A model for the time dependence of expression during exponential growth

In exponentially growing cells the steady-state concentration of β-galactosidase in the cells is determined by the synthesis rate and the dilution rate over newly synthesized cells, or the growth rate of the cells. In each cell the synthesis rate is constant and, therefore, the rate of synthesis in the culture will be proportional to the number of cells, which increases exponentially, i.e.

\[
\frac{d[LacZ]}{dt} = p \cdot N_t \cdot e^{\mu t}
\]

where \( p \) is the synthesis rate per cell, \( N_t \) the cell density at the beginning of the exponential growth phase, \( t \) is the time and \( \mu \) the growth rate constant. In this approach, we have ignored any breakdown of β-galactosidase, which is probably slow anyway. The time dependence of the β-galactosidase concentration in the culture follows from integration of equation 1. In batch culture, it is unlikely that the level of expression at the beginning of growth is the same as the steady-state level (see for example ref. 7) as the cultures are usually inoculated with cells from the stationary growth phase (‘overnights’) or with ‘uninduced’ cells. In the latter case, the β-galactosidase concentration in the inoculate is zero and integration yields

\[
[LacZ] = \frac{p \cdot N_0}{\mu} (e^{\mu t} - 1)
\]

Normalizing to the amount of cells (\( N \)) present at any time point during growth results in the expression for the cell-specific β-galactosidase concentration

\[
\frac{[LacZ]}{N} = \frac{p}{\mu} \left( e^{(\mu - \frac{1}{\mu})} - 1 \right)
\]

The time \( t_{lag} \) was introduced to account for the lag time that precedes exponential growth in a real experiment. The initial condition of zero β-galactosidase ignores the synthesis of any β-galactosidase during the lag time.

Equation 3 shows that the β-galactosidase concentration in the cell reaches a steady state concentration \( p/\mu \) after enough time has elapsed to make \( e^{(\mu - \frac{1}{\mu})} \gg 1 \). Importantly, the steady state concentration depends both on the synthesis rate \( p \) and the growth rate \( \mu \). The time dependence of the process by which the synthesis rate and the dilution rate over the new cells are balanced (the pre-steady state) only depends on the growth rate.

Pre-steady-state expression of β-galactosidase

The Mg\(^{2+}\)-citrate transporter CitM is the principal citrate transporter of Bacillus subtilis. Expression of CitM is induced by citrate and repressed by glucose, and other sugars and non-sugars present in the medium (9). Repression is mediated by carbon catabolite...
The deficiency results in relief of ccr-mediated repression. CSE minimal medium containing citrate was inoculated with a pre-culture of the strains grown in the absence of citrate and, therefore, completely devoid of any β-galactosidase activity (Figure 1a, left panel).

In both strains, the β-galactosidase activity expressed per cell density increased gradually during the exponential growth phase and reached a maximum when entering the stationary phase. Clearly, inferring promoter activity from the level of expression would significantly depend on the time point of sampling. Fitting of the data from the exponential growth phase to equation 3 gave satisfactory results (regression coefficients of $R = 0.99$ and $R = 0.98$ for strains CM002 and CM010, respectively), indicating that the data is well described by the simple model presented above (Figure 1a, right panel). It follows that the β-galactosidase synthesis rates in the two strains were constant during the exponential growth phase. The increase in cell-specific β-galactosidase merely reflects the pre-steady-state period of the expression in the cells and is an intrinsic property of the system. The analysis allows for a reliable estimate of the difference in promoter activity in the two strains that is independent of the time point at which the expression was measured. The fit revealed that the β-galactosidase synthesis rate $p$ was 3.6 times higher in strain CM010 than in strain CM002. Assuming that the difference is completely caused by the relief of ccr in the CcpA-deficient CM010 strain, this would correspond to a repression of 72% in the wild-type strain during exponential growth on CSE medium.

**β-Galactosidase expression at different growth rates**

Growth of *B. subtilis* strain CM010 in LB medium is three times faster than growth in minimal CSE medium. The steady-state level of cell-specific β-galactosidase activity in the exponential growth phase (between ~1 and 3 h) was accordingly reached faster, in agreement with equation 3 (Figure 1b, left panel). Again, a reasonable fit of the data to equation 3 was obtained ($R = 0.97$) indicating a constant rate of β-galactosidase synthesis during exponential growth. The steady-state level of expression of β-galactosidase $p/\mu$ was fitted to be 10 Miller U, which was five times lower than observed during growth on minimal CSE medium. Importantly, the β-galactosidase synthesis rates, $p$, differed only by a factor of 1.5. Following this analysis, one might argue that a constant level of expression observed at different growth rates is the result of strict regulation of the rate of transcription by the growth rate; the level of expression is independent of the growth rate when $p$ is

---

**Fig. 1.** (a, b). Left panels. Growth curves of *B. subtilis* strains CM002 (wild type, black circles) and CM010 (∆CcpA, white circles) in CSE medium (a) and LB medium (b) containing 10 mM citrate. The bars indicate the β-galactosidase activity of the cells at the indicated time points (in Miller U). Right panels. Best fit of the β-galactosidase activities from the exponential part of the growth curves to equation 3 (Figure 1b, left panel). It follows that the β-galactosidase synthesis rate $p$ was 3.6 times higher in strain CM010 than in strain CM002. Assuming that the difference is completely caused by the relief of ccr in the CcpA-deficient CM010 strain, this would correspond to a repression of 72% in the wild-type strain during exponential growth on CSE medium.

---

repression (ccr). The *B. subtilis* strains CM002 and CM010 contain a transcriptional fusion of the citM promoter region and the lacZ reporter gene. Strain CM002 is the wild-type background, while CM010 is deficient in CcpA, a central component of ccr in *B. subtilis*...
proportional to $\mu$. In contrast, an inverse relation between the level of expression and growth rate, as observed here, would be indicative of a growth-rate-independent rate of transcription (10). Such considerations may be helpful when discussing growth-rate-dependent expression (2, 3, 4, 5, 8) and, at least, show the importance of careful use of terminology. At any time, it should be kept in mind that the rate of transcription is a complex parameter containing many different contributions.

**$\beta$-Galactosidase expression and growth phase**

The dominant effect of growth on the evaluation of lacZ-promoter fusions is nicely demonstrated when mRNA levels and protein levels are compared in CM010 cells from the mid-exponential and the early-stationary growth phase in LB medium. In the stationary growth phase, the level of mRNA in the cells as determined by RT-PCR was significantly lower than observed in the exponential growth phase (Figure 1c, top left). This situation may arise from a lower rate of transcription, increased messenger instability or a combination of these. Many effects may cause the lower rate of transcription, i.e. a change in the energetic state of the cells, inactivation of the transcription machinery, reduced induction (e.g. depletion of the inducer during growth), or downregulation of transcription. Surprisingly, the cell-specific $\beta$-galactosidase activity was a factor of 3.4 higher in the cells from the stationary phase than in the cells from the exponential growth phase (Figure 1c, bottom left). Measurement of the Mg$^{2+}$-citrate transporter uptake activity in the same cells revealed an increase of a factor of six, indicating that the increase is not an artefact of the lacZ reporter gene fused behind the citM promoter region (Fig. 1c, right panel). The paradoxical relationship between the decreased mRNA level and increased protein levels should be analysed in the context of the almost complete lack of growth of the cells in stationary phase. Then, the decreased rate of protein synthesis from the lower amount of mRNA may still exceed the very low, if any, rate of dilution caused by cell division, resulting in a net increase of protein content per cell. While the rate of transcription goes down, the level of expression goes up. Importantly, changing levels of expression when cells shift from the exponential to the stationary growth phase as observed here do not necessarily indicate (a change in) regulation of expression by trans-acting elements.

**Conclusion**

In growing cells, the rate of transcription is not the same as the level of expression. The latter is determined by both the rate of transcription and the growth rate. The model presented in equation 3 (above) gives a quantitative account of the relation between rate of expression end level of expression and indicates that the pre-steady state of expression is solely dependent on the growth rate. Two questions should be kept in mind when analysing lacZ-promoter fusions by measuring the specific $\beta$-galactosidase activity of the cells. One: has the expression level reached the steady-state value? Occasionally, the exponential growth phase in batch culture may be too short to reach the steady state. Then, the steady-state level follows from extrapolation of the levels of expression during the exponential growth phase using equation 3. Two: when comparing $\beta$-galactosidase activities, are the growth rates in the two situations the same? If not, expression levels should be corrected for the growth rates to make a reliable comparison between the transcription rates. A reliable evaluation of lacZ-promoter fusions requires measurement of the time course of $\beta$-galactosidase activity during the exponential growth phase and the growth rate.

Jessica B. Warner and Juke S. Lolkema

1 Molecular Microbiology, Groningen Biomolecular Sciences and Biotechnology Institute, University of Groningen, 9751 NN Haren, The Netherlands.

Author for correspondence: Juke S. Lolkema. Tel.: +31 50 3632154. Fax: +31 50 3632154. e-mail: j.s.lolkema@biol.rug.nl


The genomes of *Pseudomonas* encode a third HU protein

HU proteins are ubiquitous DNA-binding factors that, along with other so-called nucleoid-associated proteins, are involved in the structural maintenance of the bacterial chromosome and other events that require DNA bending (1). In contrast to the structurally related integration host factor (IHF) protein, HU proteins bind DNA in a sequence-independent manner, although both produce changes in DNA structure (8, 13). In *Escherichia coli*, HU is one of the most abundant of this kind of protein, with a postulated binding site in the chromosome of 200 bp. In *Pseudomonas putida*, HU has been implicated in the transcriptional activation of the Pu and Ps promoters of the tolue degradative plasmid TOL (12). Previous work in our laboratory led to the purification and characterization of two HU-like factors named HupB and HupN (2). They were first identified biochemically on the basis of sequence-independent DNA binding and DNA-bending activity. While independent knockouts of each yield no apparent phenotype in *vivo*, simultaneous mutation of the two is lethal to the cell. This observation implies a redundancy in function, as is the case with the HupA and HupB proteins (9). In principle, these results ruled out the existence of any other HU-like protein, at least on the basis of functional screenings.

These assumptions were challenged, however, during the course of a genome-wide search for nucleoid-associated proteins of *P. putida* KT2440 (http://www.tigr.org) using the Genewise 2.0 program. This software can scan suboptimal DNA sequences using a Hidden Markov Model like that stored in Pdam database (3). This procedure is very sensitive, as it can properly handle small errors in DNA sequence (such as frameshifts due to artefactual insertions or deletions). In addition, the same software accurately finds weak similarities. This is because the program extracts all the information contained in an alignment of known members of the family, instead of using a single protein or a consensus for the search (7). Our choice was the use of the PF0012 model from the Pdam database (3), which was contrived on the alignment of all known IHF and HU-like proteins. By this

Microbiology 148, May 2002
Two of these ORFs encode the two subunits and characterized by experimental methods. Five potential HU proteins were found. Four had already been described (a) Alignment of the HupP proteins of Pseudomonas. CLUSTAL W (16) was employed to align the sequences found in the genome-wide searches described in the text. Note the nearly complete conservation of the sequence. (b) Phylogenetic tree of HU-like proteins of the Proteobacteria, the chromosomal sequences of which are known. The HU-like proteins were also aligned with CLUSTAL W. The evolutionary distances were calculated with the Protdist program. The tree shown was generated with the neighbour-joining algorithm of the PHYLIP 3.5 package. Note that HupN and HupP clearly cluster away from the archetypal HU-α and HU-β subfamilies.

Fig. 1. (a) Alignment of the HupP proteins of Pseudomonas. CLUSTAL W (16) was employed to align the sequences found in the genome-wide searches described in the text. Note the nearly complete conservation of the sequence. (b) Phylogenetic tree of HU-like proteins of the Proteobacteria, the chromosomal sequences of which are known. The HU-like proteins were also aligned with CLUSTAL W. The evolutionary distances were calculated with the Protdist program. The tree shown was generated with the neighbour-joining algorithm of the PHYLIP 3.5 package. Note that HupN and HupP clearly cluster away from the archetypal HU-α and HU-β subfamilies.

procedure, five potential HU/IHF proteins were found. Four had already been described and characterized by experimental methods. Two of these ORFs encode the two subunits of the IHF protein of P. putida (5). Two others match the HupB and HupN protein sequences (2). But, in addition, a fifth gene was found that clearly encoded an extra HU-like protein. This so-far unknown protein was used as a query with the BLAST software (provided at the NCBI site http://www.ncbi.nlm.nih.gov/Microbes_blast/unfinishedgenome.html) to search for similar proteins in other complete or incomplete bacterial genomes. Interestingly, the only close relatives were found in Pseudomonas aeruginosa, Pseudomonas fluorescens and Pseudomonas syringae (Fig. 1a). Although some of the accessible versions of P. putida, P. syringae and P. fluorescens genomes are still incomplete, the degree of conservation of the novel HU-like gene in the four Pseudomonas species is certainly very high. The similarities among correspondents in these species, and the presence of an orthologue in the P. aeruginosa chromosome (an entirely annotated genome sequence; 15), indicated that such a sequence determines a true protein. The corresponding gene was thereby named hupP, since it appeared to be habitually present in several Pseudomonas species. Further, we argue that hupP is an authentic gene encoding a genuine and thus-far unidentified HU-like protein.

The capacities of hupP are not trivial to predict. The failure of in vivo and in vitro screenings to detect either the function or the sequence of the hupP gene of P. putida (2) and P. aeruginosa (6) is intriguing. In the case of P. putida, such a screening was based on the enrichment of DNA-binding proteins from a cell lysate that stimulated a β-recombination reaction dependent on non-specific DNA bending. This procedure exclusively yielded the HupN and HupP products. Moreover, the lack of viability of double hupB/hupN mutants of P. putida suggests that their redundant functions in multiple cellular events cannot be replaced by hupP. The possibility that hupP is a pseudogene, devoid of coding capabilities, is unlikely. Unlike typical pseudo genes, both the hupP sequence and its surrounding chromosomal context are conserved in the Pseudomonas species examined. In addition, each of the four hupP genes are flanked by a sigma-70 dependent promoter and a strong Rho-independent terminator. This is predicted by the NNPP (http://www.fruitfly.org/seq_tools/promoter.html) and Terminator (4) programs. Such a resemblance is a feature rarely present in pseudogenes and argues that the corresponding sequence can be perfectly expressed. It thus appears that the hupP, hupN and hupP genes of Pseudomonas encode unequivocally distinct HU variants. One possibility is that HupB provides housekeeping HU activity (i.e., facilitating non-specific DNA bending) while HupN and/or HupP have more specialized functions as the result of a recent horizontal gene-transfer event.

Such an uneven number of HU types in the same cells is somewhat surprising. When 31 completed and published eubacterial genomes were scanned for the presence of HU-like proteins using similar procedures, most had only one (16 species) or two (6 species) HU proteins. When all the HU-like proteins from the group Proteobacteria that were found in the analysis of completed genomes were compared (Fig. 1b), the HupN and HupP proteins from the genus Pseudomonas were present in independent clusters. The hupP gene was previously annotated in the P. aeruginosa genome as hupA, a name that could be misleading, since the hupP genes of Pseudomonas clearly branch out from the HU-α group (Fig. 1b). In contrast, the HupB protein is included in a more populated group that comprises, for instance, the reference HU-1 specimen of E. coli. This classification is compatible with those described previously (10, 11) and can shed some light on the possible origin of the three HU-like protein present in Pseudomonas. The tree of Fig. 1(b) hints at HU-like proteins originating in a common ancestor (10, 11) which may have been maintained as such in some organisms of the α and ε subdivision of the proteobacteria (i.e. Rhizobiaceae, Caulobacter, Helicobacter). The ancestor of that kind could have then diverged in the γ group to give rise to the HU-α (i.e. Enterobacteriaceae, Vibrionaceae) and HU-β (i.e. Pseudomonadaceae, Xanthomonadales) branches. An enticing possibility is that, later in evolution, the enterobacteria could have horizontally acquired an extra copy of the HU-β type, thus giving rise to the heterodimeric HU that is emblematic of this family (10). In this context, it would be reasonable that HupP appeared in Pseudomonas through the duplication of HupB (after
the proposed gene-transfer event to the enterobacteria), and rapidly evolved for attaining a specialized function. The origin of HupN is more ambiguous, since it pertains to a wider class of less delineated HU species (Fig. 1b). HupN could have been procured from other soil microorganisms (i.e. Rhizobiaceae) and maintained as a backup for the equivalent functions of HupB (2).

In conclusion, three HU-like proteins were found to be encoded in the genomes of four Pseudomonas species. Two, HupN and HupP, clearly diverge from the other known members of the HU protein family, and are so far restricted to Pseudomonas. The appearance of three different HU-like proteins in a bacterial genome is not yet reported for other genera and could reflect a degree of regulatory complexity related to the adaptability of this genus to a wide range of environmental conditions (14). Although related to the other HU factors in its sequence and predicted structure, HupP may form a class of its own within the larger protein family.

Acknowledgements

Genomic sequences of Pseudomonas putida strain KT2440 were kindly provided by the Institute for Genomic Research (http://www.tigr.org), whose sequencing effort is supported by grants of the Department of Energy (DOE) and the German Ministry of Education and Research (BMBF).

Ildefonso Cases† and Víctor de Lorenzo
Centro Nacional de Biotecnología CSIC, Campus de Cantoblanco 28049 Madrid, Spain.
†Present address: European Bioinformatics Institute, Cambridge CB10 1SD, UK. Author for correspondence: V. de Lorenzo.
e-mail: vdlorenzo@cnb.uam.es