). These K<sub>d</sub> are 6 to 9-fold lower than the one that was previously determined for binding of the B. subtilis antiterminator domain of the tyrS T-Box with its cognate tRNA<sup>Tyr(A73U)</sup> (K<sub>d</sub> = 63 µM) (48). In addition, the NT-Box•Cac tRNA<sup>Asn</sup> duplex starts to form at a minimal concentration of 0.078 µM of NT-Box.

A dissociation constant of 2 - 3 µM was determined using the pure in vivo expressed Tth tRNA<sup>Asn</sup> (Supplementary Figure S9A). Since this tRNA<sup>Asn</sup> was overexpressed in E. coli, and therefore harbours E. coli tRNA<sup>Asn</sup> post-transcriptional modifications (51), our results suggest that Cac nucleotide modifications might increase the affinity of the NT-Box for its cognate tRNA<sup>Asn</sup>. Note that the effect of the tRNA nucleotide modifications on the affinity for T-Boxes has, yet, not been studied.

Nonetheless, altogether, the affinity of NT-Box for tRNA<sup>Asn</sup> is comparable to the affinities measured for aaRS•tRNA duplex formation (52), which is expected, since NT-Box has to compete with possibly two archaeal AspRSs (AspRS2o and AspRS2) and AsnRS for tRNA<sup>Asn</sup> binding, in order to trigger transcription of the aspS2ogatCABo operon.

**Size-exclusion chromatography** - We further confirmed formation of the NT-Box•Cac tRNA<sup>Asn</sup> duplex using a different approach namely the analytical size-exclusion chromatography. This technique was never applied before to study T-Box•tRNA duplex formation. Figure 7B shows that when the NT-Box was mixed with Cac tRNA<sup>Asn</sup>, a new elution peak was obtained. This peak accounts for elution of a higher molecular weight particle than that of the NT-Box alone or of Cac tRNA<sup>Asn</sup>. Denaturing PAGE analysis of the RNA species present in this peak confirmed that both the NT-Box and Cac tRNA<sup>Asn</sup> were present in the corresponding fractions (Figure 7B, lane c and d). Figure 7B and Supplementary Figure S9B show that NT-Box•Cac tRNA<sup>Asn</sup> peak was not completely symmetric, suggesting the presence of two conformations for the duplex, probably due to a high dynamics in assembly of the RNA•RNA complex.

In addition to the EMSA results, the size-exclusion chromatography assay provides new evidence showing that the two conformations of the duplex may be due to the presence of two NT-Box conformations. (Figure 7B, NT-Box elution profile). The elution profiles of the NT-Box and the NT-Box•Cac tRNA<sup>Asn</sup> duplex showed that these two alternative conformations are present, regardless of whether Cac tRNA<sup>Asn</sup> is binding or not to the NT-Box (Figure 7B). It is likely that they are provoked by a certain domain in the T-Box structure, which is not involved in tRNA<sup>Asn</sup> binding. We hypothesize that the long intermediate region, located between the specifier and T-Box domain can adopt alternative conformations, yielding the two conformers we observe both in our EMSA and size-exclusion assays. This is in agreement with previous results reporting the possibility that T-Box long intermediate regions might alternatively fold into a pseudoknot structure (16). The presence of two conformers, due to alternative conformations of the intermediate region, has also been observed in the report describing the structural probing experiments done on the glyQS T-Box in B. subtilis. In fact, Yousef and coworkers (50) showed the existence of conformational changes in the
intermediate region just upstream of the antiterminator element.

The use of the post-transcriptionally modified Tth tRNA\textsubscript{Asn} yielded the same elution profile than the transcript. However, when Cac tRNA\textsubscript{Asp} was used for assaying duplex formation with NT-Box, no elution peak relative to a particle of higher molecular weight could be detected, confirming that the NT-Box•Cac tRNA\textsubscript{Asp} is also not detected using size-exclusion chromatography (not shown). These experiments, not only confirmed the results obtained using EMSA, but also validated the use of size-exclusion chromatography for analyzing T-Box•tRNA complex formation.

Design of an in vivo NT-Box antitermination assay in a Gram-negative environment - Since the \textit{in vitro} read-through experiment was not conclusive regarding NT-Box antitermination (Supplementary Figure S10), we searched for another approach to demonstrate the capacity of the NT-Box to perform aa- and tRNA-dependent transcription antitermination.

Most of the antitermination experiments are done using aa auxotroph \textit{B. subtilis} strains, genetically modified to encode a single chromosomal copy of a \(\beta\)-galactosidase gene under the dependence of the studied T-Box. This experimental design allows to follow, \textit{in vivo}, T-Box read-through as a function of aa starvation, by measuring \(\beta\)-galactosidase activity (18). However, transplanting the same experimental design in Cac was simply not feasible, mainly due to technical constrains when working with a strict anaerobe. On the other hand, trying to transplant this experimental design in a Gram-negative environment had never been reported. We engineered an \textit{E. coli} plasmid-based system in which the \(\beta\)-galactosidase reporter gene is under the dependence of the NT-Box. Since NT-Box should, in principle, trigger read-through of the downstream gene in response to Asn starvation, we used the \textit{E. coli} asparagine auxotroph strain (\(ER\) strain). This strain was used in order to be able to control the precise amounts of the aa supplemented into the medium. This plasmid construct is based on a promoterless, low copy plasmid, pIacZFT that had already been used to measure promoter strength in Cac (25). In our construct, the \textit{lacZ} gene is preceded by the NT-Box sequence, along with its endogenous promoter (Figure 8A). Results of the \textit{in vitro} halted-complex transcription assay with the \textit{E. coli} RNAP indicated that the endogenous Cac promoter is well recognized by \textit{E. coli} RNAP. A pKK223.3 recombinant plasmid containing an IPTG-inducible Cac tRNA\textsubscript{Asn} gene was also introduced into the strain.

\textit{In vivo} NT-Box antitermination detection - Figure 8B shows, as expected, an increased \(\beta\)-galactosidase activity (measured as Miller Units) under Asn starvation (-Asn, -IPTG), showing that NT-Box read-through requires uncharged tRNA\textsubscript{Asn}, which, in these conditions, is that of \textit{E. coli} (Figure 8A). When Asn was added up to 50 \(\mu\)g/mL (+Asn, -IPTG), no detectable \(\beta\)-galactosidase activity could be measured, indicating that Asn-tRNA\textsubscript{Asn} formation could no longer trigger formation of the antiterminator conformation (Figure 8A and -B). As expected, we observed a maximum read-through when combining both Asn starvation and Cac tRNA\textsubscript{Asn} overproduction (Figure 8B: -Asn, +IPTG). Consequently, both molecules may have a cooperative effect on NT-Box antitermination. When Cac tRNA\textsubscript{Asn} expression was induced, even in the presence of Asn (+Asn, +IPTG), a \(\beta\)-galactosidase activity 2 times smaller than the activity measured without tRNA induction and without Asn (-Asn, -IPTG) was detected (Figure 8B). This shows that overexpression of tRNA triggers the adoption of the NT-Box antiterminator conformation in an aa-independent manner. Moreover, it shows that the level of tRNA expression is an important effector of T-Box antitermination. This makes sense, if one considers that endogenous AsnRS will probably be unable to aminoacylate the non physiological amounts of tRNA\textsubscript{Asn} that are produced during the overexpression of the heterologous Cac tRNA\textsubscript{Asn}. In these conditions, production of Asn-tRNA\textsubscript{Asn} is not limited by the intracellular concentration of Asn, but by AsnRS activity. Therefore, the exceeding pool of uncharged tRNA\textsubscript{Asn} will bind to the NT-Box, even if
there are physiological amounts of Asn-tRNA\textsubscript{\textsuperscript{Asn}} that are formed.

In our \textit{in vivo} experiment, we observed that the exceeding pool of the induced uncharged tRNA\textsubscript{\textsuperscript{Asn}} may be responsible for a “leaky” read-through. It is worth to note that a different sort of “leaky” read-through has also been noticed \textit{in vitro}, when using the \textit{B. subtilis} RNAP (50). As a consequence, one could speculate that it could have a biological significance. For example, the maintenance of a small amount of the operon mRNA, resulting from leaky read-through, could be used to give a quick initial response to the need for aminoacyl-tRNA synthesis, since it only needs to be translated. Meanwhile, the T-Box-controlled accumulation of the operon mRNA would be essential to prolong and amplify the response.

**CONCLUDING REMARKS**

The present report provides compelling evidence that unlike what could be deduced from Cac gene content, the direct and indirect routes to Asn and Asn-tRNA\textsubscript{\textsuperscript{Asn}} synthesis are not redundant in this organism. While our biochemical and genetic experiments could not elucidate the reason why Cac encodes extra AspRSs and GatCABs, they suggest a differential preferential use of each copy during the various metabolic phases adopted by this bacterium. More importantly, our results point toward an important and interconnected role of Asn and AspRS triplication for the switch between acidogenesis and solventogenesis, as seen by the phenotypic effect of the AspRS2 knock-out. These observations suggest that some of the redundant genes encoding enzymes of Asn/Asp-tRNA\textsubscript{formation} could participate to Cac homeostasis. However, this does still not explain the relevance of this redundancy.

The presence of an effective T-Box, tightly regulating the expression of an entire transamidation pathway is another argument in favour of Asn as a potential effector for the metabolic switch. Although we did not yet prove the involvement of the NT-Box in the metabolic switch, we think that it is very likely the case. Indeed, AspRS2 is connected to the metabolic switch but also controls the level of tRNA\textsubscript{\textsuperscript{Asn}} charging, which is crucial for \textit{aspS2ogatCABo}-mediated formation of Asn.

Considering that the quantity of uncharged tRNA\textsubscript{Asn} that governs NT-Box read-through and Asn production can potentially be controlled by the charging activities of at least three enzymes, AsnRS, AspRS\textsubscript{2o} and AspRS2, one can easily predict that AspRS2 duplication reflects their involvement in NT-Box regulation and therefore Asn synthesis. Further transcriptomic studies will be needed to decipher the intricate story of Cac Asn and Asn-tRNA\textsubscript{Asn} synthesis.

Our \textit{in vitro} studies made use of the size-exclusion chromatography as a novel method to detect T-Box\textsubscript{T}tRNA complex formation, and confirmed the presence of the two NT-Box conformers observed in the EMSA assay. The unstable conformation of the NT-Box intermediate region did not hinder tRNA\textsubscript{Asn} binding. It would be interesting to see whether this intermediate region is essential for NT-Box antitermination \textit{in vivo}. If so, the stabilization of the intermediate region conformation could necessitate the presence of protein cofactors. The latter could also be involved in hindering “leaky” read-through or promoting transcription antitermination. Finally, our study shows that T-Box riboswitches, which are essentially restricted to Gram-positive bacteria, are fully functional in a Gram-negative environment.

**REFERENCES**

1. K. Sheppard, J. Yuan, M. J. Hohn, B. Jester, K. M. Devine and D. Soll. (2008) From one amino acid to another: tRNA-dependent amino acid biosynthesis. \textit{Nucleic Acids Res} \textbf{36}, 1813-25
tRNA-dependent synthesis of Asn regulated by a T-Box riboswitch

2. H. D. Becker, J. Reinbolt, R. Kreutzer, R. Giege and D. Kern. (1997) Existence of two distinct aspartyl-tRNA synthetases in Thermus thermophilus. Structural and biochemical properties of the two enzymes. *Biochemistry* **36**, 8785-97

3. A. W. Curnow, K. Hong, R. Yuan, S. Kim, O. Martins, W. Winkler, T. M. Henkin and D. Soll. (1997) Glu-tRNA^Gln^ amidotransferase: a novel heterotrimeric enzyme required for correct decoding of glutamine codons during translation. *Proc Natl Acad Sci U S A* **94**, 11819-26

4. O. Nureki, P. O'Donoghue, N. Watanabe, A. Ohmori, H. Oshikane, Y. Araiso, K. Sheppard, D. Soll and R. Ishitani. (2010) Structure of an archaeal non-discriminating glutamyl-tRNA synthetase: a missing link in the evolution of Gln-tRNA^Gln^ formation. *Nucleic Acids Res* **38**, 7286-97

5. H. D. Becker, B. Min, C. Jacobi, G. Raczniai, J. Pelaschier, H. Roy, S. Klein, D. Kern and D. Soll. (2000) The heterotrimeric Thermus thermophilus Asp-tRNA(Asn) amidotransferase can also generate Gln-tRNA(Gln). *FEBS Lett* **476**, 140-4

6. D. L. Tumbula, H. D. Becker, W. Z. Chang and D. Soll. (2000) Domain-specific recruitment of amide amino acids for protein synthesis. *Nature* **407**, 106-10

7. H. D. Becker and D. Kern. (1998) Thermus thermophilus: a link in evolution of the tRNA-dependent amino acid amidation pathways. *Proc Natl Acad Sci U S A* **95**, 12832-7

8. H. D. Becker, H. Roy, L. Moulinier, M. H. Mazauric, G. Keith and D. Kern. (2000) Thermus thermophilus contains an eubacterial and an archaeabacterial aspartyl-tRNA synthetase. *Biochemistry* **39**, 3216-30

9. F. Leveque, P. Plateau, P. Dessen and S. Blanquet. (1990) Homology of lysS and lysU, the two Escherichia coli genes encoding distinct lysyl-tRNA synthetase species. *Nucleic Acids Res* **18**, 305-1

10. M. V. Saluta and I. N. Hirshfield. (1995) The occurrence of duplicate lysyl-tRNA synthetase gene homologs in Escherichia coli and other procaryotes. *J Bacteriol* **177**, 1872-8

11. S. Skouloubris, L. Ribas de Pouplana, H. De Reuse and T. L. Hendrickson. (2003) A noncognate aminoacyl-tRNA synthetase that may resolve a missing link in protein evolution. *Proc Natl Acad Sci U S A* **100**, 11297-302

12. K. M. Chang and T. L. Hendrickson. (2009) Recognition of tRNA^Gln^ by Helicobacter pylori GluRS2--a tRNA^Gln^-specific glutamyl-tRNA synthetase. *Nucleic Acids Res* **37**, 6942-9

13. J. Dubois, D. Qiu, J. Hitti, Y. I. Wolf, R. L. Tatusov, F. Sabathe, L. Doucette-Stamm, P. Soucaille, M. J. Daly, G. N. Bennett, E. V. Koonin and D. R. Smith. (2001) Genome sequence and comparative analysis of the solvent-producing bacterium Clostridium acetobutylicum. *J Bacteriol* **183**, 4823-38

14. F. J. Grundy, S. E. Hodil, S. M. Rollins and T. M. Henkin. (1997) Specificity of tRNA-mRNA interactions in Bacillus subtilis tyrS antitermination. *J Bacteriol* **179**, 2587-94

15. M. Wels, T. Groot Kormelink, M. Kleerebezem, R. J. Siezen and C. Francke. (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

16. A. Gutierrez-Preciado, T. M. Henkin, F. J. Grundy, C. Yanofsky and E. Merino. (2009) Biochemical features and functional implications of the RNA-based T-box regulatory mechanism. *Microbiol Mol Biol Rev* **73**, 36-61
tRNA-dependent synthesis of Asn regulated by a T-Box riboswitch

20. T. M. Henkin and F. J. Grundy. (2002) In Vitro Transcription Assay For T Box Antitermination System - Patent 7413856
21. H. Putzer, C. Condon, D. Brechemier-Baey, R. Brito and M. Grunberg-Manago. (2002) Transfer RNA-mediated antitermination in vitro. *Nucleic Acids Res* **30**, 3026-33
22. H. D. Becker, R. Giege and D. Kern. (1996) Identity of prokaryotic and eukaryotic tRNA(Asp) for aminoacylation by aspartyl-tRNA synthetase from Thermus thermophilus. *Biochemistry* **35**, 7447-58
23. P. Fechter, J. Rudinger, R. Giege and A. Theobald-Dietrich. (1998) Ribozyme processed tRNA transcripts with unfriendly internal promoter for T7 RNA polymerase: production and activity. *FEBS Lett* **436**, 99-103
24. H. Roy, H. D. Becker, M. H. Mazauric and D. Kern. (2007) Structural elements defining elongation factor Tu mediated suppression of codon ambiguity. *Nucleic Acids Res* **35**, 3420-30
25. L. Feustel, S. Nakotte and P. Durre. (2004) Characterization and development of two reporter gene systems for Clostridium acetobutylicum. *Appl Environ Microbiol* **70**, 798-803
26. J. F. Milligan, D. R. Groebe, G. W. Witherell and O. C. Uhlenbeck. (1987) Oligoribonucleotide synthesis using T7 RNA polymerase and synthetic DNA templates. *Nucleic Acids Res* **15**, 8783-98
27. J. F. Milligan and O. C. Uhlenbeck. (1989) Synthesis of small RNAs using T7 RNA polymerase. *Methods Enzymol* **180**, 51-62
28. M. Bailly, M. Blaise, H. Roy, M. Deniziak, B. Lorber, C. Birck, H. D. Becker and D. Kern. (2008) tRNA-dependent asparagine formation in prokaryotes: characterization, isolation and structural and functional analysis of a ribonucleoprotein particle generating Asn-tRNA(Asn). *Methods* **44**, 146-63
29. J. C. McDowell, J. W. Roberts, D. J. Jin and C. Gross. (1994) Determination of intrinsic transcription termination efficiency by RNA polymerase elongation rate. *Science* **266**, 822-5
30. T. M. Henkin. (2009) Analysis of tRNA-directed transcription antitermination in the T Box system in vivo. *Springer Protocols - Methods in Molecular Biology* **540**, 281-90
31. K. Thormann, L. Feustel, K. Lorenz, S. Nakotte and P. Durre. (2002) Control of butanol formation in Clostridium acetobutylicum by transcriptional activation. *J Bacteriol* **184**, 1966-73
32. J. W. Roos, J. K. McLaughlin and E. T. Papoutsakis. (1985) The effect of pH on nitrogen supply, cell lysis, and solvent production in fermentations of Clostridium acetobutylicum. *Biotechnol Bioeng* **27**, 681-94
33. J. L. Johnson. (1994) Similarity analysis of DNAs, p. 655-682. In. *P. Gerhardt, R. G. E. Murray, W. A. Wood, and N. R. Krieg* (ed.), *Methods for general and molecular bacteriology*. *American society for Microbiology*, Washington, D.C.
34. J. Sambrook, Fritsch, E. F., and Maniatis, T. (1989) Molecular cloning 2nd ed. *Cold Spring Harbor Laboratory, Cold Spring Harbor, New York*
35. M. Frechin, B. Senger, M. Braye, D. Kern, R. P. Martin and H. D. Becker. (2009) Yeast mitochondrial Gln-tRNA(Gln) is generated by a GatFAB-mediated transamidation pathway involving Arc1p-controlled subcellular sorting of cytosolic GluRS. *Genes Dev* **23**, 1119-30
36. M. Bailly, S. Giannouli, M. Blaise, C. Stathopoulos, D. Kern and H. D. Becker. (2006) A single tRNA base pair mediates bacterial tRNA-dependent biosynthesis of asparagine. *Nucleic Acids Res* **34**, 6083-94
37. N. J. P. t. Broeke-Smits, T. E. Pronk, I. Jongerius, O. Bruning, F. R. Wittink, T. M. Breit, J. A. G. v. Strijp, A. C. Fluit and C. H. E. Boel. (2010) Operon structure of Staphylococcus aureus. *Nucleic Acids Res* **38**, 3263-74
38. J. T. Heap, O. J. Pennington, S. T. Cartman, G. P. Carter and N. P. Minton. (2007) The ClosTron: a universal gene knock-out system for the genus Clostridium. *J Microbiol Methods* **70**, 452-64
39. P. Durre. (2008) Fermentative butanol production: bulk chemical and biofuel. *Ann N Y Acad Sci* **1125**, 353-62
40. H. Grupe and G. Gottschalk. (1992) Physiological Events in Clostridium acetobutylicum during the Shift from Acidogenesis to Solventogenesis in Continuous Culture and Presentation of a Model for Shift Induction. *Appl Environ Microbiol* **58**, 3896-902
41. T. M. Larsen, S. K. Boehlein, S. M. Schuster, N. G. Richards, J. B. Thoden, H. M. Holden and I. Rayment. (1999) Three-dimensional structure of Escherichia coli asparagine synthetase B: a short journey from substrate to product. *Biochemistry* **38**, 16146-57
42. J. Felton, S. Michaelis and A. Wright. (1980) Mutations in two unlinked genes are required to produce asparagine auxotrophy in Escherichia coli. *J Bacteriol* **142**, 221-8
43. S. Poggio, C. Domeinzain, A. Osorio and L. Camarena. (2002) The nitrogen assimilation control (Nac) protein represses asnC and asnA transcription in Escherichia coli. *FEMS Microbiol Lett* **206**, 151-6
44. H. Janssen, C. Doring, A. Ehrenreich, B. Voigt, M. Hecker, H. Bahl and R. J. Fischer. (2010) A proteomic and transcriptional view of acidogenic and solventogenic steady-state cells of Clostridium acetobutylicum in a chemostat culture. *Appl Microbiol Biotechnol* **87**, 2209-26
45. K. Sheppard and D. Soll. (2008) On the evolution of the tRNA-dependent amidotransferases, GatCAB and GatDE. *J Mol Biol* **377**, 831-44
46. M. Zuker. (2003) Mfold web server for nucleic acid folding and hybridization prediction. *Nucleic Acids Res* **31**, 3406-15
47. S. Griffiths-Jones, A. Bateman, M. Marshall, A. Khanna and S. R. Eddy. (2003) Rfam: an RNA family database. *Nucleic Acids Res* **31**, 439-41
48. M. S. Gerdeman, T. M. Henkin and J. V. Hines. (2002) In vitro structure-function studies of the Bacillus subtilis tyrS mRNA antiterminator: evidence for factor-independent tRNA acceptor stem binding specificity. *Nucleic Acids Res* **30**, 1065-72
49. D. Luo, C. Condon, M. Grunberg-Manago and H. Putzer. (1998) In vitro and in vivo secondary structure probing of the thrS leader in Bacillus subtilis. *Nucleic Acids Res* **26**, 5379-87
50. M. R. Yousef, F. J. Grundy and T. M. Henkin. (2005) Structural transitions induced by the interaction between tRNA(Gly) and the Bacillus subtilis glyQS T box leader RNA. *J Mol Biol* **349**, 273-87
51. M. Sprinzl and K. S. Vassilenko. (2005) Compilation of tRNA sequences and sequences of tRNA genes. *Nucleic Acids Res* **33**, D139-40
52. C. Sauter, B. Lorber, J. Cavarelli, D. Moras and R. Giege. (2000) The free yeast aspartyl-tRNA synthetase differs from the tRNA(Asp)-complexed enzyme by structural changes in the catalytic site, hinge region, and anticodon-binding domain. *J Mol Biol* **299**, 1313-24

Acknowledgments - We wish to thank C. Stathopoulos, from the University of Patras for critical reading of the manuscript.

FOOTNOTES

*This work was supported by the Association pour la Recherche sur le Cancer (ARC), the University of Strasbourg, the CNRS, ANR (ANR-09-BLAN-0091-02) and ARC. N. Saad was both a recipient of a fellowship from the Ministère de l’Education Nationale, de la Recherche et de la Technology - bourse de la Présidence de l’Université de Strasbourg, and a FEBS Summer Scholarship fellow.

To whom correspondence should be addressed: Pr. Hubert Dominique Becker, UMR ‘Génétique Moléculaire, Génomique, Microbiologie’, CNRS, Université de Strasbourg, 21 rue René Descartes, 67084 Strasbourg, France, Tel: 33 (0)3 68 85 14 70; Fax: 33 (0)3 68 85 13 65; E-mail: h.becker@unistra.fr and Dr. Bettina Schiel, Institut für Mikrobiologie und Biotechnologie,
tRNA-dependent synthesis of Asn regulated by a T-Box riboswitch

Universität Ulm, 89069 Ulm, Germany, Tel: 49 (0)731 50 22715; Fax: 49 (0)731 50 22719; E-mail: bettina.schiel@uni-ulm.de.

The abbreviations used are: Cac, *Clostridium acetobutylicum*; Tth, *Thermus thermophilus*; Hpy, *Helicobacter pylori*; Eco, *Escherichia coli*; Asn or N, asparagine; Asp, acid aspartic or aspartate; Gln, glutamine; Glu, acid glutamic or glutamate; Gly, glycine; Tyr, tyrosine; AdT or GatCAB, tRNA-dependent amidotransferase; AspRS, aspartyl-tRNA synthetase; AsnRS, asparaginyl-tRNA synthetase; AsnB or AS, asparagine synthetase; RNAP, RNA polymerase; Kd, dissociation constant; Ct, C-terminal; Nt, N-terminal.

**FIGURE LEGENDS**

**FIGURE 1.** The direct and indirect pathways for Asn and Asn-tRNA^Asn_ synthesis. A. Schematic representation of the direct pathway of Asn-tRNA^Asn_ synthesis. AsnB: asparagine synthetase B, AsnRS: asparaginyl-tRNA synthetase, AspRS1 bacterial-like aspartyl-tRNA synthetase. B. Schematic representation of the indirect pathway that usually compensates for the absence (red cross) of AsnRS or asparagine synthetase (AS). AspRS2: archaeal-like non-discriminating aspartyl-tRNA synthetase, GatCAB: bacterial trimeric tRNA-dependent amidotransferase. C. Increase in the number of protein partners when the indirect pathway is used, especially in the case of Cac. Based on the bacterium gene content, there is triplication in AspRSs (AspRS2o, AspRS2 and AspRS1), and duplication in GatCABs (GatCABo and GatCAB). Cac also encodes for two truncated asparagine synthetases (AsnB): Nt-AS and Ct-AS. D. Schematic representation of the operon organization of aspS2ogatCABo, regulated at the transcriptional level by NT-Box. The latter is found at the 5'-UTR of the operon. Asn may come from other sources then asparagine synthetase-dependent amidation of Asp.

**FIGURE 2.** Verification of the operon organization for aspS2ogatCABo and gatCAB genes using reverse transcription PCR (RT-PCR). Six RNA preparations corresponding to six growth conditions were used: MM: minimal medium without asparagine and aspartic acid, - Asp: MM in the presence of asparagine and absence of aspartic acid, - Asn: MM in the presence of aspartic acid and absence of asparagine. For the three conditions, RNAs were extracted from cells grown until acidogenesis (Ac.) or solventogenesis (S.). A. Verification of the presence of the NT-Box, using specific oligonucleotides (see materials and methods). The synthesized cDNA was of 72 bp. The same band size was obtained when genomic DNA (G) was used as a positive control. B. Verification of the presence of the NT-Box upstream of the aspS2ogatCABo mRNA. The same conditions were used as in (A.) and the amplified band size was of 207 bp. C. Verification of the operon arrangement between aspS2o and gatCABo. The same conditions were used as in A. and the amplified band size was of 167 bp. A PCR on RNA preparations was done without reverse transcription to check for DNA contamination (Control). D. Verification of the gatCAB operon arrangement. The same conditions were used as in (A.) and the amplified band size was of 175 bp. For B., C. and D., oligonucleotides were chosen to amplify the junction between two cistrons (see schematic representation).

**FIGURE 3.** Comparison of the growth curves of Cac WT and AspRS2 (CAC3564, *aspS2*) knock-out strains. The strains were grown on MM, with ammonium as the only nitrogen donor and in presence (+ Asn) and absence of Asn (- Asn). Cell density was optically measured at 600 nm.

**FIGURE 4.** Analysis of Cac Nt-AS and Ct-AS asparagine synthetase activities *in vitro* and *in vivo*. A. Complementation of the Asn auxotroph *E. coli* ER strain (*asnA*, *asnB*) by the Cac Nt- AS_pET15b and the Cac Ct-AS_pET15b constructs. The *E. coli* ER strain was transformed with either one of the two recombined pET15b vectors or with both recombinant vectors (Ct/Nt-
tRNA-dependent synthesis of Asn regulated by a T-Box riboswitch

AS_pET15b). Transformants were grown on minimal M9 medium agar plates supplemented with ampicillin and 0.5 mM IPTG in the absence (–) of Asn. B. Asparagine synthetase activity assay using Cac protein extracts (S100). Reactions were carried out using standard amidation mix (see materials and methods) and 100 µg of Sce (lane 1), Tth (lane 2), and Cac (lanes 3-8) S100. Six different Cac S100 were analyzed for their asparagine synthetase activities. Lanes 3-8: S100 taken from cells grown until acidogenesis (Ac.) or solventogenesis (S.) phase. -aa: no aa were added for the culture, +Asn: addition of 1 mM Asn, +Asp: addition of 1 mM Asp.

FIGURE 5. Cac GatCABo AdT catalyzes tRNA-dependent Asn and Gln formation. A. GatCABo tRNA-dependent Asn formation. Transamidation reaction done in the presence of GatCABo and absence of the amide group donor (glutamine: Gln) (lane 1), tRNAAsn (lane 2) and D. radiodurans ND-AspRS2 (lane 3). Lane 4: transamination reaction was carried out during 10 min at 37°C in the presence of GatCABo, D. radiodurans ND-AspRS2, tRNAAsn and 2 mM of Gln.. B. GatCABo tRNA-dependent Gln formation. Transamidation reactions were carried out as in A. using preformed H. pylori [14C]Glu-tRNA Glu and GatCABo (lanes 1 and 3), in the absence (lane 1) or presence (lane 3) of Gln. Lanes 2 and 4: control transamidation reactions without GatCABo (lane 2) or with H. pylori purified GatCAB (lane 4).

FIGURE 6. A T-Box is located in the 5'-UTR of aspS2ogatCABo. A. Model of the secondary structure of Cac aspS2ogatCABo T-Box (NT-Box). The sequence shown comprises the full length T-Box from the transcription start site (+1) through the stem I, containing the well conserved GA motif and AG box, as well as the specifier loop, which displays the (AAC) Asn codon sequence. The sequence also shows the stem II and stem III, which both form the non-conserved intermediate region of the T-Box. The sequence ends with the antiterminator region including the 14 nt conserved T-Box sequence. The alternate and more stable terminator structure is shown next to the antiterminator. Nucleotides marked with asterisks form a predicted Kink-turn (communication). B. Schematic description of the putative mechanism of regulation mediated by the NT-Box antitermination system, in response to Asn starvation or supply.

FIGURE 7. Specific binding of uncharged Cac tRNAAsn transcript to NT-Box, monitored by EMSA and size-exclusion chromatography. A. EMSA. Gel 1 shows the purified (gel-eluted) NT-Box transcript used in the experiments. L: ladder, bp: base-pair. EMSA analysis of the NT-Box•Cac tRNAAsn duplex formation (schematized on the right side of the figure) was performed as described in the Methods section. Only tRNA was 5’ [32P]-labeled. Lane 2: Cac tRNAAsp transcript, lane 3: Cac tRNAAsp transcript with 20 µM of NT-Box and lanes 4-12: Cac tRNAAsn transcript in the absence (lane 4) and presence of NT-Box transcript (lanes 5-12). Increasing concentrations of the NT-Box transcript from 0.078 to 20 µM (lane 5-12: 0.078, 0.5, 1, 2, 5, 10, 15 and 20 µM respectively) allowed the determination of a dissociation constant (Kd) for each of the two forms of the NT-Box•Cac tRNAAsn duplex indicated by arrows. Upper duplex form (black arrow) and lower duplex form (grey arrow). B. Analysis of the NT-Box•Cac tRNAAsn duplex formation by size-exclusion chromatography. Elution profiles of Cac tRNAAsn (5 µM) (△-), NT-Box transcript (5 µM) (-○-) and of a mixture of Cac tRNAAsn and NT-Box (5 µM each) (-□-) are shown. Insets represent denaturing PAGE analysis of the transcripts present in the peak fractions marked by dark arrowheads (►). Arrows on the elution profiles indicate the presence of two conformations of the NT-Box (-○-) or the NT-Box•Cac tRNAAsn duplex (-□-).

FIGURE 8. Design and analysis of the in vivo NT-Box-mediated antitermination assay in the Gram-negative E. coli bacterium. A. Schematic representation of the in vivo antitermination system in E. coli. The E. coli ER strain was co-transformed with the NT-Box_placZFT and Cac tRNAAsn_pKK223-3 recombinant plasmids. “Bac. P.” corresponds to the endogenous NT-Box promoter. “ptac” corresponds to the IPTG inducible promoter controlling Cac tRNAAsn gene.
expression. B. Schematic representation of the NT-Box-controlled lacZ expression in *E. coli* under the four conditions tested (see C.), using the two Cac NT-Box and tRNA^Asn^ (colored in black). The endogenous Eco tRNA^Asn^ (colored in light grey) has a “GUU” anticodon sequence. The tRNA^Asn^ colored in dark grey and binding to the NT-Box corresponds to the Eco and Cac uncharged tRNA^Asn^, which both can affect antitermination. T: Nt-Box terminator conformation. Anti-T: Nt-Box antiterminator conformation. C. *In vivo* NT-Box-mediated antitermination assay. Bars for each graph represent β-galactosidase activity in Miller units relative to cell density (Miller, 1972). The effect of Asn presence and tRNA induction were compared by taking 5 mL aliquots from each culture and by measuring the β-galactosidase activity after 4 hours of culture. The β-galactosidase activity value was calculated using the following equation: 1000×[(A_{420} - (1.75×A_{560})) / (T×v×A_{595})] where “A_{420}” is the absorbance of the ONP product (see materials and methods), “A_{560}” is the absorbance of the cell debris pellet, “A_{595}” is the absorbance of bacterial suspension, “T” is the reaction time in minutes, and “v” is the volume in “mL” of the treated cells used to measure β-galactosidase activity.
Figure 1.
Figure 2.

A. NT-Box transcript presence

B. NT-Box-aspS2o

C. aspS2o-gatCo

D. gatCA

Figure 3.

```
| Time (hours) | OD (600 nm) |
|--------------|-------------|
| 0            | 5           |
| 20           | 4           |
| 40           | 3           |
| 60           | 2           |
| 80           | 1           |
| 100          | 0           |

- WT/-Asn
- WT/+Asn
- AspS2 knock-out mutant/-Asn
- AspRS2 knock-out mutant/+Asn
- SD ≤ 5%
```
tRNA-dependent synthesis of Asn regulated by a T-Box riboswitch

Figure 4.

A.

- Asn/+ IPTG

B.

Cac S100

Figure 5.

A.

Asp

Asn

B.

Glu

Gln
Figure 6.

A. tRNA-dependent synthesis of Asn regulated by a T-Box riboswitch

B. Asp/Asn and tRNA$^{\text{Asn}}$ - dependent T-Box (NT-Box)
Figure 7.

A.

B.
Figure 8.

A.

B.

C.

| Asn   | tRNA induction | β-galactosidase activity/OD (MU) |
|-------|----------------|---------------------------------|
| -     | -              | 1                               |
| -     | +              | 2                               |
| +     | -              | 3                               |
| +     | +              | 4                               |
Riboswitch (T-Box)-mediated control of tRNA-dependent amidation in clostridium acetobutylicum rationalizes gene and pathway redundancy for asparagine and asparaginyl-tRNAAsn synthesis
Nizar Y. Saad, Bettina Schiel, Melanie Braye, John T. Heap, Nigel P. Minton, Peter Duerre and Hubert Dominique Becker

*J. Biol. Chem.* published online April 13, 2012

Access the most updated version of this article at doi: [10.1074/jbc.M111.332304](http://10.1074/jbc.M111.332304)

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

Supplemental material:
http://www.jbc.org/content/suppl/2012/04/13/M111.332304.DC1