RESEARCH

Ribosome pausing in amylase producing *Bacillus subtilis* during long fermentation

Yaozu Han¹, Biwen Wang¹, Alberto Agnolin¹, Gaurav Dugar¹, Frans van der Kloet¹, Christopher Sauer², Paul Igor Costea², Max Fabian Felle², Mathis Appelbaum² and Leendert W. Hamoen^{1*}

Abstract

Background Ribosome pausing slows down translation and can affect protein synthesis. Improving translation efficiency can therefore be of commercial value. In this study, we investigated whether ribosome pausing occurs during production of the α -amylase AmyM by the industrial production organism *Bacillus subtilis* under repeated batch fermentation conditions.

Results We began by assessing our ribosome profiling procedure using the antibiotic mupirocin that blocks translation at isoleucine codons. After achieving single codon resolution for ribosome pausing, we determined the genome wide ribosome pausing sites for *B. subtilis* at 16 h and 64 h growth under batch fermentation. For the highly expressed α -amylase gene *amyM* several strong ribosome pausing sites were detected, which remained during the long fermentation despite changes in nutrient availability. These pause sites were neither related to proline or rare codons, nor to secondary protein structures. When surveying the genome, an interesting finding was the presence of strong ribosome pausing sites in several toxins genes. These potential ribosome stall sites may prevent inadvertent activity in the cytosol by means of delayed translation.

Conclusions Expression of the α-amylase gene *amyM* in *B. subtilis* is accompanied by several ribosome pausing events. Since these sites can neither be predicted based on codon specificity nor on secondary protein structures, we speculate that secondary mRNA structures are responsible for these translation pausing sites. The detailed information of ribosome pausing sites in *amyM* provide novel information that can be used in future codon optimization studies aimed at improving the production of this amylase by *B. subtilis*.

Keywords Ribosome profiling, Ribosome pausing, Protein secretion, Bacillus subtilis, Amylase

*Correspondence:

Leendert W. Hamoen

l.w.hamoen@uva.nl

¹Swammerdam Institute for Life Sciences, University of Amsterdam, Science Park 904, Amsterdam 1098 XH, The Netherlands ²White Biotechnology Research, BASF SE, Carl-Bosch-Strasse 38,

Ludwigshafen am Rhein 67056, Germany



© The Author(s) 2025. **Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by-nc-nd/4.0/.



Open Access

Background

Elongation rates of ribosomes vary considerably and depend on codon usage, amino acid availability, amino acid sequences, and mRNA structures. For example, rare codons have been implicated in fine-tuning translational rates in order to favor proper protein folding [1-4], and the amino acid sequence of the nascent polypeptide can alter elongation rates by interacting with the exit tunnel of the ribosome [5–8]. In addition, stable mRNA structures can pause elongation, and can trigger rescue pathways associated with ribosome stalling and arrest [9]. Ribosome pausing due to suboptimal codon usage is often regarded as an important reason for low expression of heterologous proteins [10]. However, in many cases codon optimization does not result in better expression (e.g. [11]). Understanding the correlation between high and low frequency codons and ribosome pausing might help optimize protein production. Ribosome profiling provides a powerful technique to identify translational pausing and stall sites, and hence can help optimize protein production [12]. Here, we used ribosome profiling to investigate ribosome pausing during secretory production of the industrially relevant α -amylase AmyM in the bacterial production organism *Bacillus subtilis*, during a 64-hour long repeated batch fermentation, with the purpose to optimize future industrial enzyme production.

The α -amylase AmyM from the thermophile Geobacillus stearothermophilus hydrolyzes the internal α -1,4 glycosidic bonds of starch [13], and has applications in many industrial processes, including brewing, baking and textiles industries, and detergent formulations [14, 15]. Microbial production is the predominant method to obtain large amounts of α -amylase, and *B. subtilis* is one of the main industrial workhorses, because of its high capacity to secrete enzymes [16], its GRAS (Generally Recognized As Safe) status, and simplicity of cultivation in large fed-batch fermentations [17, 18]. In this study we focused on the following research questions: Are there ribosome pause sites in the *amyM* mRNA, and do they change over the course of a long repeated batch fermentation? Are pause sites related to rare codons and can they be predicted based on sequence motifs? And can genome-wide ribosome profile data of the host provide insights into the cellular physiology during fermentation?

Methods

Strains and growth conditions

Strains, plasmids, primers and relevant materials used in this study are listed in table S1. Plasmid pBW212 carrying amyM under control of the *aprE* promoter. The plasmid contains a kanamycin selection marker. For the ribosome profiling experiment, plasmid pBW212 was transformed into the production strain 1S145, which is a protease-lacking derivative of *B. subtilis* wild type background,

resulting in strain BY212. Fresh transformants of BY212 were streaked out on LB plates containing 20 μ g/ml kanamycin at 37 °C. The next day, cells were scraped from the plate and resuspended in 0.9% NaCl and subsequently diluted into 40 ml B3 medium containing 20 μ g/ml kanamycin to an OD₆₀₀ of approximately 0.25, and grown at 30 °C for 88 h. The cultures were grown under vigorous shaking in 250 ml Erlenmeyer flasks. The B3 medium is a rich medium containing 20 g/L glucose, 1 g/l yeast extract and 12 g/L potassium glutamate as carbon and nitrogen sources. Extra 1% (final concentration) glucose was added to the cultures after 16, 40 and 64 h growth.

Inactivation of epeX, sdpC, skfA and ykoJ

All primers used for cloning are listed in table S1. BYH43 was constructed by deleting the chromosomal epeX, sdpC and skfA gene from strain B. subtilis 1S145. Briefly, a B. subtilis gene knockout library (BKE collection) with an erythromycin-resistance cassette inserted into non-essential genes was used as template for most gene deletions [19]. The epeX::Em region was PCR amplified from chromosomal DNA of strain BKE40180 using primers YH131 and YH132, and transformed into competent 1S145 cells. epeX removal was verified by PCR, and yielded strain BYH38 (*DepeX::Em*). After that, a chloramphenicol resistance marker was PCR amplified from plasmid pBW29, using primers BW327 and BW330, and transformed into BYH38 to replace the erythromycin selection marker with a chloramphenicol selection marker, resulting in strain BYH39 ($\Delta epeX::Cm$). The sdpC::Em region was PCR amplified from chromosomal DNA of strain BKE33770 using primers YH133 and YH134, and transformed into strain BYH39 to yield strain BYH40 ($\Delta epeX::Cm \ \Delta sdpC::Em$). The erythromycin resistance marker was replaced by amplifying the spectinomycin resistance marker from plasmid pBW30 using primers BW327 and BW330, and transforming into strain BYH40, yielding strain BYH41 ($\Delta epeX::Cm$ $\Delta sdpC::Spc$). Finally, the *skfA::Em* region was PCR amplified from chromosomal DNA of strain BKE01910 using primers YH135 and YH136, and transformed into strain BYH41, yielding strain BYH42 (ΔepeX::Cm ΔsdpC::Spc $\Delta skfA::Em$). The plasmid BW212 was transformed into strain BYH42, resulting in the amylase production strain **BYH43**.

The *ykoJ* deletion in the BKE collection appeared to be incorrect and was therefore made by ligating the *ykoJ*-upstream region, a chloramphenicol selection marker, and the *ykoJ*-downstream region by overlap PCR using primers YH137 and YH142. After purification, the fused fragment was transformed into competent 1S145 cells, yielding strain BYH44 ($\Delta ykoJ$::*Cm*). The plasmid BW212 was transformed into strain BYH44, resulting in amylase production strain BYH45.

Amylase activity assay

Samples were collected at 16 h, 40 h, 64 h and 88 h to test the amylase activity. Amylase levels in the medium were assayed using Amylase test tablets (Phadebas) made of starch polymers carrying a blue dye, according to the manufacturer's protocol. Hydrolysis by α -amylase releases a water-soluble blue chromophore that absorbs light at 620 nm.

Ribosome profiling

Ribosome profiling was based on the procedure described in [20]. We uses RNase AWAY (Thermo Fisher) and 70% ethanol to clean the surface of instruments, gloves and bench before manipulations of RNA. Milli-Q water was double autoclaved to remove RNase and was used for buffer and sucrose solution preparation and pellet resuspension. Collection of cell material was performed as follows. 40 ml cultures were poured into 2 Falcon tubes (50 ml) containing each 20 ml crushed frozen PBS buffer and 1 mM (323 µg/ml) chloramphenicol. Cells were harvested by centrifugation at 10,000 rpm for 3 min at 4 °C and were then frozen in liquid nitrogen. The frozen cell pellets were transferred into a 50 ml stainless steel grinding jar (25 mm grinding ball; Retsch) together with 2 ml lysis buffer. The lysis buffer contains 100 mM NH₄Cl, 10 mM MgCl₂, 20 mM Tris pH 8.0, 0.4% Triton X-100, 0.1% NP-40, 5 mM CaCl₂, 1 mM chloramphenicol. The culture was cryogenically pulverized by using the mixer mill MM400 (Retsch) with 8 cycles of 2 min at a frequency of 20 1/s and 1 min cooling between cycles in liquid nitrogen. Pulverized cells were thawed at room temperature, 50 µl DNase I was added and incubated on ice for 5 min. The lysate was clarified by centrifugation at 15,000 rpm for 10 min at 4 °C. The clarified supernatant was collected and used immediately for ribosome isolation and total RNA isolation.

The clarified lysates were pelleted by ultracentrifugation over a 8 ml sucrose cushion (20% sucrose, 100 mM NH₄Cl, 10 mM MgCl₂, 20 mM Tris pH 8.0, 0.5 mM EDTA, 0.4% Triton X-100, 0.1% NP-40, 1 mM chloramphenicol) in OptiSeal polypropylene tubes (Beckman) using a Ti-60 rotor at 50,000 rpm for 2 h at 4 °C. Ribosome pellets were resuspended in 200 µl resuspension buffer (100 mM NH₄Cl, 10 mM MgCl₂, 20 mM Tris pH 8.0) and the concentration was determined by Nanodrop (Thermo Scientific) using a 1:10 dilution in doubleautoclaved Milli-Q water. 1 mg of RNA was digested with 16,000 gel units of micrococcal nuclease (NEB) for 30 min at 37 °C and shaking at 1,400 rpm in a thermomixer (Eppendorf). The reaction was quenched with 10 mM EGTA. Sucrose gradients (10 - 50% sucrose in 100 mM NH₄Cl, 10 mM MgCl₂, 20 mM Tris pH 8.0, 2 mM DTT) were prepared by sequentially filling Open-Top thin-wall polypropylene tubes (Beckman) with 0.9 ml of 10-50% sucrose solutions (4.5 ml in total) from the bottom using a sterile syringe with a hypodermic needle. Digested samples were loaded onto the ice-cooled 10-50% sucrose gradients, and spun using an SWTi-55 rotor at 42,000 rpm for 2.5 h at 4 °C. After centrifugation, 0.9 ml of the 10% and 20% sucrose fractions were removed by pipetting carefully along the wall from the top. 0.9 ml of the 30%, 40% and 50% sucrose fractions were collected and used for RNA extraction, performed by adding 0.9 ml pre-warmed (65 °C) phenol/chloroform/isoamyl alcohol (P/C/I) (Carlroth), incubating for 5 min at 65 °C at 1100 rpm in a thermomixer, then incubating the tubes on ice for an additional 5 min, spinning at 13,200 rpm for 5 min at 4 °C, and transferring 0.7 ml of the upper aqueous layer to fresh tubes. aqueous fractions were further purified by adding 0.7 ml P/C/I at room temperature, shaking by hand for 1 min, spinning at 13,200 rpm for 5 min at 4 °C, and collecting 0.5 ml of the upper aqueous layer. The RNA samples (0.5 ml) were precipitated by adding 0.625 ml isopropanol and 20 µl 3 M sodium acetate and stored at -80 °C overnight. RNA was pelleted by centrifugation at 15,000 rpm for 0.5 h at 4 °C, and after removing the supernatant, the pellets were washed with 0.8 ml ice-cold 70% ethanol with careful pipetting, spun at 15,000 rpm for 10 min to remove ethanol, and shortly spun at 15,000 rpm for 0.5 min to discard the last droplets of ethanol. After air drying for 10 min at room temperature, the RNA was resuspended in 8 µl double-autoclaved Milli-Q water. Ribosome protected fragments (RPFs) were isolated by electrophoresis on a 15% TBE 7 M Urea PAGE gel with 1.5 mm-thick spacers, using oligonucleotides of 18 nt, 22 nt and 34 nt in length as size markers. A maximum of 6 µl RNA samples mixed with 2x RNA loading dye (NEB) was loaded per lane, and electrophoreses was performed at 60 V for 20 min and then at 180 V for 1 h. The staining solution was prepared by adding 1 µl Sybr-Gold nucleic acid gel stain (Thermo Fisher) to 30 ml double-autoclaved Milli-Q water. The gel was transferred to a transparent plastic tray containing the staining solution and incubated for 1-2 min with shaking (90-100 rpm). The gel area between 22 and 34 nt was cut out with a clean razor blade and crushed in a 2 ml tube. The further isolation of RPF fragments from PAGE gels, the preparation of the RNA control samples, and the construction of cDNA libraries for next generation sequencing is described in the Supplementary Information.

Raw data processing

After deep sequencing, the raw sequence reads of both RPFs and unprotected mRNA fragments were uploaded into the Galaxy platform (https://usegalaxy.org/). Using the Cutadapt tool (version 4.0+galaxy0), the 3' adapter sequences were removed, and 18 to 34 nt-long reads were

selected, which on average removes approximately 15% of the total reads. After checking that the adapter has been removed from reads by using the FastQC tool (Galaxy Version 0.73 + galaxy0), the Bowtie2 tool (Galaxy Version 2.4.2 + galaxy0) was used to align reads to the *B. subtilis* genome sequence NC_000964.3 to generate SAM files, which link reads to their genomic position. The SAM files have been submitted to the Gene Expression Omnibus (GEO) with accession number GSE250314.

Calculation of the ribosomal A-site codon position with Alt_predict

A commonly used web based tool for ribosome profiling data processing is RiboGalaxy [21]. Unfortunately, during our analysis RiboGalaxy's "Create Ribosome Profiles" tool did not indicate whether the 3' end offset was used to calculate ribosomal A-sites from RPFs. The offset direction can be selected in the Trips-viz (https://www.trips.ucc.i e/) online ribosome profile visualization platform, however the platform requires a genome annotation file format that is not yet available for *B. subtilis* [22]. Therefore, we developed our own Python script as an alternative method, named Alt_predict_v2.py. An advantage of this tool is that it provides the nucleotide sequence around calculated ribosomal A-sites. The script uses SAM files and determines the ribosomal A-site by subtracting 12 nt from the 3' end of RPFs, as shown schematically in Fig. 2A. Then it creates a comma-separated values (.csv) file containing 10 columns with: (i) "genome", indicating the genome code, (ii) "position", indicating the position of the first nucleotide of the calculated ribosomal A-site on the genome, (iii) "strand", indicating the reading direction of the gene on the genome, whereby "0" stands for 5' to 3' and "16" stand for a 3' to 5' end direction relative to the genome sequence (the use of the number 16 has to do with SAM file specific coding), (iv) "gene", standing for gene name, (v) "gene_length", indicating the gene length in nucleotides, (vi) "offset", indicating the distance of the first nucleotide in the calculated ribosomal A-site relative to the first nucleotide of the start codon, (vii) "in_orf_90", indicating whether the calculated ribosomal A-site is NOT located within either the first 5% or the last 5% of the gene length (the reason for this is to dismiss, when required, ribosomal pausing sites close to translational start or stop sites), (viii) "count", indicating the number of RPF reads that have this calculated ribosomal A-site, (ix) "sequence", providing 101 nucleotide sequence centered around the first nucleotide of the calculated ribosomal A-site, and (x) "Pause_codon", presenting the actual calculated ribosomal A-site codon. For convenience the output is also stored in a bed format so it can easily be imported in the integrated genome viewer (IGV). Alt_predict_v2.py can be downloaded from the GitHub repository https://github.com/BiosystemsDataAnalysis/P ausePredictionTools.git.

Codon phasing quality check with A-site distance distribution analysis

To test the quality of the ribosomal A side calculations, we developed a small Python script named "run_a_ site_analysis.py" that displays the number of RPF reads related to the distance of the calculated first nucleotide of the A-site relative to the position of nucleotides of the different amino acids codons (Fig. 2B and S4). The highest peak should be at position 0, which represents the first nucleotide of calculated A-site, and the peaks should display a codon (3 nucleotide) periodicity, or in case of stalling at specific amino acids codons, as is the case with mupirocin, a strong peak around position 0 (Fig. 2B). The generated distribution plot shows both A-site calculations using the 3' 12 nt offset and 5' 14 nt offset. The advantage of this A-site distance distribution analysis is that it will also indicate whether the offset of either 12 or 14 nt gives the best results. The program can be downloaded from GitHub (https://github.com/BiosystemsData Analysis/PausePredictionTools.git).

Visualization of ribosome pause sites around start codons using the A-site analysis script

Figure 4B shows the distribution of RPFs, based on their calculated ribosomal A-sites, within 20 nt upstream and downstream of all start codons in the genome. This figure has been made using the same "run_a_site_analysis.py" script with the same input files but with the additional "-orf" switch. The program can display both the A-site calculation using the 3' 12 nt offset or using the 5' 14 nt offset rule. In addition, it shows whether the gene is on the left or right side of the circular genome, and whether it reads in the clockwise or counterclockwise direction. We found that these characteristics did not affect the distribution shown in Fig. 4B.

Ribosome profile generation

The ribosome profiles shown in the figures were made with GraphPad and the Alt_predict .csv output files, using gene name information ("gene"), gene length information ("gene_length"), distance of the first nucleotide in the calculated ribosomal A-site relative to the first nucleotide of the start codon ("offset") and the number of RPF reads that have this calculated ribosomal A-site ("count"). The number of RPF reads were normalized using the total number of unique reads (mapping stats in Bowtie2 tool). The highest number of 16 h mRNA unique reads was used as reference number.

It can be useful to examine ribosome pause sites on a genomic scale using the integrated genome browser IGB [23], which enables a quick visual inspection of genes and

genome regions. To enable this using the Alt_predict data files, we have to convert them first into bedGraph files. This file type can be read by IGB but can also be edited using Excel. First a SAM file is converted to a bedGraph file using bamCoverage tool (Galaxy Version 3.5.2+galaxy0), and the bedGraph file is then opened in Excel, depicting 4 data columns. The first column contains the genome code, NC 000964.3 in our case. The data in the second column is replaced by the genome position of the calculated ribosomal A-site, which can be found under the "position" data column in the Alt_predict csv file. Importantly, these values have to be subtracted by the value 1 (thus, "position"-1) because IGB starts counting from 0 and not 1, consequently, the first nucleotide of the genome receives the value 0 in IGB. Since we only want to display the first nucleotide of the A-site, the third data column should contain the value of the second column + 1. The data in the fourth column indicates the read numbers (peak height) and should therefore be replaced by "count" data from the Alt_predict csv file. Subsequently, saving (do not use "Save as") maintains the bed-Graph configuration of the file. These new bedGraph files were then imported into IGB, including the reference genome NC_000964.3 FASTA file and its general feature GFF3 file, to visualize ribosome pausing positions on the genome.

Calculating codon frequency and codon pause score

To calculate codon frequency in the *B. subtilis* genome (NC_000964.3 from NCBI GenBank), we used the Sequence Manipulation Suite (https://www.bioinform atics.org/sms2/codon usage.html). To calculate codon pause scores, we used the output file of Alt_predict (see Ribosome profile generation). The pause score is than



Fig. 1 Growth and amylase production during repeated batch fermentation. *B. subtilis* BY212 containing the *amyM* overexpression plasmid pBW212 was grown in rich medium for 88 h at 30 °C under continuous shaking. Glucose was added at 16, 40 and 64 h (1% final concentration)

calculated for every peak by dividing the "count" by the ribosomal coverage of the gene that contains that peak. Ribosomal coverage of the gene is calculated by summing the "count" of the gene divided by the gene length (nucleotides). Then, the codon pause score is calculated by summing the pause scores of the same codons of the genome. The codon pause score is divided by the total number of RPF reads in ORFs to obtain the normalized codon pause score.

Protein secondary structure and signal peptide predictions Protein secondary structures were predicted using PredictProtein (https://predictprotein.org/) [24], and putativ e signal peptides using SignalP 6.0 [25].

Ribosome occupancy and highest ribosome pause peaks

To calculate the ribosome occupancy of genes, the sum of the read counts for a gene ("count" data column) was divided by the nucleotide length of the gene ("gene_ length" data column). The highest ribosome pause peak of each gene was determined on the basis of the highest "count" number for a gene. The ribosome occupancy and highest ribosome pause peak for all genes is listed in Table S5. The information on gene function and length were obtained from the SubtiWiki database [26].

Results and discussion

Assessment of ribosome stalling using mupirocin

To analyze ribosome pausing in *B. subtilis* under repeated batch conditions, a simple fermentation setup was established by using rich medium to which 1% glucose was added after 16, 40 and 64 h. The culture was grown at 30 °C under continuous shaking, resulting in around 28 OD₆₀₀ units after 64 h (Fig. 1). B. subtilis strain BY212 was used for amyM expression. This strain contains a deletion in the sigma factor gene sigF to prevent sporulation, and lacks the main secreted proteases, including NprE, AprE, Epr, Mpr, NprB, Vpr and Bpr, to stabilize the secreted enzyme [27]. BY212 was transformed with a high-copy expression plasmid harboring *amyM* under control of the *aprE* promoter, which becomes active during stationary growth [28, 29]. As shown in Fig. 1, the amylase activity in the medium starts to increase steadily after 16 h.

To assess the quality of our ribosome profiling procedure, we examined the effect of mupirocin, an antibiotic that inhibits isoleucyl-tRNA synthetases, resulting in ribosomes stalling at isoleucine codons [20, 30, 31]. After 16 h growth, cells were incubated for 10 min with 10 μ M mupirocin, and the reaction was stopped by adding crushed frozen buffer. Cells were harvested by a short centrifugation and the pellet was stored by flash freezing in liquid nitrogen. We chose this procedure instead of collecting cells by filtering, since the latter procedure has been shown to create bias in ribosome pause sites in *Escherichia coli* [20]. Subsequently, sucrose density gradient centrifugation was used to isolate ribosomes, and micrococcal nuclease digestion was employed to obtain ribosome protected fragments (RPF), which were isolated by polyacrylamide gel electrophoresis (Fig. S1). All ribosome profiling experiments were repeated to provide biological replicates. The isolated RPFs showed a size distribution around 25 nt (Fig. S2A), and subsequent next generation sequencing revealed that 80% of the reads were ribosomal RNA fragments, and around 15% (~ 1 M reads) uniquely mapped to the genome (Fig. S2B, Table S2).

Curiously, more than 90% of the ribosomal RNA fragments consisted of a single 23 S fragment with the sequence TGCCTCTTGGGGTTGTAGGACACT. Interestingly, the eukaryotic 28 S homologue of this sequence was also heavily enriched in a ribosome profiling study with yeast [32], suggesting that this motif is somehow highly stable.

To determine the ribosomal A-site codon position, we tried different offsets and found that either an offset of 14 nt from the 5' end, and especially an offset of 12 nt from the 3' end, gave the highest read scores for isoleucine codons in the ribosomal A-site (Fig. 2). This is in line with a previous studies, describing that a nucleotide offset from the 3' end of RPFs results in a more accurate prediction of the ribosomal A-site codon when using micrococcal nuclease [20, 30]. When the antibiotic chloramphenicol was used to fixate ribosomes, no exclusive peak at isoleucine codons was observed (Fig. S3A), confirming the activity of mupirocin.

Different protection at translation initiation sites

Having validated the experimental pipeline, we performed a standard ribosome profiling with a sample taken after 16 h growth, and used chloramphenicol together with direct cooling with crushed PBS buffer ice to instantly block translation [20, 33]. Analogously to the previous benchmark, the application of an offset of 12 nucleotides from the 3' end of RPF reads correlates best with ribosomal A-site positions (Fig. S3B). As shown in Fig. 3A, *amyM* contains 12 clear ribosome pause peaks with more than 150 RPF reads. The nucleotide sequence and codon context of these peaks are depicted in Fig. 3B. The biological replicate showed comparable peaks except for peak 11 (Fig. S7). To be sure that we were looking at ribosome pause sites, we also sequenced ribosomeunprotected mRNA fragments isolated from polyacrylamide gels (Fig. 3A, lower panel). A genome wide assessment of read peaks indicated that these mRNA control profiles showed fewer peaks compared to the ribosome profiles in open reading frames (Fig. S4). However, the mRNA control also showed high read numbers at the location of peak 6 (Fig. 3B), suggesting that this peak might be caused by a PCR amplification bias and not necessarily by ribosome pausing.

The strongest ribosome pause sites, P1 and P2, are located upstream of the ATG start codon (Fig. 3B), indicating strong binding of ribosomes to the Shine Dalgarno sequence. Such a peak between the Shine Dalgarno sequence and the ATG start codon was observed for many strongly expressed housekeeping genes. However, the exact nucleotide position of these ribosome pause peaks were neither determined by the Shine Dalgarno motif nor by the start codon, as is illustrated in Fig. 4A. Interestingly, when we performed a genome-wide analysis of these translation start peaks counting from the 5'



Fig. 2 Comparison of mupirocin ribosome profiling data by counting from either 3' or 5' end of RPF reads. (**A**) Schematic diagram for calculating the position of the first nucleotide in the ribosomal A-site codon from an RPF fragment from mupirocin treated cells when counting from either the 3' or 5' end of RPFs. In this example the ribosomal A-site contains the isoleucine codon ATC. (**B**) A-site distance distribution analysis. Average RPF reads of calculated A-sites for every nucleotide upstream and downstream of isoleucine codons. 0 indicates the first nucleotide of the calculated A-site of isoleucine codons. The A-site was determined by counting from either the 3' end (offset 12 nt) or 5' end (offset 14 nt). The advantage of this method is that it also indicates whether the offset used (12 and 14 nt) gives the best score



Fig. 3 Ribosome pause sites in amyM. (A) Upper profile: ribosome profile of amyM mRNA after 16 h growth. Peaks of more than 150 RPF reads are numbered. Lower profile: unprotected mRNA control sample. The signal peptide of AmyM spans amino acids 1–33. The first peaks, P1 and P2, are located before the start codon (B) Sequence details of the 12 ribosome pause sites. Ribosome density, based on RPF read numbers representing A-site codons (first nucleotide position), are shown in the first two lines, representing the 2 biological replicates (Ribo 1 and Ribo 2). The 3^{de} line shows the read numbers from the unprotected mRNA control sample (mRNA). Red boxes indicate the Shine Dalgarno sequence and start codon. A-site codons of the peaks are bold and underlined. Peak 6 is marked by an asterisk since the high mRNA read numbers suggests this peak does not reflect a ribosome pause site

А	Shine-Dalgarno sequence (SD)									Start Codon										
amyM	-15	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5
	G	G	A	G	G	A	G	A	G	T	G	A	G	A	T	A	T	G	A	A
	0	7	35	138	77	4131	75	61	<mark>198</mark>	187	30	1301	<mark>371</mark>	40	41	2	94	23	22	0
cspD	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6	7
	G	G	A	A	T	T	A	A	G	C	A	A	T	A	T	G	C	A	A	A
	9	9	1	53	37	3	1	0	0	7	58	346	5	1	0	0	0	0	0	0
tufA	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6
	A	G	G	A	G	G	A	T	T	T	T	A	G	A	A	T	G	G	C	T
	0	0	0	3	14	45	173	58	72	118	45	220	38	43	39	5	6	0	0	0
rpsB	-15	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5
	T	T	A	G	G	A	G	G	A	A	A	T	A	A	C	A	T	G	T	C
	0	0	0	0	0	0	0	0	1	5	34	158	2	0	8	3	3	0	0	0
gapA	-16	-15	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4
	A	A	A	G	G	A	G	G	A	A	A	C	A	A	T	C	A	T	G	G
	0	0	0	0	0	1	0	15	45	7	10	127	11	0	0	1	3	0	2	5
recA	-17	-16	-15	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3
	A	A	G	G	A	G	G	A	A	A	A	A	A	T	A	G	A	A	T	G
	0	0	0	0	1	0	1	65	0	5	7	77	3	1	9	24	0	24	0	0
glnA	-19	-18	-17	-16	-15	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1
	A	G	A	G	G	A	G	G	A	A	T	T	T	T	A	C	C	A	A	A
	0	0	0	0	7	1	0	2	2	2	3	294	55	15	7	98	9	6	19	3
rpIJ	-14	-13	-12	-11	-10	-9	-8	-7	-6	-5	-4	-3	-2	-1	1	2	3	4	5	6
	C	A	G	G	A	G	G	T	G	T	A	A	C	C	A	T	G	A	G	C
	0	0	0	0	3	0	0	34	11	3	17	61	40	9	0	0	0	10	2	0
B	4(- 0C				\bigwedge	(3' 5'	-er -er	าd าd			
ď	1	00 00 0	20	_] -1	0			ba	10	√	2	✓	5	∧ − 30		<u>∽</u> 4	1 10		
					[Dist	an	ce	to	sta	rt	cod	on	(n	t)					

Fig. 4 Different protection at translation initiation sites. (A) Ribosome density based on RPF read numbers at the translation initiation site of *amyM* and several genes coding for abundant proteins. Red boxes indicate Shine Dalgarno sequences and start codons. (B) Average ribosome density around start codons by assigning the A-site from either the 3' or 5' end, using an offset of 12 nt and 14 nt, respectively

end, a narrower peak distribution was obtained (Fig. 4B). This difference is likely caused by the conformational change of the ribosome from an initiation complex to a translating ribosome [34], resulting in a different protection of the mRNA from micrococcal nuclease digestion.

It is also clear from Fig. 4B that the codon periodicity of calculated A-sites after the start codon is best observed using the offset from the 3' end.



Fig. 5 Correlation between codon frequency and codon pause score. Scatter plots showing the correlation between A site codon pause score and genome wide codon frequency at 16 h. Genome wide ribosome frequencies of all the 61 codons are shown in Table S3. Glutamate codons (gaa, gag) and aspartate codons (gac, gat) are indicated. Regression line and R² value are indicated. Red dots are codons at the main ribosome pause sites of *amyM* (Fig. 3). Biological replicates are shown in Fig. S5A

Codon motifs at ribosome pause sites

The presence of clear ribosome pause sites in amyM begged the question what the cause is of translation pausing at these locations. A well-known trigger for ribosome pausing is the presence of consecutive proline codons [5– 8], but none of the ribosome peaks in *amyM* are found at or close to proline codons (Fig. 3B). Another mechanism could be the presence of rare codons, although for some organisms, including Saccharomyces cerevisiae and E. coli, it has been shown that there is no increase in pausing at rare codons [20, 35, 36]. Indeed, when we plotted the codon pause score, i.e. the fraction of RPFbased A-site reads for each codon in the genome, against the codon frequency in the genome, rare codons did not show higher pause scores (Fig. 5). A similar result was observed for the biological replicate (Fig. S5). In fact, none of the codons at the center of the ribosome pause peaks in *amyM* are rare codons (Fig. 3 and Table S3). The most visible deviation from the genomic codon frequency is the slight overrepresentation of aspartate and glutamate codons at ribosome pause sites (Fig. 5 and Table S4), which might suggest relative low levels of these amino acids in the medium at the time of sampling.

Correlation between pause sites and secondary protein structure

It is assumed that ribosome pausing may facilitate protein domain folding [37], thus ribosome pause sites might be found close to the beginning or end of secondary



Fig. 6 Ribosome pause sites in *amyM* mRNA after 64 h fermentation. Ribosome profiles of *amyM* after 16 h and 64 h growth under repeated batch fermentation conditions. The signal peptide of AmyM spans amino acids 1–33. Main ribosome pause peaks are numbered, and new peaks are indicated in red. Profiles of the biological replicate are shown in Fig. S7. RFP read numbers and sequence details at the peak positions, including those of the biological replicates, are shown in Fig. S8

protein structures. Several of the ribosome pause peaks seem to be located close to a putative alpha helix or beta sheet in AmyM (Fig. S6A, B). However, these secondary structures are abundant and correlation with ribosome pause peaks appeared to be statistically insignificant (Fig. S6C). Of course, this does not exclude a relation between ribosome pausing and protein folding, but it shows that such correlation cannot be deduced from knowledge of basic secondary protein structure.

Codon occupancy stability during fermentation

After 64 h of growth the optical density of the cell culture does not increase any further and the amylase production has peaked (Fig. 1). Nevertheless, ribosome profiling at this time point showed that the *amyM* mRNA still contained clear ribosome pause sites with many of the same peaks that were found for the 16 h sample, and a few additional peaks (Fig. 6 and Fig. S7). The genome wide codon occupancy at 64 h showed only mild differences compared to the 16 h sample, with the strongest increase (1.6x) for cystine codons (Table S4).

Correlation between ribosome occupancy and mRNA levels

The clear ribosome pause sites in *amyM* at 64 h suggest that the gene is still actively translated, yet the amount of amylase in the medium does not further increase (Fig. 1). We were therefore curious how the ribosome occupancy scales to the mRNA control read numbers, and analyzed this for all genes. As shown in Fig. 7, both the 16 h and 64 h samples displayed a strong correlation between ribosome occupancy and mRNA levels.

Although the ribosome occupancy did not change much over time, the mRNA levels are clearly lower at 64 h compared to the 16 h sample. This reduction is measurable since mRNA reads comprise only a small fraction compared to ribosomal RNA reads (Fig. S2C and D), and



Fig. 7 Correlation between ribosome occupancy and mRNA abundance. Correlation of ribosome occupancy and mRNA abundance of genes at 16 h and 64 h growth. Genes coding for the 100 most abundant proteins in *B. subtilis* are indicated in blue. *amyM* shown by the encircled white dot. Notice the shift towards lower relative mRNA reads in the 64 h sample

Table 1 mRNAs with highest ribosome occupancy. 20 genes with the highest ribosome occupancy are listed. Ribosome occupancy is based on the number of RPF reads divided by gene length, and the average value from the two biological replicates is listed including the standard deviation (SD). Of note, the ribosome unprotected mRNA control reads, also divided by gene length, are based on a single experiment. mRNAs that also contain the highest ribosome pause peaks, listed in table 2, are indicated in bold

Gene	annotated function	gene length	16 h occupancy	16 h mRNA reads	64 h occupancy	64 h mRNA reads
ереХ	antimicrobial peptide	150	127.6 +/- 11.9	73.5	83.5 +/- 26.8	23.8
skfA	killing of non-sporulating sister cells	168	47.6 +/- 5.0	20.9	42.9 +/- 9.9	19.5
yuzK	unknown	138	37.7 +/- 8.7	10.6	29.1 +/- 18.5	11.7
rpmA	translation	285	18.9 +/- 10.1	5.9	16.8 +/- 0.0	3.1
ykoJ	unknown	513	11.0 +/- 3.5	2.8	55.6 +/- 19.2	7.4
cspC	RNA chaperone	201	10.7 +/- 1.4	3.9	13.0 +/- 9.0	1.5
cspB	RNA chaperone	204	10.3 +/- 1.2	6.4	5.8 +/- 0.9	1.4
cspD	RNA chaperone	201	9.7 +/- 1.8	4.0	8.5 +/- 2.6	1.5
veg	control of biofilm formation	261	8.1 +/- 0.3	24.7	5.7 +/- 0.0	4.2
sdpC	killing of non-sporulating sister cells	612	8.5 +/- 1.1	4.5	2.1 +/- 1.2	0.4
gInA	regulation glutamine biosynthesis	1335	10.7 +/- 4.3	5.8	6.2 +/- 0.8	1.5
abh	regulation during transition from growth to stationary phase	279	4.4 +/- 1.8	0.3	6.1 +/- 1.1	0.1
phrC	control of ComA activity	123	8.1 +/- 4.2	1.8	3.0 +/- 0.9	0.6
rpmB	translation	189	3.2 +/- 2.5	1.1	1.8 +/- 1.9	0.3
htrA	protein quality control	1350	4.7 +/- 0.1	1.7	1.5 +/- 1.1	0.4
htrB	protein quality control	1377	2.0 +/- 0.2	1.4	1.4 +/- 0.5	0.4
spx	regulator of many genes	396	5.8 +/- 1.2	1.3	8.2 +/- 3.0	0.8
spoVG	control of sporulation initiation	294	4.6 +/- 1.6	7.5	7.1 +/- 4.0	8.3
yhfH	unknown	141	1.4 +/- 0.4	0.4	54.7 +/- 68.4	6.2
rpsU	translation	174	4.6 +/- 0.7	2.0	6.6 +/- 1.4	1.0

decreased from 8.2 to 4.8% for the 16 h and 64 h samples, respectively (Table S2). This result suggests that the lack of amylase production after 64 h is due to low transcription levels and not due to a reduction in translation capacity of cells, which is in line with previous observations showing that even after long periods of starvation *B. subtilis* cells can still produce high amounts of a protein when transcription is induced [38].

Strong translation pausing in other genes

Since the translation capacity of the cell will affect AmyM production, it might be relevant to identify the most highly translated genes in the genome. In Table 1 we have listed 20 genes with the highest ribosome occupancy for the 16 h sample. In Table 2, we have listed 20 genes with the highest ribosome pause peaks. Their ribosome profiles are shown in Fig. S9. 14 of these genes are also

Table 2 mRNAs with highest ribosome pause peaks. 20 genes with the highest ribosome pause peak are listed. The nucleotide position (peak position), and average RPF read numbers (peak height) of the peaks are listed. Standard deviations (SD) are included. Genes with the highest ribosome occupancy listed in table 1 are indicated in bold

Gene	annotated function	gene length	16 h peak position	16 h peak height	64 h peak position	64 h peak height
ереХ	antimicrobial peptide	150	145	5281 +/- 2476	145	2912 +/- 3019
rpmA	translation	285	64	1905 +/- 1713	65	1171 +/- 299
skfA	killing of non-sporulating sister cells	168	34	2322 +/- 1006	34	1273 +/- 463
mdh	TCA cycle	939	835	2501 +/- 668	835	671 +/- 805
htrA	protein quality control	1350	990	1886 +/- 1464	990	433 +/- 504
htrB	protein quality control	1377	4	405 +/- 18	4	775 +/- 16
yuzK	unknown	138	82	2328 +/- 351	82	1183 +/- 829
sdpC	killing of non-sporulating sister cells	612	69	1365 +/- 482	69	165 +/- 112
abh	regulation during transition from growth to stationary phase	279	22	1119 +/- 554	22	1613 +/- 303
gInA	Regulation glutamine biosynthesis	1335	469	1303 +/- 138	469	952 +/- 646
sdpB	maturation of the SdpC toxin	972	930	1222 +/- 19	930	314 +/- 412
aroA	biosynthesis of aromatic amino acids	1077	535	1017 +/- 252	535	755 +/- 617
recA	DNA repair/ recombination	1047	9	685 +/- 551	9	210 +/- 158
ykoJ	unknown	513	232	674 +/- 443	232	2932 +/- 856
gapA	catabolic enzyme in glycolysis	1008	157	925 +/- 8	157	81 +/- 64
spx	regulator of many genes	396	313	766 +/- 43	313	813 +/- 100
yerl	unknown	1011	7	688 +/- 103	7	543 +/- 588
cspC	RNA chaperone	201	145	696 +/- 56	148	789 +/- 800
cspB	RNA chaperone	204	61	347 +/- 127	61	226 +/- 66
yhfH	unknown	141	134	131 +/- 48	134	6702 +/- 8846



Fig. 8 Ribosome profiles of toxin encoding genes displaying strong ribosome pause sites. Ribosome profiles of the toxin encoding *epeX*, *sdpC*, *skfA* genes, and the unknown genes *yuzK* and *yhfH* after 16 h growth. The amino acid size of the related proteins are indicated between brackets. Biological replicates are shown in Fig. S10

among the top 20 genes with the highest ribosome occupancy (Table 1). For a complete list of all *B. subtilis* genes, see Table S5. The *htrA* and *htrB* genes in Table 1 encode the membrane-anchored protein quality control proteases, whose transcription is strongly induced by the overexpression of amylase [39, 40]. *ykoJ* is another gene which has been shown to be induced upon overexpression of an amylase [41]. HtrA/B are required for high level amylase production [42], but nothing is known about YkoJ. According to SignalP 6.0 [25], this 513 amino acid long protein contains a putative signal peptide, thus could be secreted. To examine whether YkoJ is important for the overproduction of AmyM, we deleted its gene, but this mutation did not change amylase levels in the medium (data not shown).

Interestingly, among the strongest translated genes are three genes coding for small secreted antimicrobial peptides: *epeX*, *sdpC* and *skfA*. Figure 8 shows the position of the strong ribosome pause sites in these genes. It is tempting to speculate that these are ribosome stall sites, preventing full synthesis of these toxins when they have not yet docked with their dedicated secretion systems. Tables 1 and 2 contain two other small genes with a single strong ribosome pausing site, *yuzK* and *yhfH*, coding for 45 and 46 amino acid long peptides, respectively (Fig. 8). It will be interesting to see whether these genes show some resemblance to toxins. Finally, we wondered whether removing the most strongly translated mRNAs would increase overall translation capacity and increase amylase production. To test this, we constructed a strain that lacked the three toxin genes with the highest ribosome occupancy, *epeX*, *sdpC* and *skfA*, however, the production of AmyM did not increase in this strain (data not shown), indicating that the high ribosome occupancy of these genes does not pose a burden on the translation machinery.

Conclusions

Our ribosome profiling analysis provided useful insights into the translation activity during repeated batch fermentation of B. subtilis, and suggests that even after 64 h, many genes are still being translated, including amyM mRNA. Interestingly, these ribosome pausing sites cannot be predicted based on nucleotide or amino acid sequences, suggesting that the mRNA structure might play a pivotal role in ribosome pausing. Of note, we could not find a clear correlation between strong ribosome pausing sites in *amyM* and its predicted secondary mRNA structure (Fig. S11). It will be interesting to determine whether synonymous codon changes can remove these pause sites, and whether this improves overall expression. However, it can also be that these ribosomal pausing sites facilitate the binding of SecA and/or other chaperones necessary for efficient secretion of the amylase. In any case, the presence of ribosome pausing sites in *amyM* provide novel information that should be considered when performing codon optimization studies to improve the production of this enzyme by *B. subtilis*, as these pausing sites might be important for efficient folding of the protein.

Supplementary Information

The online version contains supplementary material available at https://doi.or g/10.1186/s12934-025-02659-3.

Supplementary Material 1

Supplementary Material 2

Acknowledgements

We thank the members of the Bacterial Cell Biology group for useful and constructive discussions, and Selina van Leeuwen from the Dutch Genomics Service & Support Provider at UvA for providing excellent sequencing services.

Author contributions

YZ performed and YZ, BW, AA, GD and LWH designed the experiments. YZ, AA, FK, CS, PIC, MFF, MA and LWH analysed data. YZ and LWH, with the help of the other co-authors, wrote the paper.

Page 12 of 13

Funding

This research was partially funded by a CSC China Scholarship Council (CSC) fellowship (BW and YH), and a BASF SE grant (LWH).

Data availability

All data analysed in this study are included in the manuscript. The raw transcriptome datasets is available in the Gene Expression Omnibus (GEO) with accession number GSE250314. Strains can be obtained upon request. Original software is available from the GitHub repository.

Declarations

Ethics approval and consent to participate Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Received: 11 October 2024 / Accepted: 20 January 2025 Published online: 27 January 2025

References

- Kim SJ, Yoon JS, Shishido H, Yang Z, Rooney LA, Barral JM, Skach WR. Translational tuning optimizes nascent protein folding in cells. Science. 2015;348:444–8.
- Kimchi-Sarfaty C, Oh JM, Kim I-W, Sauna ZE, Calcagno AM, Ambudkar SV, Gottesman MM. A silent polymorphism in the MDR 1 gene changes substrate specificity. Science. 2007;315:525–8.
- Komar AA. A pause for thought along the co-translational folding pathway. Trends Biochem Sci. 2009;34:16–24.
- Li S-c, Zhou Z-q, Li L-p, Xu Z-h, Zhang Q-q. Shi S-s: risk assessment of water inrush in karst tunnels based on attribute synthetic evaluation system. Tunn Undergr Space Technol. 2013;38:50–8.
- Ito K, Chiba S. Arrest peptides: cis-acting modulators of translation. Annu Rev Biochem. 2013;82:171–202.
- Leininger SE, Rodriguez J, Vu QV, Jiang Y, Li MS, Deutsch C, O'Brien EP. Ribosome elongation kinetics of consecutively charged residues are coupled to electrostatic force. Biochemistry. 2021;60:3223–35.
- Krafczyk R, Qi F, Sieber A, Mehler J, Jung K, Frishman D, Lassak J. Proline codon pair selection determines ribosome pausing strength and translation efficiency in bacteria. Commun Biology. 2021;4:589.
- Buskirk AR, Green R. Ribosome pausing, arrest and rescue in bacteria and eukaryotes. Philosophical Trans Royal Soc B: Biol Sci. 2017;372:20160183.
- 9. Doma MK, Parker R. Endonucleolytic cleavage of eukaryotic mRNAs with stalls in translation elongation. Nature. 2006;440:561–4.
- Al-Hawash AB, Zhang X, Ma F. Strategies of codon optimization for high-level heterologous protein expression in microbial expression systems. Gene Rep. 2017;9:46–53.
- Overkamp W, Beilharz K, Detert Oude Weme R, Solopova A, Karsens H, Kovács ÁT, Kok J, Kuipers OP, Veening J-W. Benchmarking various green fluorescent protein variants in Bacillus subtilis, Streptococcus pneumoniae, and Lactococcus lactis for live cell imaging. Appl Environ Microbiol. 2013;79:6481–90.
- 12. Ingolia NT, Ghaemmaghami S, Newman JR, Weissman JS. Genome-wide analysis in vivo of translation with nucleotide resolution using ribosome profiling. *science* 2009, 324:218–223.
- Sundarram A, Murthy TPK. α-amylase production and applications: a review. J Appl Environ Microbiol. 2014;2:166–75.
- Gopinath SC, Anbu P, Arshad MM, Lakshmipriya T, Voon CH, Hashim U, Chinni SV. Biotechnological processes in microbial amylase production. Biomed Res Int, 2017.
- Farooq MA, Ali S, Hassan A, Tahir HM, Mumtaz S, Mumtaz S. Biosynthesis and industrial applications of α-amylase: a review. Arch Microbiol. 2021;203:1281–92.
- Phan TTP, Nguyen HD, Schumann W. Development of a strong intracellular expression system for Bacillus subtilis by optimizing promoter elements. J Biotechnol. 2012;157:167–72.

- 17. Su Y, Liu C, Fang H, Zhang D. Bacillus subtilis: a universal cell factory for industry, agriculture, biomaterials and medicine. Microb Cell Fact. 2020;19:1–12.
- Abuhena M, Al-Rashid J, Azim MF, Khan MNM, Kabir MG, Barman NC, Rasul NM, Akter S, Huq MA. Optimization of industrial (3000 L) production of Bacillus subtilis CW-S and its novel application for minituber and industrial-grade potato cultivation. Sci Rep. 2022;12:11153.
- Koo B-M, Kritikos G, Farelli JD, Todor H, Tong K, Kimsey H, Wapinski I, Galardini M, Cabal A, Peters JM. Construction and analysis of two genome-scale deletion libraries for Bacillus subtilis. Cell Syst. 2017;4:291–305. e297.
- Mohammad F, Green R, Buskirk AR. A systematically-revised ribosome profiling method for bacteria reveals pauses at single-codon resolution. Elife. 2019;8:e42591.
- Michel AM, Mullan JP, Velayudhan V, O'Connor PB, Donohue CA, Baranov PV. RiboGalaxy: a browser based platform for the alignment, analysis and visualization of ribosome profiling data. RNA Biol. 2016;13:316–9.
- Kiniry SJ, Judge CE, Michel AM, Baranov PV. Trips-Viz: an environment for the analysis of public and user-generated ribosome profiling data. Nucleic Acids Res. 2021;49:W662–70.
- 23. Freese NH, Norris DC, Loraine AE. Integrated genome browser: visual analytics platform for genomics. Bioinformatics. 2016;32:2089–95.
- Bernhofer M, Dallago C, Karl T, Satagopam V, Heinzinger M, Littmann M, Olenyi T, Qiu J, Schütze K, Yachdav G. PredictProtein-predicting protein structure and function for 29 years. Nucleic Acids Res. 2021;49:W535–40.
- Teufel F, Almagro Armenteros JJ, Johansen AR, Gíslason MH, Pihl SI, Tsirigos KD, Winther O, Brunak S, von Heijne G, Nielsen H. SignalP 6.0 predicts all five types of signal peptides using protein language models. Nat Biotechnol. 2022;40:1023–5.
- Pedreira T, Elfmann C, Stülke J. The current state of Subti Wiki, the database for the model organism Bacillus subtilis. Nucleic Acids Res. 2022;50:D875–82.
- 27. Schroeder JW, Simmons LA. Complete genome sequence of Bacillus subtilis strain PY79. Genome Announcements. 2013;1:e01085–01013.
- 28. Diderichsen B, Christiansen L. Cloning of a maltogenic alpha-amylase from Bacillus stearothermophilus. FEMS Microbiol Lett. 1988;56:53–9.
- Ferrari E, Henner D, Perego M, Hoch J. Transcription of Bacillus subtilis subtilisin and expression of subtilisin in sporulation mutants. J Bacteriol. 1988;170:289–95.
- Mohammad F, Woolstenhulme CJ, Green R, Buskirk AR. Clarifying the translational pausing landscape in bacteria by ribosome profiling. Cell Rep. 2016;14:686–94.
- Woolstenhulme CJ, Guydosh NR, Green R, Buskirk AR. High-precision analysis of translational pausing by ribosome profiling in bacteria lacking EFP. Cell Rep. 2015;11:13–21.

- 32. Gerashchenko MV, Lobanov AV, Gladyshev VN. Genome-wide ribosome profiling reveals complex translational regulation in response to oxidative stress. Proc Natl Acad Sci. 2012;109:17394–9.
- Oh E, Becker AH, Sandikci A, Huber D, Chaba R, Gloge F, Nichols RJ, Typas A, Gross CA, Kramer G. Selective ribosome profiling reveals the cotranslational chaperone action of trigger factor in vivo. Cell. 2011;147:1295–308.
- Basu RS, Sherman MB, Gagnon MG. Compact IF2 allows initiator tRNA accommodation into the P site and gates the ribosome to elongation. Nat Commun. 2022;13:3388.
- Galmozzi CV, Merker D, Friedrich UA, Döring K, Kramer G. Selective ribosome profiling to study interactions of translating ribosomes in yeast. Nat Protoc. 2019;14:2279–317.
- Wu CC-C, Zinshteyn B, Wehner KA, Green R. High-resolution ribosome profiling defines discrete ribosome elongation states and translational regulation during cellular stress. Mol Cell. 2019;73:959–70. e955.
- Samatova E, Daberger J, Liutkute M, Rodnina MV. Translational control by ribosome pausing in bacteria: how a non-uniform pace of translation affects protein production and folding. Front Microbiol. 2021;11:619430.
- Gray DA, Dugar G, Gamba P, Strahl H, Jonker MJ, Hamoen LW. Extreme slow growth as alternative strategy to survive deep starvation in bacteria. Nat Commun. 2019;10:890.
- Stephenson K, Harwood CR. Influence of a cell-wall-associated protease on production of α-amylase by Bacillus subtilis. Appl Environ Microbiol. 1998;64:2875–81.
- Hyyryläinen HL, Bolhuis A, Darmon E, Muukkonen L, Koski P, Vitikainen M, Sarvas M, Prágai Z, Bron S, Van Dijl JM. A novel two-component regulatory system in Bacillus subtilis for the survival of severe secretion stress. Mol Microbiol. 2001;41:1159–72.
- Lulko AT, Veening J-W, Buist G, Smits WK, Blom EJ, Beekman AC, Bron S, Kuipers OP. Production and secretion stress caused by overexpression of heterologous α-amylase leads to inhibition of sporulation and a prolonged motile phase in Bacillus subtilis. Appl Environ Microbiol. 2007;73:5354–62.
- 42. Öktem A, Núñez-Nepomuceno D, Ferrero-Bordera B, Walgraeve J, Seefried M, Gesell Salazar M, Steil L, Michalik S, Maaß S, Becher D. Enhancing bacterial fitness and recombinant enzyme yield by engineering the quality control protease HtrA of Bacillus subtilis. Microbiol Spectr. 2023;11:e01778–01723.

Publisher's note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.