SUPPLEMENTARY INFORMATION

Maturation of atypical ribosomal RNA precursors in *Helicobacter pylori*

Isabelle Iost*, Sandrine Chabas and Fabien Darfeuille

ARNA Laboratory, Inserm U1212, CNRS UMR 5320, Université de Bordeaux, France

* To whom correspondence should be addressed.

This supplement contains:

  Table S1
  Figures S1 to S9
Table S1: List of oligonucleotides

| Name   | Sequence 5' -> 3' | Description                                                                 |
|--------|------------------|-----------------------------------------------------------------------------|
| FD378  | GGAATTCATATGAAAAACACGCTCATAAAC (NdeI site underlined)                     | Cloning of rnc                                                             |
| FD380  | CGC TTCATATGCGGTGTTGTAATTATAGACATCCAATGCG (BamHI site underlined)        | Cloning of rnc                                                             |
| FD533  | AGCTTCTCTACTGCTGTTGTTAAACCG (T→C mutation of the asRNA promoter is underlined, complementary to FD622) | Mutagenesis of the asRNA promoter                                           |
| FD640  | GGATCGTTCCATATGCGGTGTTGTAATTATAGACATCCAATGCG (KpnI and BamHI sites are underlined; a T→C mutation was introduced to avoid transcription from the opposite strand of the cloned asRNA) | Deletion of the 23S-5S locus (kan<sup>R</sup> cassette)                    |
| FD660  | GGATCGTTCCATATGCGGTGTTGTAATTATAGACATCCAATGCG (KpnI and BamHI sites are underlined; a T→C mutation was introduced to avoid transcription from the opposite strand of the cloned asRNA) | Deletion of the 23S-5S locus (kan<sup>R</sup> cassette)                    |
| FD690  | GGAATTCATATGCGGTGTTGTAATTATAGACATCCAATGCG (KpnI and BamHI sites are underlined; a T→C mutation was introduced to avoid transcription from the opposite strand of the cloned asRNA) | Deletion of the 23S-5S locus (kan<sup>R</sup> cassette)                    |
| FD689  | GGAATTCATATGCGGTGTTGTAATTATAGACATCCAATGCG (KpnI and BamHI sites are underlined; a T→C mutation was introduced to avoid transcription from the opposite strand of the cloned asRNA) | Deletion of the 23S-5S locus (kan<sup>R</sup> cassette)                    |
| FD708  | GGAATTCATATGCGGTGTTGTAATTATAGACATCCAATGCG (KpnI and BamHI sites are underlined; a T→C mutation was introduced to avoid transcription from the opposite strand of the cloned asRNA) | Deletion of the 23S-5S locus (kan<sup>R</sup> cassette)                    |
| FD718  | GGAATTCATATGCGGTGTTGTAATTATAGACATCCAATGCG (KpnI and BamHI sites are underlined; a T→C mutation was introduced to avoid transcription from the opposite strand of the cloned asRNA) | Deletion of the 23S-5S locus (kan<sup>R</sup> cassette)                    |
| FD622  | GCCTTTGAGCGTTTTATGG                                | Amplification of the kan<sup>R</sup> cassette |
| FD255  | CAGGTACTAAAAACATTCATCC                            | Amplification of the apra<sup>5</sup> cassette |
| FA302  | CAATACGAATGCGAAAAAGCCG                            |                                       |
| FA304  | TCACGCCAATCAGCTGCGGAG                            |                                       |
| FD371  | GACTAAAAATGCTTGACCCCTATTC (probe ‘a’ to detect pre-23S) | Probes for Northern |
| FD359  | CTTCAGTGCTACCAGCTGCTT (probe ‘b’ to detect pre-23S) |                                       |
| FA357  | GCTTGTTTTTTTTTTTTTT (probe ‘b’* to detect pre-23S) |                                       |
| FA167  | TCGGAATGGTTAACGGGTAG (probe ‘c’ to detect 5S rRNA) |                                       |
| FD397  | GCTATCCTTGTTAGGTATCAAG (probe ‘f’ to detect pre-16S) |                                       |
| FD399  | GCTTTAGCTATCACAAGAAGCTTC (probe ‘g’ to detect pre-16S) |                                       |
| FA156  | ATGAATATAAGCACAGAGTTATG (probe ‘h’ to detect pre-23S) |                                       |
| FD372  | GATAGGGATCAAGCATTTTTAGTC (probe ‘asRNA’*) |                                       |
| FD500  | CTTCCTTTAAGGTTTAAC (probe ‘asRNA-5’*) |                                       |
| FD373  | GCATAGCTTATCAGCCAGTCTAGT (primer ‘d’) |                                       |
| FA168  | TTTAACAAAGAAGATATTTAGTA (primer ‘e’) |                                       |
| FA203  | GTAAACCTTCTTAAGAAAGAGC (primer ‘i’) |                                       |
| FA197  | GAAATTTAATAGCACTCAATAGGTAAACCCTTTAAAGAGAC (the T7 promoter sequence is underlined) | Oligos for primer extension |
| FA234  | ACATTTGCTTTTTAAAGCAGAC |                                       |
| FA88   | GAAATTTAATAGCACTCAATAGGTAAACCCTTTAAAGAGAC (the T7 promoter sequence is underlined) |                                       |
| FA156  | ATGAATATAAGCAGAGTTATG |                                       |
| FA88   | GAAATTTAATAGCACTCAATAGGTAAACCCTTTAAAGAGAC (the T7 promoter sequence is underlined) |                                       |
| FD373  | GATAGCTTATCAGCAGTCTAGT |                                       |
| FA193  | ACTAGACGTGATAAGCTATGCGGTGTTGTAGGTAGATGTG (underlined sequence is complementary to FD373) |                                       |
| FA194  | GATAACGAGCTAAAGATAAACG |                                       |
| FA310  | GAAATTTAATAGCACTCAATAGGTAAACCCTTTAAAGAGAC (the T7 promoter sequence is underlined) |                                       |
| FA234  | ACATTTGCTTTTTAAAGCAGAC |                                       |
| FA311  | GAAATTTAATAGCACTCAATAGGTAAACCCTTTAAAGAGAC (the T7 promoter sequence is underlined) |                                       |
| FA156  | ATGAATATAAGCAGAGTTATG |                                       |
Supplementary figures

Figure S1: Supplementary data related to Fig. 1

Fig. S1A: Location of the rRNA genes on the *H. pylori* B128 genome.

The B128 genome (NCBI Reference Sequence: NZ_CP019700) contains two copies of each rRNA gene (19, 20). The HPB8-r1 and HPB8-r2 genes encoding the 16S rRNA are located at nt 97,496-99,004 and 323,581-325,089, respectively; the HPB8-r3 and HPB8-r6 genes encoding the 23S rRNA are located at nt 448,705-451,593 and 1,529,464-1,532,352, respectively and the HPB8-r4 and HPB8-r5 genes encoding the 5S rRNA are located at nt 451,833-451,950 and 1,529,107-1,529,224, respectively.

The 23S-5S rRNA operons and the 16S rRNA genes (denoted as gray arrows) lie far apart. While the closest 16S and 23S genes are separated by ~123,600 bp, the two other copies are more distant and in opposite orientations. In this manuscript (Fig.6, Fig. S6 and Fig. S8C), the 23S-5S copy located downstream position 400kb will be named ‘*rrn1*’, whereas the other copy will be named ‘*rrn2*’.

Fig. S1B: Representative growth curves of *H. pylori* wt and ∆*rnc* B128 strains.

The wt and ∆*rnc* B128 strains were grown in BHI + SVF medium at 37°C. The absorbance at 600nm (A\textsubscript{600nm}) was measured at different time points. The doubling times were estimated to be 2.7 h for the wt (blue curve) and 3h for ∆*rnc* (red curve).
Fig. S1C: Transcriptional start sites (TSS) of the 23S-5S rRNA precursor and of the cis-encoded antisense RNA (asRNA).

The screenshots of RNA-seq data from IGB show the cDNA coverage plots for the forward (top) and reverse (bottom) strands above the rRNA leader region of the 23S-5S operon. The blue and red bars correspond to wild-type and \(\Delta rnc\) libraries, respectively. The TATAAT promoter sequences and the TSS (« +1 ») of the 23S-5S precursor and of the asRNA are indicated. The green arrows indicate two RNase III cleavages that are described in Fig. 8.
Fig. S1E: Expressing RNase III from a plasmid corrects the \( H. \) pylori \( \Delta rnc \) defects.

The B128 \( \Delta rnc \) strain was transformed with the plLL2157-\( rnc \) plasmid in which the \( rnc \) ORF was cloned downstream a \( ureI \)-derived promoter (22). As a control, the strain was also transformed with the empty plLL2157 plasmid. RNA was extracted from B128 wt, \( \Delta rnc \) and plLL2157 containing \( \Delta rnc \) strains. Eight \( \mu \)g of total RNA were separated on a 1% agarose gel and stained with ethidium bromide. The mature 23S and 16S rRNAs and the p1 and p2 rRNA precursors are indicated.
Fig. S2: Screenshot of RNA-seq data in the 5S rRNA region.
The green arrows indicate the RNase III cleavage sites upstream of mature 5S rRNA. The cleavage located 4 nt upstream of the mature 5’ end (‘M’) was confirmed by primer extension (Fig. 2A, primer ‘c’). Symbols are in Fig. S1C.

Fig. S3: Screenshot of RNA-seq data in the leader and 3’ end regions of 16S rRNA.
The TSS of the 16S rRNA precursor (‘+1’), the promoter sequence (TATACT) and the RNase III cleavage sites (green arrows) are indicated. Symbols are as in Fig.S1C.
Figure S4 : Unprocessed 5S rRNAs are found in polysomes.

RNA extracted from wt and Δrnc polysome profiles (Fig. 4A) were separated on a denaturing 8% polyacrylamide gel and analyzed by Northern blot. The probe ‘c’ (see Fig. 1A) was used to reveal 5S rRNA. The fraction samples are the same than the ones used in Fig. 4. However, in this experiment, equal volumes (instead of equal amount of total RNA) were loaded on the gel. The low signal observed for the second fraction of the 50S peak is due to a lower recovery of RNA.

The first lanes of each panel (« E ») correspond to the extract before fractionation on sucrose gradient. The top of the gradient and the 30S, 50S, 70S and polysomes are indicated. The main species detected in both strains corresponds to the 5S mature form (« M »). Larger 5S precursors (p5S) that accumulate in the Δrnc strain are indicated. The M+4 and M+1/+2 species correspond to 5S precursors containing either 4 or 1 and 2 additional nt at the 5’ end in wt and Δrnc strains, respectively (see Fig. 2A primer ‘c’).
Fig S5A: The main asRNA species share the same 5’ end.

The B128 ∆rnc strains transformed either with the pILL2157-rnc (lane 1) or pILL2157 (lane 2) plasmid were grown in BHI + SVF + chloramphenicol medium until A\textsubscript{600nm} reached ~1. Total RNA was extracted and analyzed as in Fig.5. The membrane was hybridized with the asRNA probe (used also for Northern in Fig. 5, 6, 7, S5C and S8C) and then with the asRNA-5’ probe, which is complementary to the first 20 nt of the transcript. The location of the two probes (blue and orange small boxes) is indicated. A 5S probe was used to control the loading.

Three main transcripts of ~175, 130 and 70 nt (depicted as horizontal arrows) are detected at low levels in the ∆rnc complemented strain (lane 1), as observed for the B128 wt strain (see Fig. 5). In contrast, the ~175 and 130 nt transcripts accumulate in absence of RNase III (lane 2). All these transcripts are detected by the asRNA-5’ probe, indicating that they start at the TSS of the asRNA.
Fig. S5B: Conservation of the asRNA promoter in various *H. pylori* strains.

The first 300 nt of the 23S-5S rRNA leader from B128 strain was used to retrieve homologous sequences using BlastN. Multiple alignment of leader sequences from diverse *H. pylori* strains was computed using MultAlin (1). “+1” corresponds to the transcription start site of the primary transcript. The asRNA promoter is highlighted within a box. Fully, partially and poorly conserved nucleotides are indicated in red, blue and black, respectively.

(1) Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering. *Nucleic Acids Res.*, 16, 10881–10890.
Fig. S5C: The deletion of RNase III leads to stabilization of the asRNA in both B128 and X47-2AL *H. pylori* strains.

RNA was extracted from B128 and X47-2AL strains and their ∆*rnc* derivatives. Two independent ∆*rnc* X47-2AL clones were used (∆*rnc*1 and ∆*rnc*2). Northern blot analysis was performed as described in Fig. 5. The vertical line means that lanes from the same gel have been brought together at this position.
**Fig S6**: The single point mutation in the *rrn1* promoter impairs transcription.

A. The *rrn1* and *rrn2* are arbitrary names for the two 23S-5S operons (see Fig. S1A).

B. Total RNA extracted from BC36 and BC41 strains were separated on an 1% agarose gel and stained with ethidium bromide. In the BC36Δ*rnc* strain, the large p1 and p2 rRNA precursors accumulate, as observed for a B128Δ*rnc* strain (see Fig. 1B). In the BC41Δ*rnc* strain, the p1 precursor is no more visible, indicating that the mutation of the *rrn1* promoter inhibits transcription. Instead, a truncated precursor form (p1Δ) that is transcribed from the *rrn2* promoter, accumulates. This species lacks the 190 nt corresponding to the deletion that removes both the promoter and the first 160 nt of the asRNA sequence. The amount of the p2 precursor is not affected in the BC41Δ*rnc* strain, indicating that its production is not dependent on the cis-encoded asRNA.

C. The membrane used in Fig. 6 has been hybridized with the probe 'a' to verify the absence of transcription from the mutated *rrn1* promoter. The position of the probe is shown by a small black box in (A). In strain BC36 *rnc*+, two main transcripts of ~280 and ~210 nt are detected (lane 1). In the Δ*rnc* background, the 280 nt transcript is no more visible but larger species of ~300 and ~350 nt accumulate (lane 3). In the BC41 strain, no transcript is visible in the *rnc*+ background (lane 2) whereas a faint signal is observed in the Δ*rnc* background (lane 4) indicating that the transcription from the *rrn1* locus is severely but not completely inhibited.
Fig. S7A: Structure probing of the asRNA/rRNA precursor complex.

A. The 5’ end-labeled asRNA was incubated without (lanes ‘-’) or with (lanes ‘+’) an excess of unlabeled rRNA precursor (p280) and subjected to enzymatic (T1 and TA) or lead (Pb²⁺) hydrolysis. For lead-induced cleavage, complex formation was performed at 37°C for 10-60 minutes. ‘Ctl’ lane: asRNA alone; ‘N’ lane: alkaline hydrolysis; ‘G’ lane: T1 digestion under denaturing conditions; lanes ‘T1’ and ‘TA’: digestion under native conditions. RNase T1 and TA cleave single-stranded RNA after guanosine and adenosine residues, respectively. Positions with increased reactivity are indicated on the right of the gel.

B. Reciprocal experiment of (A). The rRNA precursor (p280) was 5’ end-labeled and incubated without or with an excess of unlabeled asRNA.
**Fig. S7B: Predicted secondary structure of the asRNA.**

The secondary structure was predicted using Mfold, in agreement with the experimental data obtained by enzymatic footprinting, and visualized with the VARNA applet (1). The ‘g’ in lower-case letters correspond to guanine cleaved by RNase T1 in Fig. S7A.

(1) Darty, K., Denise, A. and Ponty, Y. (2009) VARNA: Interactive drawing and editing of the RNA secondary structure. *Bioinformatics*, **25**, 1974–1975.
Fig. S8A: Overexpression of the asRNA does not affect the amount of 23S rRNA.

The B128 strains transformed with either the asRNA overexpressing vector (pILL-asRNA) or the control vector (pILL) were grown in BHI + SVF + chloramphenicol medium at 37°C. Total RNA was extracted and 0.4 µg was separated on an 1% agarose gel and stained with ethidium bromide.

The different panels correspond to RNA samples derived from independent experiments. The A_{600nm} of the culture at which RNA was extracted is indicated below the gels. The 23S and 16S rRNAs are indicated.
B128 wt and ∆rnc strains transformed with either the asRNA overexpressing vector (pILL-asRNA) or the empty plasmid (pILL) were grown in BHI + SVF + chloramphenicol medium at 37°C until an A_{600nm} of ~0.75 - 1.0. Total RNA was extracted and subjected to primer extension analysis as in Fig. 2. The dotted vertical line indicates different exposition times of the same gel. The location of the primers used is shown on the map of the the 23S rRNA leader region. A strong signal at position 166-167 (indicated by an asterisk) observed with the primer ‘h’ is likely due to the arrest of the RT enzyme by the duplex. Indeed, this stop increases when the asRNA is overexpressed and is enhanced in the ∆rnc strain. The other arrows indicate RT stops that are significantly increased when the asRNA is overexpressed and which are absent in the ∆rnc strain. These stops may correspond to RNase III cleavages induced by the asRNA. The cleavages at positions 106 and 161 producing the p75 and p122 precursor fragments (see Fig. 8) are indicated. The cleavage at nt 106 results in a shift of the RNAseq reads (Fig. S1C). A shift 2 nt upstream is also observed on the other strand and likely corresponds to the cleavage of the asRNA leading to the 70 nt product.
Figure S8C: Inhibition of asRNA transcription abolishes specific processing cleavages in the rRNA precursor.

(a) Schematic view of the 23S-5S rRNA operon from BΔ23S-5S-29 and BΔ23S-5S-30 strains. The BΔ23S-5S-30 strain carries a mutation in the asRNA promoter (TATAAT → TATAAC), whereas the isogenic BΔ23S-5S-29 strain contains the wt promoter. Both strains contain only one 23S-5S rRNA operon (‘rrn1’ copy has been deleted, see Fig.S1A for the location of the rrn genes). The location of the probes ‘a’ and ‘h’ is indicated. Symbols are as in Fig. 1A.

(b) RNA was extracted from both strains and Northern analysis was performed as described in Fig. 5. The 5’ end-labelled probes were specific to the asRNA (left) and to two different regions of the rRNA precursor (probes ‘a’ and ‘h’). The same membrane was used for the different probes. As a control, we also used a B128 strain containing only one rrn copy, like the BΔ23S-5S-29 and BΔ23S-5S-30 strains, but lacking the cmR resistance cassette (“Ctrl” lane).
Fig. S8D: The p75 and p122 processed species are absent in a Δrnc strain.

The B128 wt and Δrnc strains transformed with either the asRNA overexpressing vector (pILL-asRNA) or the empty plasmid (pILL) were grown in BHI + SVF + chloramphenicol medium at 37°C until an $A_{600\text{nm}}$ of ~0.8 – 0.9. Total RNA was extracted and analyzed by Northern blot as in Fig. 5. The precursor specific probes ‘a’ and ‘h’, which are located at different positions of the leader region (see panel S8C) revealed the asRNA-induced p75 and p122 products, respectively. These products were absent in the Δrnc strain indicating that they are generated by RNase III. The saturated pixels are highlighted in red. The vertical lines mean that the lanes from the same gel were brought together at this position.
Fig. S9A: The absence of the asRNA does not affect the growth rate.

The strains BΔ23S-5S-29 (‘wt’) and BΔ23S-5S-30 (‘Δas’) were grown in BHI + SVF medium at 37°C. The ‘Δas’ strain contains a point mutation in the asRNA promoter that inhibits its transcription (see Fig. S8C). No significant difference in the growth rate was observed whether the strain expresses the asRNA (‘wt’, green curve) or not (‘Δas’, pink curve).

Fig. S9B: The absence of the asRNA does not affect the amount of rRNA.

Total RNA from strains BΔ23S-5S-29 (‘wt’) and BΔ23S-5S-30 (‘Δas’) was extracted and 0.4 µg was separated on an 1% agarose gel and stained with ethidium bromide. The different panels correspond to RNA samples derived from independent experiments. The A_{600nm} of the culture at which RNA was extracted is indicated below the gels. The 23S and 16S rRNAs are indicated.