AEM Accepted Manuscript Posted Online 26 June 2015
Appl. Environ. Microbiol. doi:10.1128/AEM.00807-15
Copyright © 2015, American Society for Microbiology. All Rights Reserved.

Accepted Manuscript Posted Online
Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Polar fixation of plasmids during recombinant protein production in *Bacillus megaterium* results in population heterogeneity

Running Title: Unequal distribution of multi-copy plasmids

Karin M. Münch[1], Johannes Müller[2,3], Sarah Wienecke[1], Simone Bergmann[1], Steffi Heyber[1], Rebekka Biedendieck[1,4], Richard Münch[1,4,#], Dieter Jahn[1,4]

[1]Institute of Microbiology, Technische Universität Braunschweig, Spielmannstr. 7, D-38106 Braunschweig, Germany;

[2]Centre for Mathematical Sciences, Technische Universität München, Boltzmannstr. 3, D-85747 Garching/Munich, Germany;

[3]Institute of Computational Biology, Helmholtz Zentrum München - German Research Center for Environmental Health, Ingolstädter Landstr. 1, D-85764 Neuherberg, Germany;

[4]Braunschweig Integrated Centre of Systems Biology (BRICS), Technische Universität Braunschweig, D-38106 Braunschweig, Germany.

[#]Corresponding author:

Email: r.muench@tu-bs.de

Tel.: +49-531-391-5847

Fax: +49-531-391-5854

Accepted Manuscript Posted Online

ABSTRACT

During the last two decades, *Bacillus megaterium* has been systematically developed for the gram-per-liter scale production of recombinant proteins. The plasmid-based expression systems employed use a xylose-controlled promoter. Protein production analyses at the single cell level using green fluorescent protein as a model product revealed a cell culture heterogeneity characterized by a significant proportion of low-producing bacteria. Due to the enormous size of *B. megaterium,* such bistable behavior seen in subpopulations was readily analyzed by time-lapse microscopy and flow cytometry. Cell culture heterogeneity was not simply caused by plasmid loss: Instead, an asymmetric distribution of plasmids during cell division was detected during the exponential growth phase. Multi-copy plasmids are generally randomly distributed between daughter cells. However, *in vivo* and *in vitro* experiments demonstrated that under conditions of strong protein production, plasmids are retained at one of the cell poles. Furthermore, it was found that cells with accumulated plasmids and high protein production ceased cell division. As a consequence, the overall protein production of the culture was mainly achieved by the subpopulation with a sufficient plasmid copy number. Based on our experimental data, we propose a model whereby the distribution of multi-copy plasmids is controlled by polar fixation under protein production conditions. Thereby, cell lines with fluctuating plasmid abundance arise, which results in population heterogeneity. Our results provide initial insights into the mechanism of cellular heterogeneity during plasmid-based recombinant protein production in a *Bacillus* species.

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

INTRODUCTION

The Gram-positive *Bacillus* group is central to the industrial and biotechnological production of proteases, amylases, antibiotics and special chemicals on a ton scale. *Bacillus subtilis*, *Bacillus licheniformis*, *Bacillus amyloliquefaciens* and *Bacillus megaterium* are only some examples of this commercially important class of bacteria (1, 2). More specifically, *B. megaterium* is an attractive host for heterologous protein production. In particular, its lack of endogenous endotoxins and alkaline proteases, its stable maintenance and replication of plasmids, a strong protein secretion system and its capacity to grow on various cheap carbon sources are important criteria for its successful application in biotechnology industries (3, 4). At approximately 4 x 1.5 µm in size, it is one of the largest known bacteria, exhibiting a 100-times greater volume than *Escherichia coli* (5).

We constructed a series of plasmids with the aim of producing recombinant protein. These expression systems utilized a strong xylose-inducible promoter, P_{xylA}, part of the *B. megaterium* *xylABT* operon (6). The *xylA* and *xylB* genes encode enzymes involved in xylose degradation, while *xylT* encodes a xylose transporter. In the absence of xylose, the expression of the operon is repressed by the xylose repressor, XylR, while in the presence of xylose, the expression of the operon is derepressed. The *xylR* gene is located upstream to the *xylABT* operon, is transcribed in a divergent direction and is negatively autoregulated (7). For the construction of a xylose-inducible expression system, the *xylR* gene, with its corresponding promoters P_{xylR} and P_{xylA}, were cloned into a freely replicating, broad-host-range plasmid and further optimized by introducing strong promoter elements (8). The plasmidless *B. megaterium* strain, DSM319, is usually employed as production host (9). As a model protein product for the expression system, the gene for the readily detectable green fluorescent protein (GFP) was used. In a previous study, up to 1.25 g

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

GFP/L were produced using the developed vector system (10). However, the culture showed a significant amount of protein production heterogeneity at the single cell level. Flow cytometry analyses of GFP-producing *B. megaterium* grown in a bioreactor revealed a stable subpopulation of about 30% low producers, even under strong, selective conditions (11). Low-producing cells were found alive, indicating the stable formation of a subpopulation within a culture of clonal cells. Moreover, these cells were still proliferating, which excluded persistence as a reason for low production. These observations are of significant commercial interest, since the growth of low-producing subpopulations consumes valuable resources during protein production processes.

Here, we investigated this bimodal production behavior of individual cell lineages to elucidate its underlying mechanistic principles. Subpopulations were analyzed via single cell analyses using flow cytometry and time-lapse microscopy. In order to study, at the molecular level, the influence of plasmid copy number on bimodal production behaviour, we observed plasmid abundance and localization. In addition, fluorescence *in situ* hybridization (FISH) was employed for the direct cellular localization of plasmids. Our data suggest that the observed bimodality was not a product of differential gene regulatory circuits but, rather, a matter of unequal plasmid distribution between daughter cells, even under selective conditions. In particular, the common assumption concerning free plasmid diffusion is questioned and a mechanism leading to the unequal distribution of plasmids is proposed. Our results provide new insights into the distribution of heterologous multi-copy plasmids during the process of recombinant protein production.

Applied and Environmental Microbiology

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

MATERIALS AND METHODS

Strains, media, plasmids and primers

Bacillus megaterium DSM319 (9) was grown in A5 medium (8) containing 30 g/L fructose instead of glucose and supplemented with 10 µg/mL tetracycline. Cell cultures were performed either as a batch culture or in a BioLector microbioreactor system (m2p-labs, Baesweiler, Germany). The *E. coli* strain DH10B (Invitrogen Life Technologies, Carlsbad, CA, USA) was used for all cloning purposes. All plasmids and primers used in this study are described in Table 1. All shown studies were based on the pSSBm85 plasmid or derivatives of it (10). This plasmid carries the *gfp*+-gene (herein referred to as *gfp,* with GFP for the encoded protein, respectively), a variant of the wild-type *gfp* whose gene product exhibits increased fluorescence (12). This *gfp* variant was originally described for the pMUTIN-GFP+ plasmid (13). For the construction of a C-terminal XylR-mCherry fusion encoded in our expression system, the following steps were carried out: Starting with the pSSBm85 plasmid (10), the stop codon of *xylR* was replaced by a new KpnI site. This was accomplished by the introduction of a synthetic AflII/SpeI fragment comprising the whole *xylR* gene (GeneArt Life Technologies, Carlsbad, CA, USA) resulting in pKMMBm1. The oligonucleotides, xylR-mCherry for and xylR-mCherry rev, were used as primers to amplify the *mcherry* gene from the pJS72 plasmid (14) with the introduction of flanking AflII and KpnI restriction sites. The *xylR-mCherry* fusion (pKMMBm2) was realized by cloning the *mcherry* encoding AflII/KpnI fragment downstream of *xylR* into an AflII/KpnI cut pKMMBm1 (Fig. S1). The Δ*xylR* mutant plasmid, pKMMBm5, was created by cutting the XmaI/KpnI fragment from pKMMBm1, thus removing the *xylR* gene, generating blunt ends

Accepted Manuscript Posted Online

using Klenow polymerase and religating the plasmid. For pRBBm99, pSSBm85 was cut with AflII/SphI, blunt ends were generated and the plasmid was religated.

Flow cytometry and cell sorting

Flow cytometry and cell sorting were performed using a Cube8 (Sysmex Partec, Münster, Germany) and a FACSAria II (Becton-Dickinson, Mountain View, CA) flow cytometer, respectively. For this purpose, *B. megaterium* cells were grown for six h, diluted to an OD_{595} of 0.05 and further cultured in A5 media supplemented with 10 µg/mL tetracycline. Induction of gene expression with 0.5 % xylose was carried out at an OD_{595} of 0.2. The culture was harvested and washed twice with 1× PBS. For flow cytometry, cells were sonicated and diluted to a final concentration of 10^6 cells/sample. Finally, 10^5 events were counted. For cell sorting, approximately $10^6 – 10^7$ cells were recovered. Data analyses were performed in R Bioconductor using flowCore and flowViz software packages (15, 16).

Quantitative reverse-transcription polymerase chain reaction (qRT-PCR)

B. megaterium carrying the pSSBm85 plasmid was cultivated aerobically in A5 medium containing 10 µg/mL tetracycline. A five-hour preculture was diluted to an OD_{578} of 0.05 and gene expression was induced by the addition of 0.5% xylose after 1.5 h in the early exponential phase. Producing cells were harvested and sorted, after an additional hour, into fractions of GFP-

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

high and low producers. Total RNA was isolated from the sorted cells, purified using an innuPREP RNA Mini Kit and DNA removed with an innuPREP DNase I Digest Kit (Analytik Jena, Jena, Germany). RNA purity was checked using a 2100 Bioanalyzer (Agilent Technologies, Santa Clara, CA, USA); RNA integrity numbers (RINs) of between 8.2 – 10.0 were achieved. First-strand cDNA synthesis was performed with SuperScript II reverse transcriptase and hexameric random primer oligonucleotides (Invitrogen Life Technologies, Carlsbad, CA, USA). RNA was degraded using 200 mM NaOH and 100 mM EDTA, neutralized with 200 mM HEPES and 1.2 M sodium acetate, and was further purified using a PCR-Purification Kit (Qiagen, Hilden, Germany). Quantitative real-time PCR was performed on a CFX96 real-time PCR detection system (Bio-Rad, Hercules, CA, USA). The 20 µL final reaction mixture contained 5 ng of cDNA, 10 µL of SsoFast EvaGreen supermix (Bio-Rad) and 20 pmol of each appropriate primer (Table 1). The relative quantification of gene expression was determined using the efficiency-corrected relative quantification method (17), and a model-based estimation of real-time PCR amplification efficiency (18) improved with a log-logistic 5 parameter model (19). Reference genes were determined and validated using the geNorm algorithm in the R Bioconductor package SLqPCR (20). As reference genes, *rpoB* (RNA polymerase beta subunit), *gyrB* (DNA gyrase subunit B), *heli1* (SNF2 helicase-associated protein) and *heli2* (SNF2 family helicase) were finally chosen. Data analysis for the calculation of C_q values, relative expression ratios and statistics were carried out using the qpcR package (21).

Microscopy and image analysis

Applied and Environmental Microbiology

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

149 Fluorescence time-lapse microscopy was performed with an Axiovert 200M microscope (Zeiss,
150 Jena, Germany), an Axiocam CCD camera (Zeiss) and a Heating System 1 microscope incubator
151 (ibidi, Martinsried, Germany). An agar block method for the preparation of cells and live-cell
152 imaging was used (22). The agarose pad was prepared using A5 medium supplemented with
153 1.5% agarose, 0.5% xylose and 10 µg/mL tetracycline in low 35 mm µ-dishes coated with
154 ibiTreat (ibidi Labware, Martinsried, Germany). For DAPI staining, living cells taken from a
155 culture two h after gene induction were incubated for 5 min in 300 nM DAPI dihydrochloride
156 (Invitrogen Life Technologies, Carlsbad, CA, USA). Cells were washed three times with A5
157 media and pipetted onto an agarose pad as previously described (22). Automated image capture
158 was performed every 5 min at an incubation temperature of 37°C. Image analyses were done
159 using Axiovision 4.8 (Zeiss) and TLM-Tracker software (23). Bimodal cell lineage distributions
160 were analyzed using the diptest R package[1].

161

162

163 **Fluorescence *in situ* hybridization**

164 The preparation of FISH probes was performed as described previously (24, 25). For this
165 purpose, pRBBm99 or pBC16 (26) vector DNA was digested in fragments ranging from 20–
166 280 bp using 4-base cutting restriction enzymes (*Alu*I, *Hae*III, *Mse*I, *Msp*I, *Rsa*I and *Sau*3AI) and
167 subsequently labeled with Cy3-dCTP (GE Healthcare, Amersham, UK) by terminal transferase
168 (New England Biolabs, Ipswich, MA, USA). FISH was performed as detailed previously (27),
169 with modifications made in washing steps (25). In addition, fixed cells were treated with
170 lysozyme (2 mg/mL) for 10 min at 37°C. Fixation and staining were performed on a flat, 18-well

[1] http://CRAN.R-project.org/package=diptest

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

171 µ-slide coated with poly-L-lysine (ibidi Labware). The detection of cell boundaries was
172 performed using TLM-Tracker image processing software (23).

173

174

175 **Data analysis and mathematical modeling**

176 Data analysis and mathematical modeling were performed using R (R Development Core Team,
177 2010), Bioconductor (28) and simbTUM[2] software. Employed R packages were listed under the
178 respective methods.

[2] http://sourceforge.net/projects/simbtumtum/

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

RESULTS

Heterogeneity of recombinant GFP production by *B. megaterium*

We previously observed a continuous population heterogeneity during high level GFP production by *B. megaterium* growing in a bioreactor (11). In order to understand the dynamics of this phenomenon, a growing population was further investigated at the single cell level. Firstly, recombinant gene expression of a *B. megaterium* culture containing the GFP production plasmid, pSSBm85 (10), was induced by the addition of xylose in the early exponential growth phase, and cells analyzed afterwards by flow cytometry at various time points. As expected, prior to the induction of gene expression, cells did not produce a significant amount of GFP product. However, a large fraction of cells (about 72%) showed very low GFP levels, indicating a basal expression rate by the P_{xylA} promoter in the absence of xylose (Fig. 1A). With time, an increasing fraction of cells began to produce high amounts of GFP; however, for considerable subpopulation proportions of 29.1% (one h after induction), 16.9 % (two h after induction) and 11.1% (three h after induction), respectively, no or low levels of GFP protein were detected (Fig. 1B – D). Five h after xylose addition and upon entering the stationary growth phase, major parts of the population shifted into a production state (Fig. 1E). Finally, in the late stationary phase, the recombinant GFP production rate became reduced again and regressed towards the initial GFP production pattern observed prior to xylose addition (Fig. 1A + F). Thus, a marked protein production heterogeneity was obvious during the plasmid-based recombinant production of GFP in *B. megaterium*.

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

Bistability is not mediated by positive feedback via the xylose transporter

The non-uniform gene expression in clonal cell cultures that results in distinct subpopulations is commonly known as bistability. This phenomenon is usually caused by positive feedback loops in gene-regulatory circuits (29). Bistable behavior was also observed for the structurally similar *lac* operon of *E. coli* after the induction of gene expression with lactose (30). In the *lac* system, the repressor LacI becomes inactivated by lactose and derepresses genes for lactose catabolism. Here, bistability is mediated by positive feedback via the increased formation of the lactose permease, LacY, which facilitates lactose uptake into the cell, which, in turn, inhibits repression by LacI. In order to test if the outlined scenario applies to our observation for the *B. megaterium* *xyl* operon, we investigated this transporter-inducer feedback loop in our system. Expression levels of the xylose repressor gene, *xylR,* and the xylose transporter gene, *xylT,* in high- and low-producing cells were analyzed in greater detail. Induced cells were sorted according to their GFP content into two fractions. Firstly, sorted low-producing cells were plated onto A5 media agar plates supplemented with 0.5% xylose. However, it became obvious that these cells were still capable of producing large quantities of GFP, which excluded mutations or complete plasmid loss. The qRT-PCR analyses of sorted cells from both subpopulations revealed mRNA ratios of the genes involved in the expression system. In high-producing cells, the expression of *xylR* and *gfp* were induced about 150- and 500-fold, respectively, while the xylose transporter gene, *xylT*, remained unaffected (Fig. 2). Although, *xylT* is the last gene of the *xylABT* operon, it seems to be uncoupled from the XylR-mediated regulation of *xylA*. This finding can be explained by the existence of an internal transcriptional terminator within the operon upstream of *xylT* (31). In agreement, *xylT* mutants of *B. megaterium* continued to show the observed culture heterogeneity

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

(unpublished data). Thus, in contrast to the *lac* system, the observed phenotype seems not be mediated via induction of the sugar transporter. The observation that high amounts of GFP product correlated with strong induction ratios of both *gfp* and *xylR* led to the idea that the arising heterogeneity is a gene dose effect, most likely mediated by unevenly distributed plasmids.

Unequal distribution of recombinant plasmids to daughter cells after cell division

In order to obtain deeper insights into the dynamics of the emerging heterogeneity in the bacterial population, time-lapse microscopy movies were recorded of a growing microcolony producing GFP in response to xylose. *B. megaterium* is a bacterium particularly well-suited for live cell imaging due to its large size relative to other bacteria. We observed that a few scattered cells produced GFP during the first 2 h after xylose addition (Fig. 3A). However, what was particularly striking in regards to GFP production by these cells was that cell division in a GFP-producing cell was followed by one daughter cell continuing GFP production, while the second, seemingly genetically identical, daughter cell pausing in its production of GFP. In addition, not only was GFP production interrupted in the second daughter cell, GFP abundance even decreased, suggesting its dilution or even its degradation after cell division (Fig. 3B – E). Such phenomena were observed, almost without exception, for all cells in the growing *B. megaterium* microcolony (Movie S1). As a result, some individual cell lines produced large amounts of GFP while the rest of the population remained low producers (Fig. 3F). We hypothesized that such an observation could only be explained by the unequal distribution of cellular components required for recombinant protein production. However, during growth of the microcolony, an increasing

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

246 number of cell lines apparently reached the high-producing state, which finally lead to a shift of
247 the overall population towards GFP production.

248 Possible candidates that might cause such an uneven distribution of cellular components resulting
249 in the observed production differences of GFP were either the recombinant plasmid itself and/or
250 the XylR repressor. To investigate the influence of XylR, we first constructed a plasmid lacking
251 the *xylR* gene, named pKMMBm5. After transformation of our wild-type strain, we did not obtain
252 a real null-mutant because one *xylR* copy was still present in the genome of DSM319. However,
253 the number of free XylR proteins in this strain in relation to the available XylR-binding sites on
254 the multi-copy plasmid was postulated to be negligible. This was confirmed by analyzing a
255 completely *xylR* deficient mutant, WH321 (*ΔxylR*) (32) carrying pKMMBm5, for which culture
256 heterogeneity remained (Fig. S2). In agreement with this assumption, *B. megaterium* DSM319
257 carrying the pKMMBm5 plasmid permanently produced GFP in the absence of xylose. An
258 analysis of time-lapse movies revealed that, in this case, a separation into high- and low-producer
259 cell lines also took place (Movie S2). Thus, the behavior of these strains was virtually identical to
260 that of the parental strain expressing *xylR* from the pSSBm85 plasmid (Movie S1). In order to
261 monitor the distribution of recombinant plasmid and XylR repressor to daughter cells during cell
262 division, we fused XylR at its C-terminus to the red fluorescent protein, mCherry. The resulting
263 recombinant plasmid was termed pKMMBm2 (Fig. S1). With this approach, it was possible to
264 determine the relationship between GFP production and XylR abundance by measuring green and
265 red fluorescence, respectively. Additionally, the *in vivo* visualization and localization of plasmids
266 was possible by the XylR-mCherry bound to plasmid-encoded, XylR binding sites. Our approach
267 was sufficiently sensitive, with only two XylR binding sites per plasmid, due to the multi-copy
268 nature of the used plasmid. We were fully aware of the localization problems related to mCherry

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

fusion proteins (33). Consequently, an independent experimental approach via FISH was also employed (see below).

Time-lapse movies with *B. megaterium* carrying the pKMMBm2 plasmid revealed a unique spatial expression pattern. After the induction of gene expression, GFP production by various cell lineages was identical to the parental strain, with native XylR encoded by the pSSBm85 plasmid. In addition, some cells then generated strong red tips, indicating XylR-mCherry-plasmid complexes asymmetrically distributed to the cell poles, while other cells showed a weak, but equal, distribution of XylR-mCherry. On closer inspection of the chronological sequence of images, red tips were seen to frequently occur at the marginal cells of growing cell chains (Fig. 4A – D, Movie S3, S4). At first glance, this surprising pattern strongly suggested that after cell division, XylR-plasmid complexes accumulated at the old cell poles and that segregation of plasmids by free diffusion to the new cell pole was somehow inhibited. At a later growth phase, some red-tip cells became deeply red, lacked GFP, and stopped dividing, while neighboring producer cells continued to grow (Fig. 4E, Movie S5). The latter, deeply red cells appeared to accumulate plasmid XylR complexes while GFP became degraded. The noted cessation of cell division was in good agreement with the observation that, in later growth phases, cells with strong red fluorescence were almost exclusively found at the end of cell chains (Fig. S5A).

Cell lineage analysis revealed asymmetric plasmid distribution and plasmid accumulation at cell poles

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

In order to better understand the nature of plasmid transmission, cell lineage analyses of a growing microcolony of *B. megaterium* DSM319 carrying the pKMMBm2 plasmid (producing the XylR-mCherry variant) were performed. For this reason, image analyses of successive frames derived from time-lapse movies were carried out (23). This approach involved a segmentation step, in which cells were spatially recognized, and a tracking step, in which cells were temporally connected from frame to frame. In addition, red and green fluorescence in relation to the pole age were determined for each cell at each time point. The final result was a cell lineage tree giving detailed information about the dynamics of cell divisions events, plasmid distribution and GFP production. The appearance of red cell lines in the tree reflected XylR-mCherry-plasmid complexes and the tree also showed the corresponding asymmetrical, plasmid distribution behavior (Fig. 5A and S3).

In order to determine the fraction of plasmids inherited from the mother cell, the ratio between the red fluorescence of each daughter cell divided by the sum of the red fluorescence of both daughter cells was calculated. Thus, values around 0.5 reflected equal distribution while significantly lower and higher values indicated unequal distributions. Histograms of the calculated segregation ratios showed a bimodal distribution for the strongly fluorescent cells, while mother cells with low fluorescence revealed normal distributions (Fig. 5B). These results confirmed the observation from the preceding section that asymmetric red tip cells were only produced in the presence of strong XylR-mCherry signals. Moreover, the applied co-localization technique was not sensitive enough to visualize single scattered plasmids within the cell, so strong noise around the mean was to be expected at low signals.

The Hartigan's dip test was applied to quantify the observed bimodality (34). For strong fluorescence signals (5% of the highest fluorescent mother cells), a *P* value of 6.3E-4 was found,

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

313 indicating a high significance for a bimodal distribution. The applied mixture model used to
314 characterize the total population (Fig. 5B) provided evidence for a distribution mechanism
315 dependent on the cellular plasmid concentration. When the plasmid concentration was rather low,
316 both daughter cells usually received almost the same number of plasmids. At higher cellular
317 plasmid concentrations, an unequal distribution led to considerable differences in the plasmid
318 concentration of the daughter cells. Unequally distributed plasmid complexes were found at both
319 the old and new cell pole. Thus, the described phenomenon seemed to be of a stochastic nature
320 and not related to cell pole aging.

321

322

323 **Retention of plasmids is linked to protein biosynthesis**

324 In order to measure the influence of protein production on the distribution of the plasmids, we
325 performed FISH analyses for the direct detection of cellular plasmids (25). Using Cy3-labeled
326 plasmid probes, we directly visualized the pSSBm85 plasmid in cells during the exponential
327 growth phase. Results clearly confirmed that under conditions of induced GFP production,
328 plasmid DNA became asymmetrically localized to one of the cell poles in the largest proportion
329 of cells (Fig. 6A). Moreover, a few cells exhibited an extremely strong FISH signal, confirming
330 plasmid accumulation in these cells as detected by time-lapse microscopy. In contrast,
331 recombinant *B. megaterium* cells grown without the addition of the inducing xylose were largely
332 missing the observed unequal distribution of plasmids (Fig. 6B). This led to the conclusion that
333 protein biosynthesis strongly influenced the distribution mechanism of the plasmids. As a further
334 control, we performed FISH experiments with *B. megaterium* carrying the 4.6 kB pBC16
335 plasmid, which is a precursor of our 7.2 kB pSSBm85 GFP production plasmid (26). In this case,

Accepted Manuscript Posted Online

a solely equal plasmid distribution was also observed. Interestingly, signals were detected as distinct groups of foci distributed all over the cell (Fig. 6C). We also tested pBC16 in *Bacillus subtilis* with similar results (Fig. 6D).

Nucleoid overlaps with plasmid location but is restricted to central parts of the cell

Next, we analyzed if plasmid localization was influenced by the chromosomal DNA of the nucleoid. It has been shown that for smaller bacteria such as *E. coli*, high-copy plasmids simply occupy the nucleoid-free space within the bacterial cell (35). For the *in vivo* co-localization of plasmid and nucleoid DNA, the pKMMBm2 system was used in combination with DAPI staining of DNA. Various stages of cell division were analyzed. Prior to cell division, the nucleoid was localized in the central part of the cell (Fig. 7, 1st row). At this stage, the XylR-mCherry-plasmid complexes were mainly distributed all over the cell, including the cell poles. In the course of cell division, plasmid complexes were subsequently localized at one of the cell poles of the daughter cells, whereas the nucleoid was segregated close to the new poles (Fig. 7, 2nd row). Finally, after cell division, the whole cell was occupied by the plasmid DNA whereas the nucleoid DNA was restricted to the middle part. Due to asymmetric plasmid distribution, terminal cells of cell chains, especially, contained high amounts of XylR-plasmid complexes (Fig. 7C, 3rd row).

Influence of unequal plasmid distribution on protein production dynamics

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

The possible influence of population heterogeneity on heterologous protein production processes was, subsequently, systematically analyzed. Two striking effects with major influences on protein production dynamics have been observed so far. Firstly, flow cytometry data showed that the degree of heterogeneity was highest during the first three h after induction, comprising the early exponential growth phase (Fig. 1). During this stage, the population consisted of high-producing cells with a high plasmid copy number and low-producing cells with low plasmid copy number. Secondly, a subpopulation of cells with highly accumulated plasmids was formed which finally lead to a termination of cell division (Fig. 4E). As a result, after induction, only the subpopulation carrying many plasmids produced large amounts of product. Time-lapse movies showed that these early producers showed reduced growth and cells with too many accumulated plasmids even stopped growing, whereas the subpopulation carrying fewer plasmids continued to grow (Movie S5). In the short term, both a shift and selection towards low producers were expected. In the long term, the number of low-producing cells with accumulated plasmids was expected to increase, which is not desirable for biotechnological processes. Consequently, we expected a complex relationship between growth rate, plasmid abundance and GFP production rate. In order to experimentally investigate this effect in more detail, we measured the growth behavior and GFP production of batch cultures in a micro bioreactor at various times after the induction of GFP production by xylose. The first striking observation was that, independently of the time of gene induction, cells stopped growing after 4 – 5 h. At the same time, GFP production stopped and, in due course, GFP became actively degraded. A disintegration of GFP could be excluded as the *in vivo* half-life of GFP in *B. megaterium* was determined to be at least 9.5 h (Fig. S4). The effect described above, namely the shift between high and low producers, could indeed be observed in the form of a delayed production showing an inflection point, especially for early induction times (Fig. 8). This production behavior fitted a mathematical model dealing with two

Applied and Environmental Microbiology

Accepted Manuscript Posted Online

subpopulations (supplemental material, Fig. S6). Another pertinent observation was that for early time points, the noise of the measured GFP production was greater, suggesting a stochastic plasmid distribution process. In summary, the observed unequal plasmid distribution obviously influences the dynamics and yield of the recombinant GFP production process in *B. megaterium*.

Model of plasmid distribution in recombinant protein-producing *B. megaterium*

We integrated our results into a model that explained the observed population heterogeneity during recombinant protein production (Fig. 9). Starting from a state of equal plasmid distribution, cells began to produce GFP. Prior to cell division, a major proportion of plasmids remained at the cell poles. Subsequently, this polar fixation led to daughter cells with high and low plasmid abundance. In turn, these differences in gene dose caused by different copy numbers directly influenced the degree of recombinant protein production. As a consequence, cell lines with high- and low-producing cells arose. The observed, characteristic pattern of neighboring producing- and non-producing cells in combination with red tipped cells in growing cell chains of *B. megaterium* mirrored our model directly (Fig. 3 and 4, Movie S1 – S4).

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

DISCUSSION

We provide evidence to show that unequal plasmid distribution by polar fixation is the major reason for population heterogeneity during recombinant plasmid-based protein production in *B. megaterium*. Our postulated mechanism of plasmid distribution contradicts the commonly assumed free diffusion of multi-copy plasmids with respect to the protein production conditions used. In our study, an uneven distribution of plasmids led to an immediate decrease in product production rates in daughter cells with lower plasmid concentrations. Given that in the stationary phase most cells were apparent producers, one might consider this effect as an insignificant, transient phenomenon. However, in growing cells likely present in chemostats or fed-batch bioreactors, a stable subpopulation of low producers permanently exists (11). Moreover, high producers accumulated plasmids and ceased growing, which finally led to a breakdown in production.

Such a described phenomenon should not be confused with the well-known segregational instability describing complete plasmid loss in subpopulations (36); in contrast to our observed unequal plasmid distribution, this effect is commonly observed under non-selective conditions. However, the factors affecting segregational instability may also be related to our observations. Recently, a relationship between gene expression and plasmid segregation was shown for ColE1-like plasmids in *E. coli* (37): it was argued that transcription and translation interfere with plasmid replication leading to low copy numbers and finally to plasmid loss. We did not detect complete plasmid loss and a highly-related plasmid has been described as being stably maintained in *B. megaterium*, even without selective pressure (38). We eventually concluded that the transcriptional and translational machinery slows down the distribution of plasmids in *B. megaterium*. In another study in *E. coli*, strong transcription was shown to drive plasmids to a

Applied and Environmental Microbiology

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

location of high transcriptional activity close to the cell poles (39). In our *B. megaterium* system, rather larger patches of plasmid complexes were detected which remained at one cell pole after cell division, with a preference for when the expression system was induced. This is in good agreement with results reported for *B. subtilis,* where different studies showed that active ribosomes are located at the cell poles of this bacterium (40), and that the cellular localization of DNA and ribosomes overlap, indicating a coupled transcription/translation process (41).

However, despite the formation of plasmid clusters at the cell poles, it is somehow ensured that at least one plasmid enters the new cell after cell division. Although it was previously shown that plasmids are preferentially located near the cell poles and grouped in clusters of foci (25, 35), it is generally assumed that high-copy plasmids without a partitioning system segregate by random diffusion. In contrast, low-copy plasmids segregated by a partitioning system were found localized near the center or quarter position of the cell, and behaved like mini-chromosomes (42).

We also tested the distribution behavior of the much smaller pBC16 precursor plasmid, which belongs to a widely dispersed plasmid family of small, multi-copy plasmids of Gram-positive bacteria mediating naturally occurring tetracycline resistance among bacilli and other soil bacteria (43). In the latter case, several plasmid foci were found randomly distributed all over the cytoplasm and, thus, we argue that plasmid size may play a crucial role in the formation of distinct foci. Due to the fact that plasmids generally occurred in clustered foci, and that free diffusion within living cells is generally constrained, the random diffusion model was previously questioned; it was postulated that chromosome-encoded proteins recruit at least one plasmid copy from the cluster and mediate the partitioning via a, so far, unknown mechanism (44, 45).

Previous reports revealed that the pUB110 plasmid, which is related to our employed expression plasmid, was found bound to the cell membrane via specific attachment regions (46, 47).

Applied and Environmental Microbiology

Accepted Manuscript Posted Online

444 However, confocal laser scanning microscopy (CLSM) analyses in combination with 3D image
445 stack reconstruction suggested that the plasmids were not bound to the membrane but, instead,
446 arranged in a central region of the transaxial plane (Fig. S5B). A plausible explanation may be
447 the reduced number of membrane attachment sites in pKMMBm2 compared to those found in the
448 original plasmid (Fig. S1).

449 In the long term, the accumulation of plasmids negatively affected the cell division rate of the
450 bacteria. One reason could be an increased metabolic burden, which may bring the production
451 capacity of a cell to its limit. Another explanation might be associated with bacterial cell aging.
452 Since plasmids also accumulated at the old poles, it can be expected that the vitality of mother
453 cells will decrease with increasing number of generations (48, 49). However, in this study we
454 showed that the cell pole was randomly chosen, and therefore, this effect should only be
455 marginal.

456 So far, very little is known about the distribution of high-copy number plasmids, especially in
457 Gram-positive bacteria (50). Our findings show a polar fixation for the plasmid used in the
458 production of high amounts of recombinant proteins in *B. megaterium*. However, the molecular
459 mechanisms behind these observations remain largely unknown. Future studies are required to
460 shed light onto the molecular principles involved.

461

462

463 **ACKNOWLEDGEMENTS**

464 This work was funded by the German Research Foundation (DFG), within the priority program,
465 SPP1617 “Phenotypic heterogeneity and sociobiology of bacterial populations”.

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

466 We would like to thank Maria Höxter and Lothar Gröbe for their support in cell sorting and
467 Irmtraud Steinmetz for confocal laser scanning microscope imaging. We are also grateful to Joe
468 Pogliano for providing the FISH protocol and to Peter Graumann for providing the pJS72
469 plasmid. Additionally, we owe thanks to Gerald Seidel for supplying the *xylR* mutants of
470 *B. megaterium* (strain WH321) and to Birte Bieser for critical reading of the manuscript.

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

REFERENCES

1. **Dong H, Zhang D.** 2014. Current development in genetic engineering strategies of *Bacillus* species. Microb. Cell Fact. **13:**63.

2. **Liu L, Liu Y, Shin H-D, Chen RR, Wang NS, Li J, Du G, Chen J.** 2013. Developing *Bacillus* spp. as a cell factory for production of microbial enzymes and industrially important biochemicals in the context of systems and synthetic biology. Appl. Microbiol. Biotechnol. **97:**6113–6127.

3. **Terpe K.** 2006. Overview of bacterial expression systems for heterologous protein production: from molecular and biochemical fundamentals to commercial systems. Appl. Microbiol. Biotechnol. **72:**211–222.

4. **Bunk B, Schulz A, Stammen S, Münch R, Warren MJ, Rohde M, Jahn D, Biedendieck R.** 2010. A short story about a big magic bug. Bioeng. Bugs. **1:**85–91.

5. **Vary PS, Biedendieck R, Fuerch T, Meinhardt F, Rohde M, Deckwer W-D, Jahn D.** 2007. *Bacillus megaterium*–from simple soil bacterium to industrial protein production host. Appl. Microbiol. Biotechnol. **76:**957–967.

6. **Rygus T, Hillen W.** 1991. Inducible high-level expression of heterologous genes in *Bacillus megaterium* using the regulatory elements of the xylose-utilization operon. Appl. Microbiol. Biotechnol. **35:**594–599.

7. **Rygus T, Scheler A, Allmansberger R, Hillen W.** 1991. Molecular cloning, structure, promoters and regulatory elements for transcription of the *Bacillus megaterium* encoded regulon for xylose utilization. Arch. Microbiol. **155:**535–542.

Accepted Manuscript Posted Online

8. **Malten M, Hollmann R, Deckwer W-D, Jahn D.** 2005. Production and secretion of recombinant *Leuconostoc mesenteroides* dextransucrase DsrS in *Bacillus megaterium*. Biotechnol. Bioeng. **89:**206–218.

9. **Eppinger M, Bunk B, Johns MA, Edirisinghe JN, Kutumbaka KK, Koenig SS, Creasy HH, Rosovitz MJ, Riley DR, Daugherty S, Martin M, Elbourne LD, Paulsen I, Biedendieck R, Braun C, Grayburn S, Dhingra S, Lukyanchuk V, Ball B, Ul-Qamar R, Seibel J, Bremer E, Jahn D, Ravel J, Vary PS.** 2011. Genome sequences of the biotechnologically important *Bacillus megaterium* strains QM B1551 and DSM319. J. Bacteriol. **193:**4199–4213.

10. **Stammen S, Müller BK, Korneli C, Biedendieck R, Gamer M, Franco-Lara E, Jahn D.** 2010. High-yield intra- and extracellular protein production using *Bacillus megaterium*. Appl. Env. Microbiol. **76:**4037–4046.

11. **Biedendieck R, Yang Y, Deckwer W-D, Malten M, Jahn D.** 2007. Plasmid system for the intracellular production and purification of affinity-tagged proteins in *Bacillus megaterium*. Biotechnol. Bioeng. **96:**525–537.

12. **Scholz O, Thiel A, Hillen W, Niederweis M.** 2000. Quantitative analysis of gene expression with an improved green fluorescent protein. Eur. J. Biochem. **267:**1565–1570.

13. **Kaltwasser M, Wiegert T, Schumann W.** 2002. Construction and application of epitope- and green fluorescent protein-tagging integration vectors for *Bacillus subtilis*. Appl. Env. Microbiol. **68:**2624–2628.

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

14. **Soufo HJD, Graumann PL.** 2010. *Bacillus subtilis* MreB paralogues have different filament architectures and lead to shape remodelling of a heterologous cell system. Mol. Microbiol. **78:**1145–1158.

15. **Hahne F, LeMeur N, Brinkman RR, Ellis B, Haaland P, Sarkar D, Spidlen J, Strain E, Gentleman R.** 2009. flowCore: a Bioconductor package for high throughput flow cytometry. BMC Bioinformatics **10:**106.

16. **Sarkar D, Le Meur N, Gentleman R.** 2008. Using flowViz to visualize flow cytometry data. Bioinformatics **24:**878–879.

17. **Pfaffl MW.** 2001. A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Res. **29:**e45.

18. **Tichopad A, Dilger M, Schwarz G, Pfaffl MW.** 2003. Standardized determination of real-time PCR efficiency from a single reaction set-up. Nucleic Acids Res. **31:**e122.

19. **Spiess A-N, Feig C, Ritz C.** 2008. Highly accurate sigmoidal fitting of real-time PCR data by introducing a parameter for asymmetry. BMC Bioinformatics **9:**221.

20. **Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A, Speleman F.** 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. Genome Biol. **3:**research0034.1–0034.11.

21. **Ritz C, Spiess A-N.** 2008. qpcR: an R package for sigmoidal model selection in quantitative real-time polymerase chain reaction analysis. Bioinformatics. **24:**1549–1551.

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

Accepted Manuscript Posted Online

22. **Young JW, Locke JCW, Altinok A, Rosenfeld N, Bacarian T, Swain PS, Mjolsness E, Elowitz MB.** 2012. Measuring single-cell gene expression dynamics in bacteria using fluorescence time-lapse microscopy. Nat. Protoc. **7:**80–88.

23. **Klein J, Leupold S, Biegler I, Biedendieck R, Münch R, Jahn D.** 2012. TLM-Tracker: software for cell segmentation, tracking and lineage analysis in time-lapse microscopy movies. Bioinformatics **28:**2276–2277.

24. **Dernburg AF, Sedat JW.** 1998. Mapping three-dimensional chromosome architecture *in situ*. Methods Cell Biol. **53:**187–233.

25. **Pogliano J, Ho TQ, Zhong Z, Helinski DR.** 2001. Multicopy plasmids are clustered and localized in *Escherichia coli*. Proc. Natl. Acad. Sci. USA **98:**4486–4491.

26. **Bernhard K, Schrempf H, Goebel W.** 1978. Bacteriocin and antibiotic resistance plasmids in *Bacillus cereus* and *Bacillus subtilis*. J. Bacteriol. **133:**897–903.

27. **Jensen RB, Shapiro L.** 1999. The *Caulobacter crescentus smc* gene is required for cell cycle progression and chromosome segregation. Proc. Natl. Acad. Sci. USA **96:**10661–10666.

28. **Gentleman RC, Carey VJ, Bates DM, Ben Bolstad, Dettling M, Dudoit S, Ellis B, Gautier L, Ge Y, Gentry J, Hornik K, Hothorn T, Huber W, Iacus S, Irizarry R, Leisch F, Li C, Maechler M, Rossini AJ, Sawitzki G, Smith C, Smyth G, Tierney L, Yang JY, Zhang J.** 2004. Bioconductor: open software development for computational biology and bioinformatics. Genome Biol. **5:**R80.

29. **Dubnau D, Losick R.** 2006. Bistability in bacteria. Mol. Microbiol. **61:**564–572.

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

30. **Ozbudak EM, Thattai M, Lim HN, Shraiman BI, Van Oudenaarden A.** 2004. Multistability in the lactose utilization network of *Escherichia coli*. Nature **427:**737–740.

31. **Schmiedel D, Kintrup M, Küster E, Hillen W.** 1997. Regulation of expression, genetic organization and substrate specificity of xylose uptake in *Bacillus megaterium*. Mol. Microbiol. **23:**1053–1062.

32. **Rygus T, Hillen W.** 1992. Catabolite repression of the *xyl* operon in *Bacillus megaterium*. J. Bacteriol. **174:**3049–3055.

33. **Landgraf D, Okumus B, Chien P, Baker TA, Paulsson J.** 2012. Segregation of molecules at cell division reveals native protein localization. Nat. Methods. **9:**480–482.

34. **Reyes-Lamothe R, Tran T, Meas D, Lee L, Li AM, Sherratt DJ, Tolmasky ME.** 2014. High-copy bacterial plasmids diffuse in the nucleoid-free space, replicate stochastically and are randomly partitioned at cell division. Nucleic Acids Res. **42:**1042-1051.

35. **Gordon GS, Sitnikov D, Webb CD, Teleman A, Straight A, Losick R, Murray AW, Wright A.** 1997. Chromosome and low copy plasmid segregation in *E. coli*: visual evidence for distinct mechanisms. Cell **90:**1113–1121.

36. **Hartigan JA, Hartigan PM.** 1985. The Dip Test of Unimodality. Ann. Stat. **13:**70–84.

37. **Summers DK**. 1996. The Biology of Plasmids, p. 68-82, Blackwell Science Ltd., Oxford, UK.

38. **Popov M, Petrov S, Nacheva G, Ivanov I, Reichl U.** 2011. Effects of a recombinant gene expression on ColE1-like plasmid segregation in *Escherichia coli*. BMC Biotechnol. **11:**18.

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

572 39. **Sánchez-Romero M-A, Lee DJ, Sánchez-Morán E, Busby SJW.** 2012. Location and
573 dynamics of an active promoter in *Escherichia coli* K-12. Biochem. J. **441:**481–485.

574 40. **Sojka L, Fucík V, Krásný L, Barvík I, Jonák J.** 2007. YbxF, a protein associated with
575 exponential-phase ribosomes in *Bacillus subtilis*. J. Bacteriol. **189:**4809–4814.

576 41. **Mascarenhas J, Weber MH, Graumann PL.** 2001. Specific polar localization of
577 ribosomes in *Bacillus subtilis* depends on active transcription. EMBO Rep. **2:**685–689.

578 42. **Meinhardt F, Stahl U, Ebeling W.** 1989. Highly efficient expression of homologous and
579 heterologous genes in *Bacillus megaterium*. Appl. Microbiol. Biotechnol. **30:**343–350.

580 43. **Polak J, Novick RP.** 1982. Closely related plasmids from *Staphylococcus aureus* and soil
581 bacilli. Plasmid. **7:**152–162.

582 44. **Million-Weaver S, Camps M.** 2014. Mechanisms of plasmid segregation: have multicopy
583 plasmids been overlooked? Plasmid **75:**27–36.

584 45. **Nordström K, Gerdes K.** 2003 Clustering versus random segregation of plasmids lacking a
585 partitioning function: a plasmid paradox? Plasmid **50:**95–101.

586 46. **Korn R, Winston S, Tanaka T, Sueoka N.** 1983. Specific *in vitro* binding of a plasmid to
587 a membrane fraction of *Bacillus subtilis*. Proc. Natl. Acad. Sci. USA. **80:**574–578.

588 47. **Tanaka T, Sueoka N.** 1983. Site-specific *in vitro* binding of plasmid pUB110 to *Bacillus*
589 *subtilis* membrane fraction. J. Bacteriol. **154:**1184–1194.

590 48. **Stewart EJ, Madden R, Paul G, Taddei F.** 2005 Aging and death in an organism that
591 reproduces by morphologically symmetric division. PLoS Biol. **3:**e45.

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

592 49. **Lindner AB, Madden R, Demarez A, Stewart EJ, Taddei F.** 2008. Asymmetric
593 segregation of protein aggregates is associated with cellular aging and rejuvenation. Proc.
594 Natl. Acad. Sci. USA **105:**3076–3081.

595 50. **Wang JD, Rokop ME, Barker MM, Hanson NR, Grossman AD.** 2004. Multicopy
596 plasmids affect replisome positioning in *Bacillus subtilis*. J. Bacteriol. **186:**7084–7090.

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

TABLES

Table 1 Strains, plasmids and primers used in this study.

Strains:

Strain name	Description	Reference or source
B. megaterium DSM319	Wild-type strain	(9)
B. megaterium WH321	genomic *xylR* mutant	(32)
B. megaterium WH377	genomic *xylT* mutant	(31)
B. subtilis DSM402	*B. subtilis* subsp. *subtilis* strain 168 carrying the pBC16 plasmid (DSM No.: 23521)	DSMZ[1]

Plasmids:

Plasmid name	Description	Reference or source
pSSBm85	Promoter-optimized shuttle vector for xylose-inducible production of GFP; P_{xylA}-*gfp*	(10)
pRBBm99	Derivative of pSSBm85 lacking *xylR* and P_{xylA}-*gfp*	This work
pKMMBm1	Introduced XmaI site at the 5'-end of *xylR* and replaced *xylR* stop codon with a KpnI site in pSSBm85	This work
pKMMBm2	*mCherry* cloned into KpnI and AflII sites of pKMMBm1; *xylR-mCherry* fusion	This work
pKMMBm5	pSSBm85 Δ*xylR* mutant	This work

Accepted Manuscript Posted Online

pJS72	pETDuet vector (Novagen) containing *mCherry*	(14)
pBC16	natural plasmid isolate originating from *Bacillus cereus*, precursor for pSSBm85	(26)

Primers:

Primer name	**Sequence**[2]	
xylR-mCherry_for	5'-tatcaggtaccgtgagcaagggcgaggag-3'	This work
xylR-mCherry_rev	5'-tatcacttaagttacttgtacagctcgtccatg-3'	This work
xylT_for	5'-cggagggttactgtttggatatgac-3'	This work
xylT_rev	5'-ccgtgtgctaaagatcctaacccta-3'	This work
xylR_for	5'-gccttgtagatcgtcatcagcaaa-3'	This work
xylR_rev	5'-cgtatgctcctgcattagcttcat-3'	This work
gfp_for	5'-tctcggacacaaactcgagtacaac-3'	This work
gfp_rev	5'-ctgctagttgaacggatccatcttc-3'	This work
rpoB_for	5'-ggcgacgaagtagtaaaaggtgaga-3'	This work
rpoB_rev	5'-ggcatcctcatagttgtaaccatcc-3'	This work
gyrB_for	5'-tacatggtgtaggtgcctcagttgt-3'	This work
gyrB_rev	5'-acttttaagtcagcagccggtacac-3'	This work
heli1_for	5'-gatgtaatccatacgtcagctgtgc -3'	This work
heli1_rev	5'-ccggcttccttgatttataactgg-3'	This work
heli2_for	5'-gtcgaacctgtacacggtgacttttt-3'	This work
heli2_rev	5'-gcttccgtaataggtctgttcatgc-3'	This work

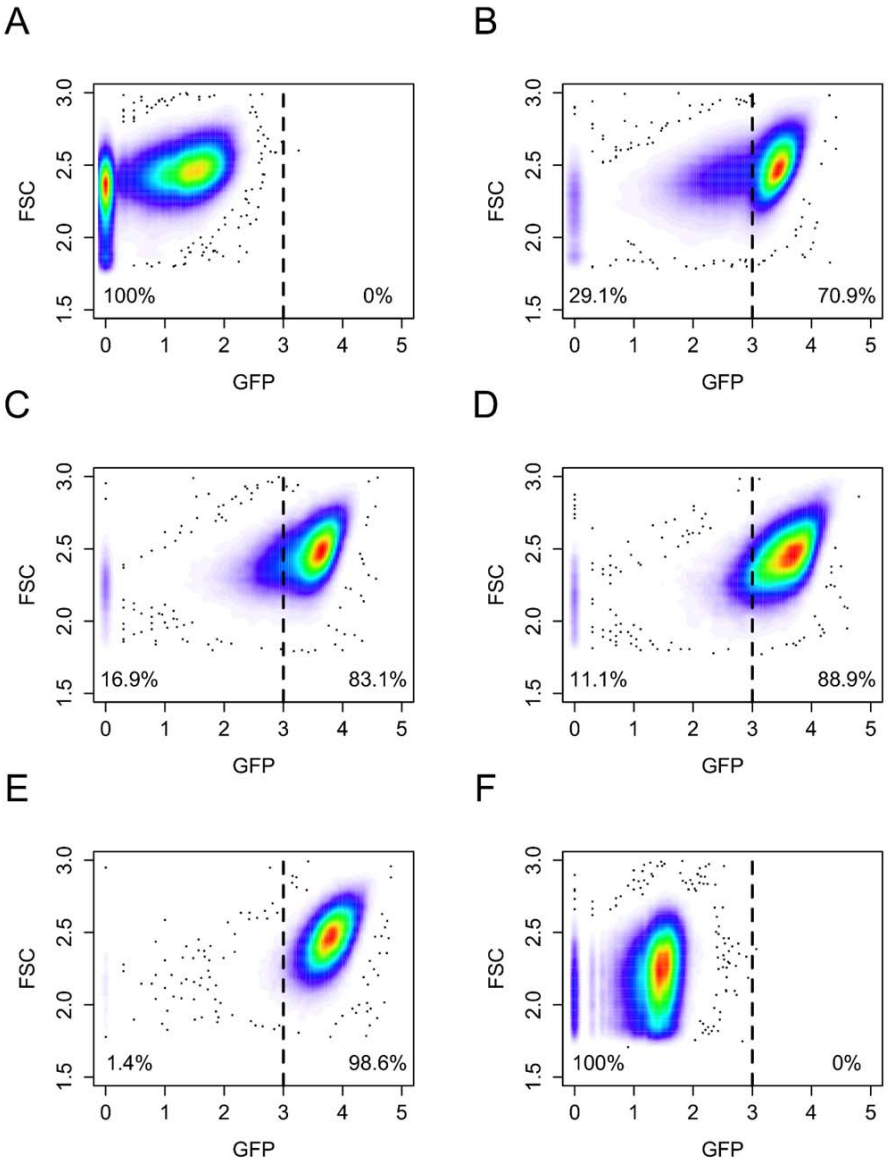
599 [1]DSMZ-German Collection of Microorganisms and Cell Cultures

600 [2]Specific restriction enzymes used for cloning are underlined.

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online
Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

601 **FIGURES**



602

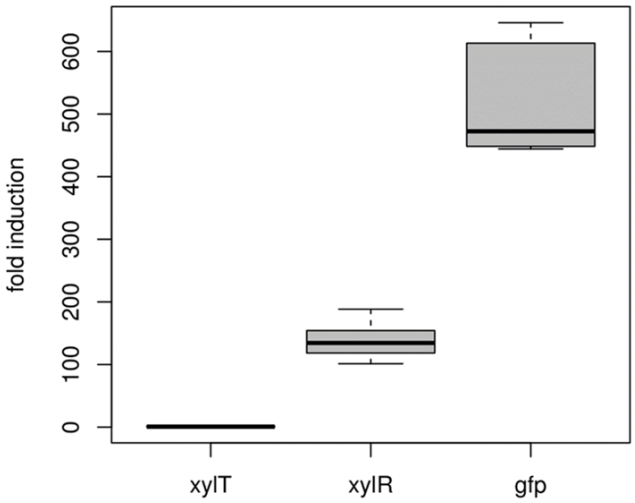
603 **Fig. 1** Time course of heterologous GFP production by recombinant *B. megaterium* and analysis

604 of arising subpopulations using flow cytometry. Before gene induction by the addition of xylose,

Accepted Manuscript Posted Online

605 the majority of the population did not produce GFP, with the exception of a small fraction of low-
606 producing cells (A). Between 1, 2 and 3 h after the induction of gene expression (B – D), the
607 population shifted towards high-producing cells, but with a significant subset of low producers.
608 After 5 h (E), when cells had already entered the stationary phase, almost all cells reached their
609 highest production levels. In the late stationary phase (22 h after induction), production levels
610 decreased and the population shifted towards its initial state (F). The dotted line indicates the two
611 gates "non/low GFP-producing cells" and "GFP-producing cells" (cut-off level) which were
612 defined at the beginning of the flow cytometry analyses. GFP: fluorescence intensity of GFP
613 given in arbitrary units; FSC: forward scatter.

614



615

616 **Fig. 2** Comparison of gene expression in sorted *B. megaterium* cells (producer compared to low
617 producer subpopulations) using quantitative RT-PCR. Genes encoding the main components of
618 the pSSBm85 expression system (*xylT*, *xylR* and *gfp*), were investigated. The xylose repressor,
619 *xylR*, and target gene, *gfp*, were induced about 150× and 600×, respectively. In contrast, the
620 xylose transporter, *xylT*, was expressed constitutively.

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

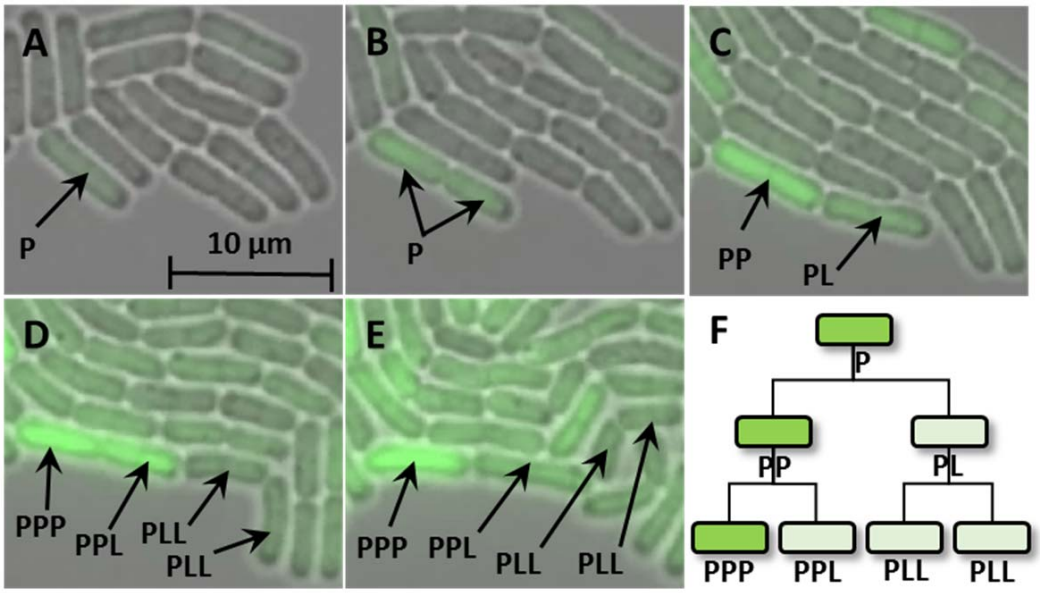
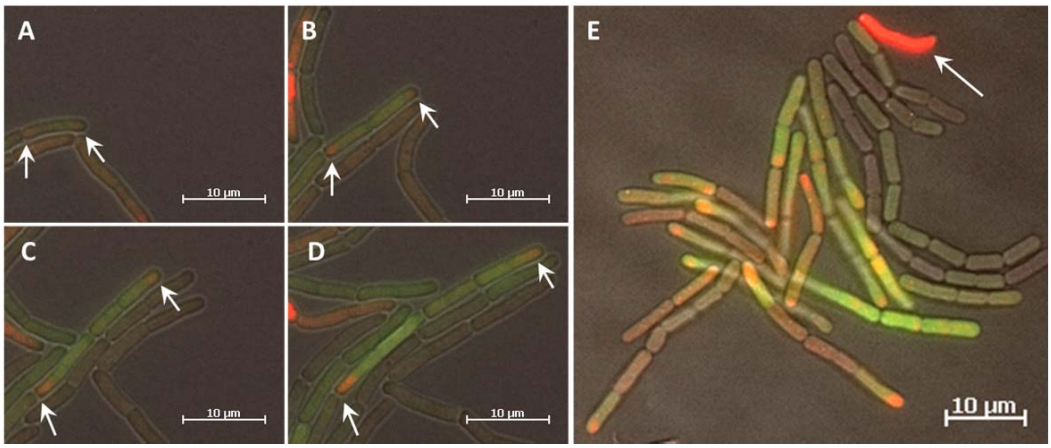


Fig. 3 Images taken from a time-lapse microscopy movie showing a growing, GFP-producing microcolony of *B. megaterium* carrying the pSSBm85 plasmid (A – E: 120, 150, 180, 210 and 260 min after the induction of gene expression by the addition of xylose). Each producer cell (P) generated daughter cells: a producer cell and a low-producing cell (L). The resulting cell lineage tree (F) schematically shows a lineage of high-producing cells.

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online



627

628 **Fig. 4** Time series of *B. megaterium* carrying the pKMMBm2 plasmid (A – D: 120, 160, 190 and
629 210 min after gene induction). Producer cells developed red tips (arrows) indicating XylR-
630 mCherry-plasmid complexes located at the old cell poles. E) Microcolony of high and low
631 producers. Occasionally, deeply red cells arose lacking GFP and which probably accumulated
632 plasmid-XylR complexes (arrow); these cells stopped proliferating.

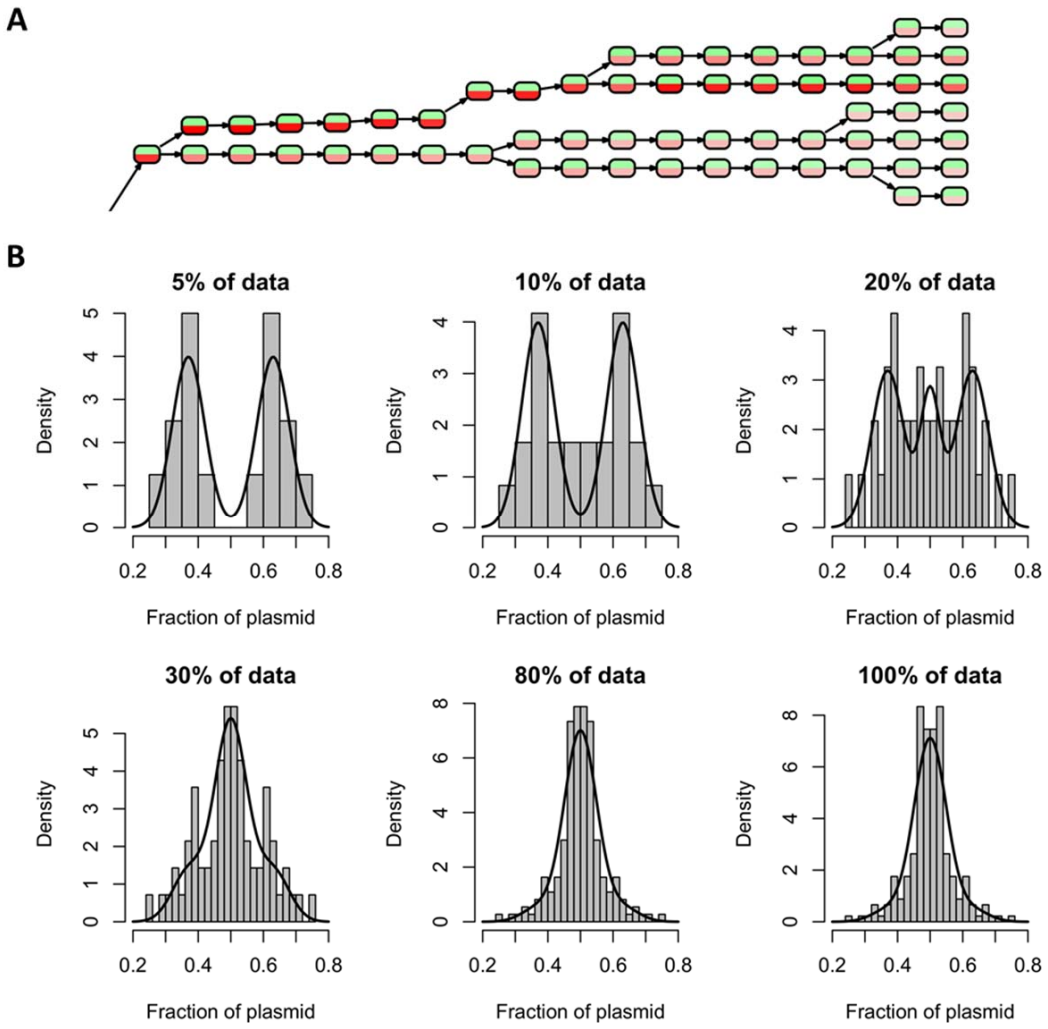
Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology



633

634 **Fig. 5:** A) Part of the reconstructed cell lineage tree derived from a time-lapse movie of *B.*
635 *megaterium* carrying the pKMMBm2 plasmid (Movie S3). Nodes represent single cells at
636 specific time points. Red and green coloring of nodes represent the respective fluorescence
637 intensities. The complete tree is available in the supplemental material (Fig. S3). B) Relationship
638 of bimodal plasmid distribution to the amount of plasmid per cell. Histograms show the fraction
639 of segregated plasmids in relation to the percentage of cells with the strongest XylR-mCherry

Accepted Manuscript Posted Online

signal. Bimodal distributions were found for 5 – 20 % of cells, all containing large amounts of XylR-mCherry. Fitted curves were generated using a mixture-model consisting of the 10% most fluorescing cells, and the rest. Highly fluorescing cells were fitted using a bi-normal distribution superposition of two normal distributions, localized at 0.5 ± 0.13 (and variance 0.05). The remaining mother cells with low fluorescence were fitted to one normal distribution (with mean 0.5 and variance 0.03).

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

AEM

Applied and Environmental Microbiology

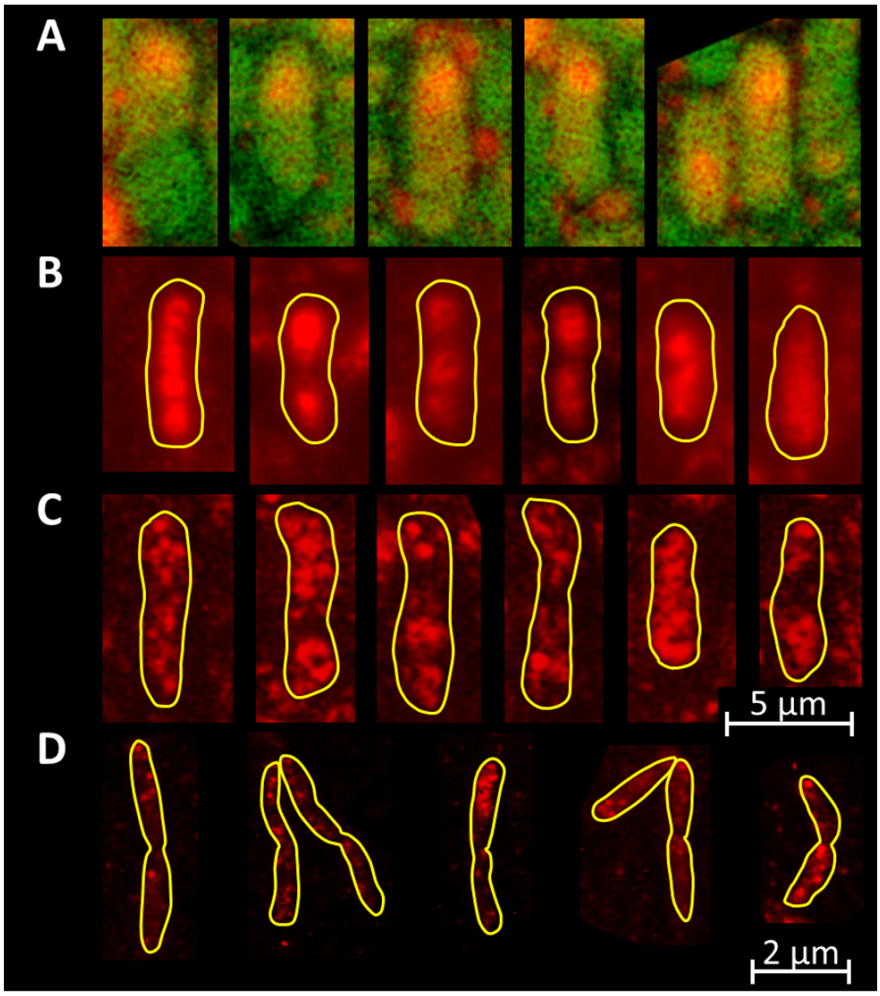


Fig. 6 Localization of plasmids using FISH. A) *B. megaterium* carrying pSSBm85, induced with xylose, showed an unequal distribution of plasmids. These preferentially remained at one cell pole. B) In the absence of gene induction, the pSSBm85 plasmid distribution is hardly affected. C) The smaller pBC16 precursor plasmid was found clustered in several foci distributed all over the cell. D) The same behavior was seen in *B. subtilis* carrying the pBC16 plasmid, resulting in distinct foci of plasmid accumulation.

Accepted Manuscript Posted Online

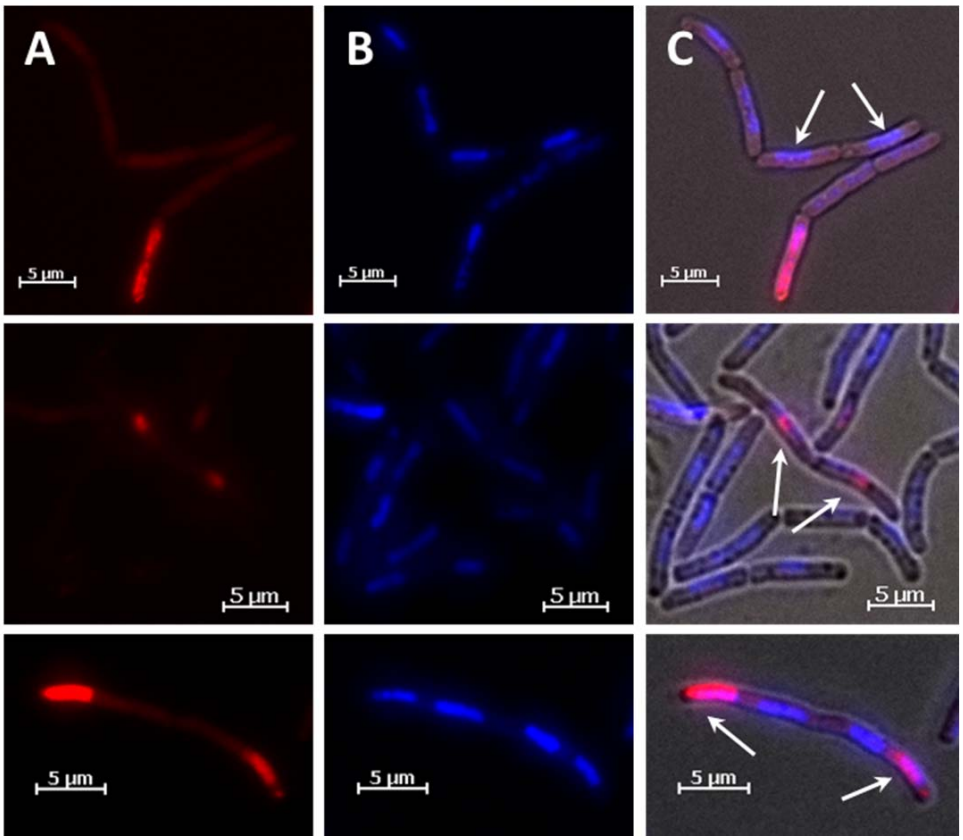


Fig. 7: Coordinated distribution of plasmid and chromosomal DNA in protein-producing *B. megaterium*. Different stages of plasmid and nucleoid positioning during cell division of *B. megaterium* carrying pKMMBm2 were visualized. The figure shows plasmid-XylR-mCherry complexes (column A), the nucleoid visualized via DAPI staining (column B) and the overlay of both images with the inclusion of the bright field image (column C). Before cell division, the nucleoid was localized in the central part of the cell, excluding the poles (arrows in row 1, column C) while XylR-mCherry was mainly distributed all over the cell. During cell division the plasmid-XylR-mCherry complexes were localized at the old cell poles (arrows in row 2, column

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology

Accepted Manuscript Posted Online

Applied and Environmental Microbiology

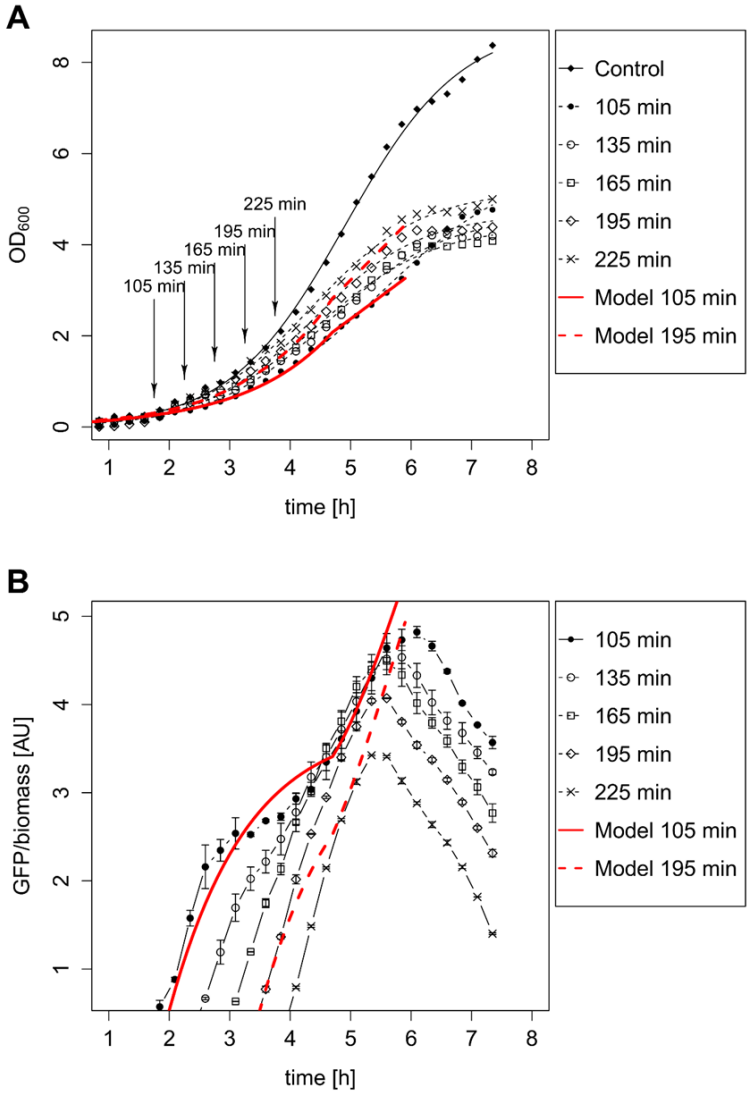
AEM

Applied and Environmental Microbiology

662 C) while the nucleoid was segregated near the new pole. This process led to bacterial chains, with
663 terminal cells containing high amounts XylR-plasmid complexes (arrows in row 3, column C).

664

665



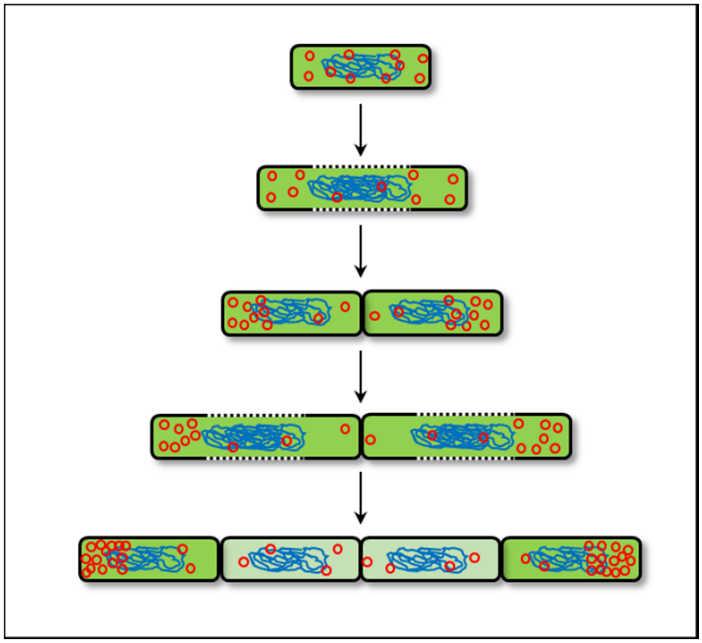
666

Accepted Manuscript Posted Online

667 **Fig. 8** GFP production of *B. megaterium* carrying pSSBm85 in response to the time of induction
668 and growth phase. A) Growth curves at various times of induction, and control without inducer.
669 B) GFP abundance normalized to biomass for various induction times. The red curves show the
670 modeled growth (A) and production behavior (B) for the induction times, 105 min (solid line)
671 and 195 min (dashed line).

672

673



674

675 **Fig. 9** Proposed model of constrained plasmid distribution that explains the population
676 heterogeneity during recombinant protein production. Polar fixation of plasmids (red circles)
677 after cell division led to cell lines with high and low production performance. In this schema,
678 plasmids were fixed to the old cell pole. However, polar fixation can affect either the old or the
679 new cell pole.

Applied and Environmental Microbiology
AEM
Applied and Environmental Microbiology