Fluorescent Reporter Genes and the Analysis of Bacterial Regulatory Networks

Hidde de Jong[1(✉)] and Johannes Geiselmann[1,2]

[1] INRIA, Grenoble - Rhône-Alpes Research Centre, 655 avenue de l'Europe, Montbonnot, 38334 Saint-Ismier Cedex, France
Hidde.de-Jong@inria.fr

[2] Laboratoire Interdisciplinaire de Physique (LIPhy, CNRS UMR 5588), Université Grenoble Alpes, 140 Avenue de la Physique - BP 87, 38402 Saint Martin d'Hères, France
Hans.Geiselmann@ujf-grenoble.fr

Abstract. The understanding of the regulatory networks controlling the adaptation of bacteria to changes in their environment is critically dependent on the ability to monitor the dynamics of gene expression. Here, we review the use of fluorescent reporter genes for dynamically quantifying promoter activity and other quantities characterizing gene expression. We discuss critical physical and biological parameters in the design, development, and use of fluorescent reporter strains. Moreover, we review measurement models that have been proposed to interpret primary fluorescence data and inference methods for estimating gene expression profiles from these data. As an illustration of the use of fluorescent reporter strains for analyzing bacterial regulatory networks, we consider two applications in the model bacterium *Escherichia coli* in some detail: the joint control of gene expression by global physiological effects and specific regulatory interactions, and the importance of protein stability for the inference and analysis of transcriptional regulatory networks. We conclude by discussing some current trends in the use of fluorescent reporter genes.

Keywords: Fluorescent reporter genes · Bacterial regulatory networks · Growth · Gene expression · Bioinformatics · Systems biology

1 Introduction

Bacterial cells are capable of surviving in an enormous variety of conditions by adapting their functioning to changes in the environment. These adaptive capabilities emerge from complex regulatory networks that control and coordinate the different functions of the cell, including transport of substrates into the cell, the metabolism of these substrates to produce energy and molecular building blocks for growth, gene expression to convert the molecular building blocks into proteins, and the replication of DNA [40]. On the molecular level, these regulatory networks involve interactions between, for example, proteins and DNA

O. Maler et al. (Eds.): HSB 2013 and 2014, LNBI 7699, pp. 27–50, 2015.
DOI: 10.1007/978-3-319-27656-4_2

(transcriptional regulation) and between metabolites and proteins (transcription factor activity and enzyme activity). While for some model organisms the regulatory networks have been mapped in quite some detail, in most cases the structure of interactions is largely unknown. Moreover, even when the structure is known, the precise role of these interactions in bringing about a specific physiological response to a change in the environment remains a very difficult question for which only partial questions are available [2]. In order to reconstruct regulatory networks from experimental data, and understand the role they play in controlling the responses of bacteria to changes in their environment, we need to be able to dynamically quantify the metabolites, proteins, and mRNAs involved in cellular processes.

Since most of the mass contents of a cell are proteins and RNA, it is particularly important to measure gene expression over time. A large variety of technologies have been developed to this purpose, with particularly striking advances over the past decade. For example, DNA microarrays, RNA sequencing, and RT-qPCT technologies allow the quantification of cellular transcripts, often on a genome-wide scale [13,32]. Quantitative proteomics makes it possible to measure the protein contents of a cell [11], while ribosome profiling provides an estimate of the translation rates [20]. These technologies have provided a wealth of information on gene expression changes in bacteria and the regulatory networks that control these changes. However, for the purpose of reconstructing regulatory interactions from the data or observing precise temporal phenomena, they have limitations as well. For example, high-resolution time-series data are currently difficult to obtain, for both financial and technological reasons. Moreover, the above technologies involve many steps for extracting the cellular content and analyzing its constituents, thus potentially introducing biases. Furthermore, due to the requirement of having to remove a sample for analysis, these techniques do not allow monitoring the time course of gene expression in a single bacterial culture or in individual bacterial cells.

Fluorescent reporter genes provide an alternative and a complement to the above technologies, making it possible to monitor gene expression in real time and *in vivo*, both in single cells and on the population level. Reporter genes are classical tools in molecular biology, especially fusions of promoters and other regulatory regions with the *lacZ* gene. This gene encodes the protein β-galactosidase, the activity of which can be quantified by colorometric assays [42]. The breakthrough of fluorescent (or luminescent) reporter genes lies in combining the same principle of fusing a genomic region to a reporter gene with a non-intrusive technology for protein quantification [10,17,47]. When excited at a specific wavelength, the fluorescent protein encoded by the reporter gene produces light with an emission peak shifted toward the red with respect to the excitation wavelength. The emission peak can be easily captured and its intensity is proportional to the amount of fluorescent protein in the sample. The use of fluorescent reporter genes as tools in biology has enormously benefited from the engineering of new fluorescent proteins with desired spectral and other characteristics, molecular cloning techniques for constructing reporter systems inside

the cell, and the miniaturization and automation of culturing bacterial cells and measuring emitted fluorescence. On the population level, microplate readers allow monitoring of the expression of fluorescent reporter genes in about one hundred bacterial cultures in parallel, whereas on the single-cell level advances in microfluidics and fluorescent microscopy make it possible to follow gene expression in single cells over many generations [29,49].

Reporter gene experiments generate huge amounts of data, the analysis of which poses subtle problems that are often not straightforward to solve. The major difficulty in the analysis of the results of reporter gene experiments, usually consisting of emitted fluorescent levels and the optical density or absorbance of a growing culture (in the case of population-level measurements), lies in the indirect relation between the primary data and the biological quantities of interest. Making sense of the primary data requires mathematical models of reporter gene expression and statistical techniques to estimate the quantities in the models from the data [1,3,12,15,26,28,37,38,48,51]. Moreover, computer tools for the efficient and user-friendly application of these methods are needed, in order to make the analysis methods available to biologists [1,7,51].

The aim of this review is to give a short overview of the process of obtaining reliable data about gene expression by means of fluorescent reporter genes and to illustrate the use of the resulting data with some representative examples from the literature, in particular work carried out by the authors over the past few years. We will first discuss the physical and biological principles underlying fluorescent reporter genes and outline which experimental set-ups are typically used to follow bacterial gene expression over time. Our focus will be on population-level measurements of gene expression and we refer to other reviews for specific issues involved in single-cell experiments [36,49]. Second, we present different aspects of the analysis of the primary data thus obtained, in particular the measurement models used to interpret the primary data, the statistical methods to infer quantities in the model from the data, and computer tools to apply these methods in a rigorous and automated way. Third, we discuss how the use of fluorescent reporter genes has shed new light on a number of well-known model systems, such as the regulatory circuits involved in growth transitions in *E. coli* and chemotaxis and motility in the same organism.

2 Fluorescent Reporter Gene Experiments in Bacteria

2.1 Reporter Genes

Reporter genes are a very convenient and widely-used tool for measuring the dynamics of gene expression. The principle consists in duplicating the elements that control the expression of the gene of interest upstream of a readily detectable gene, the reporter gene. Depending on the control elements studied, different portions of the original gene are cloned upstream of the reporter gene [44]. For example, if we want to measure the combined effect of transcriptional and translational regulation, we would "fuse" the reporter protein to the gene of interest, creating a chimeric protein. This type of construction is called a translational

fusion. Even though the basic principles are identical for all types of reporter gene constructions, we will focus here on the first, and most highly regulated step of gene expression: the initiation of transcription [8,40]. In order to study regulation of this process, we create "transcriptional fusions".

The fundamental principle is illustrated in Fig. 1. The elements that control transcription initiation are located in the promoter of the gene (red in Fig. 1). The time-dependent activity of the promoter determines the rate of production of the corresponding messenger RNA (mRNA), which is subsequently translated into protein. However, in almost all cases, the gene of interest codes for a protein that is not easily detectable. We therefore introduce into the cell a second copy of the promoter, cloned upstream of a gene coding for an easily detectable protein. This second copy can either be placed on the chromosome or carried on a plasmid [41]. The advantage of integrating the reporter construct into the chromosome is to ensure that in all conditions the number of copies of the reporter gene are well controlled. This criterion is important, for example, when using reporter genes for assessing the noise of gene expression in cells. However, extra-chromosomal copies carried on plasmids are easier to construct and give higher signal intensities due to the multiple copies of the plasmids. For the promoter of *E. coli*, a complete library fusing all intergenic regions to a *gfp* reporter gene have been constructed [50]. The copy-number of these plasmids remains constant over different growth conditions [6], thus avoiding the major potential artifact of the use of reporter plasmids.

2.2 Different Types of Reporter Genes

The most commonly used reporter proteins are fluorescent proteins [17,44]. These proteins contain a fluorophore that emits visible light shifted towards longer wavelengths when irradiated with light corresponding to excitation wavelengths of the molecule. The original green fluorescent protein (GFP) had an excitation maximum around 480 nm and an emission peak around 510 nm [47]. Today, variants of the protein are available that emit at virtually all visible wavelengths [10]. Since very few cellular components emit fluorescence in the visible part of the spectrum, the concentration of these proteins can be quantified within the living cell.

However, the formation of the fluorophore can be a relatively slow enzymatic reaction and the final step of the reaction sequence involves molecular oxygen [47]. These fluorescent reporters will therefore not function in an anaerobic environment. The engineered fluorescent proteins used today have maturation times on the order of 15 min, short compared to the dynamics of gene expression with characteristic times on the order of an hour. However, the bias of maturation rates should be taken into account and measured for each fluorescent protein and strain since this parameter can vary depending on experimental conditions [18]. When the maturation of the active fluorophore is incorporated into the measurement model, we can deduce the production rate of GFP from the experimentally observed fluorescence (see Sect. 3.1). Additional care must be taken when creating translational fusions since the attachment of the fluorescent protein to the

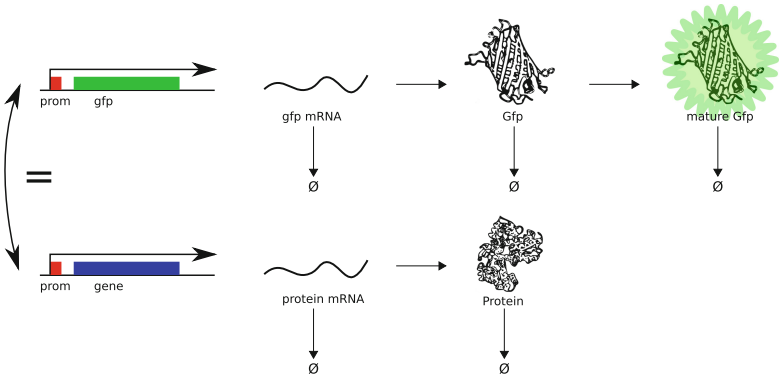

Fig. 1. Reporter gene measurements. The target gene (blue) is transcribed from the promoter (red, prom) to give the corresponding messenger RNA (mRNA). The mRNA is translated into the corresponding protein. However, the target protein is generally not easily detectable. The promoter of the target gene is therefore cloned upstream of the gene coding for the green fluorescent protein (green, GFP). Since the promoter region contains all of the control elements of transcription initiation, the rate of production of the *gfp* mRNA is in principle the same as the rate of production of the mRNA of the target gene. The *gfp* mRNA is translated into GFP, which is subsequently converted to the mature, fluorescent form by an autocatalytic process. All constituents of the process are degraded by growth dilution or natural degradation (Color figure online).

native protein can perturb the folding of GFP. Specific variants of GFP greatly alleviate this problem [46]. Even though the fluorescence of GFP allows the specific detection of this protein in a living cell, other cellular components, such as flavins, also emit fluorescence at similar wavelengths [28]. This autofluorescence has to be incorporated into the measurement model (Sect. 3.1).

Another type of reporter that gives a unique signal in living cells is luciferase. Most bacteria do not spontaneously emit light and bioluminescence is therefore an ultra-sensitive method for detecting the reporter protein in a bacterial population. The commonly used firefly luciferase is a monomeric protein that emits light when oxidizing the substrate, luciferin [39]. Bacterial luciferase is a heterodimer that emits light when oxidizing the substrate, a long-chain aldehyde [35]. Both of these enzymes can therefore not function in an anaerobic environment. Furthermore, the substrate of the bioluminescence reaction has to be continuously provided. This is difficult in the case of luciferin, but can be easily accomplished in the case of bacterial luciferase. We simply clone the entire luciferase operon, comprising genes *luxCDABE*, downstream of the promoter of interest. The genes *luxAB* code for the luciferase, whereas the genes *luxCDE* reduce cellular fatty acids to the aldehyde substrate of luciferase [35]. Reporter plasmids based on these constructs are readily available [33,34]. Light emission is therefore continuous. However, light emission depends on the concentration of the substrate and the concentration of luciferase. This dependency can be

calibrated to derive a reliable measure of the concentration of luciferase from the measured bioluminescence intensity [33]. However, for both types of luciferases, the reaction depends on ATP and changes in the bioluminescence signal may reflect changes in intracellular metabolites. In the case of bacterial luciferase, the generation of the substrate is furthermore sensitive to the redox potential of the cell [25]. Even though bioluminescence detection is extremely sensitive, for all these reasons, fluorescent proteins are generally preferred as reporter genes.

2.3 Experimental Setup

A typical experiment of measuring the dynamics of gene expression therefore consists in constructing the appropriate plasmid-based or chromosomally integrated reporter construct (Sect. 2.1) and quantifying the fluorescence per cell as a function of time. The measurements can target single cells under the microscope [36,49] or an entire bacterial population. In single-cell measurements the total fluorescence of the cell is divided by the cell volume in order to obtain the fluorescence intensity per unit cell volume. Here, we will focus on population measurements, where the total cell volume is approximated by the absorbance of the culture. The task therefore consists in measuring at regular time intervals the absorbance and the total fluorescence of the bacterial culture. The most convenient setup for this task consists of an automated, thermostated plate-reader that keeps the samples at a constant growth temperature and agitates the microplate in order to prevent sedimentation of the bacteria and provide sufficient mixing of the culture (Fig. 2). Typically, a 96-well microplate is agitated for about one minute, the absorbance signal of all wells is read, agitation is resumed and the fluorescence is measured. The entire growth and measurement cycle takes about 5 min, which is a sufficient sampling rate since changes in gene expression take place on a time-scale of several tens of minutes or even hours.

The goal of such experiments is to monitor the dynamics of gene expression. This implies that the environmental conditions change during the experiment. This is achieved by carefully selecting the initial conditions. For example, the bacteria could be grown overnight in an Erlenmeyer flask until they have exhausted the available nutrient source (Fig. 2). The experiment is started by diluting this pre-culture into fresh growth medium. The changes in gene expression then reflect the dynamics of adaptation of the bacteria to the new environment. In the example of Fig. 2, after adapting to new environmental conditions the bacteria reach a steady state where gene expression remains constant. In this example, the new growth medium contains two nutrient sources. When the preferred nutrient source is again exhausted, the bacteria continue growth, at a lower rate, on the less preferred nutrient and the expression of the target gene dynamically adapts to the new environment. A typical experiment of this type lasts for about 12 h. In principle, 96 different reporter gene constructs can be measured in parallel in a 96-well microplate. However, in order to analyze the data, we need to include wells of bacteria that do not express the fluorescent protein (for fluorescence background correction) and wells that do not contain bacteria in order to correct the measured absorbance for the absorbance of the

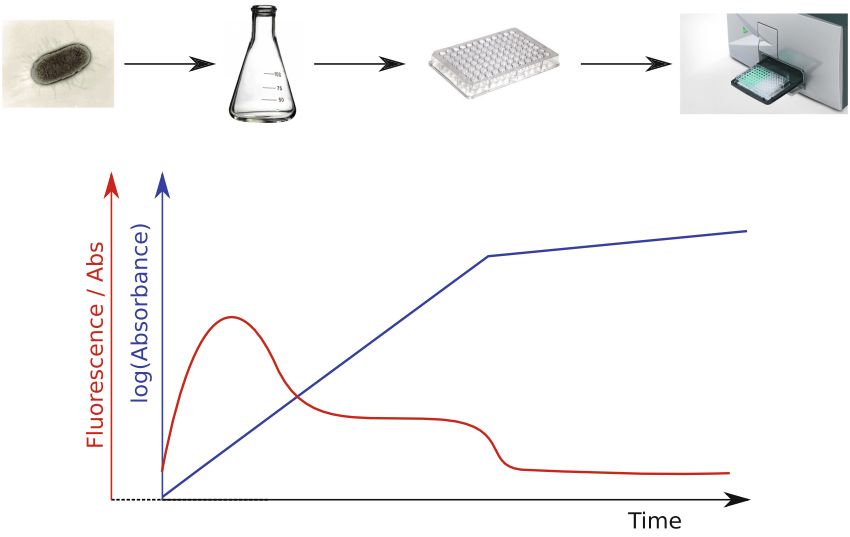

Fig. 2. A typical experiment for measuring the dynamics of gene expression using reporter genes and an automated microplate reader. The top row shows the main steps of the experiment. Bacteria are generally grown in a culture tube or an Erlenmeyer flask to a defined starting state. Aliquots of different pre-cultures are dispatched into the wells of a microtiter plate. The microplate is then read in an automated plate-reader that measures the absorbance and fluorescence signals of all the wells at regular intervals (typically five minutes). An example curve is shown in the bottom part of the figure. Here, the bacterial culture grows exponentially until a preferred nutrient source is exhausted. The bacteria continue to grow at a lower rate on a less favorable nutrient source. The absorbance measurements (blue) are on a logarithmic scale. The fluorescence measurements are shown in red on a linear scale. For clarity, we show the relative fluorescence, *i.e.*, the absolute fluorescence measured by the plate-reader divided by the absorbance (Color figure online).

growth medium (Sect. 3). It is even possible to incorporate two or three different reporter gene constructs, using fluorescent proteins of different color, into the same bacterial cell. In this case, each measurement cycle comprises fluorescence measurements at different wavelengths. An experiment thus generates a large amount of data that have be analyzed in an automated manner (Sect. 3.3).

3 Analysis of Fluorescent Reporter Gene Data

3.1 Mathematical Models of Reporter Gene Expression

The interpretation of reporter gene data is usually based on mathematical models describing the dynamics of the expression of reporter genes. The most commonly found models in the literature distinguish between the stages of transcription, translation, and maturation, giving rise to a three-variable system of ordinary

differential equations (ODEs). Here, we follow with some modifications the model of [12]. The variables are $m(t)$, $r_u(t)$, and $r(t)$, representing the time-varying concentrations of reporter mRNA and immature and mature reporter protein, respectively (μM), with time variable t (min).

The accumulation of reporter mRNA involves a balance between synthesis and decay. mRNA synthesis is defined as the product of a maximum synthesis constant k_m (μM min^{-1}) and a dimensionless promoter activity $a(t)$, scaled between 0 and 1. Notice that the promoter activity is a time-dependent input variable and that the regulatory interactions shaping the temporal variation of the promoter activity are not explicitly modeled. In fact, as will be seen below, the very objective of reporter gene experiments is usually to infer $a(t)$ from the data. mRNA decay involves a first-order term representing physical degradation, with a rate constant d_m (min^{-1}) and growth dilution, determined by the time-varying growth rate $\mu(t)$ (min^{-1}). In practice, the growth dilution term is often omitted since mRNA decay is dominated by physical degradation: mRNA in bacteria is generally unstable, with half-lives on the order of a few minutes [5], much shorter than typical cell doubling times. The above considerations give rise to the following equation describing the transcription of a gene:

$$\frac{d}{dt}m(t) = k_m\, a(t) - (d_m + \mu(t))\, m(t). \tag{1}$$

The dynamics of the concentration of unfolded (immature) protein is also defined as a balance between protein synthesis and decay. The synthesis rate is given by a first-order term, the multiplication of the reporter mRNA concentration $m(t)$ with a protein synthesis constant k_u (min^{-1}), which can be interpreted as the translational activity (quantity of protein produced per quantity of mRNA per min). Here we will assume that k_u is constant [24]. Note that this assumption implies that there is no (active) regulation of translation. The decay of immature protein also involves physical degradation and growth dilution, like for mRNA, with the degradation constant d_r (min^{-1}). The half-live of a protein is defined in terms of the latter constant as $\ln 2/d_r$. In addition, the maturation of the unfolded protein leads to another first-order decay term, involving the maturation constant k_r (min^{-1}). Maturation times reported in the literature are usually given as the half-time of maturation, $\ln 2/k_r$. The translation equation thus becomes

$$\frac{d}{dt}r_u(t) = k_u\, m(t) - (d_r + k_r + \mu(t))\, r_u(t). \tag{2}$$

Mature protein accumulates through maturation and decays by physical degradation and growth dilution, where we assume that the mature and immature proteins have the same half-lives and thus degradation constants:

$$\frac{d}{dt}r(t) = k_r\, r_u(t) - (d_r + \mu(t))\, r(t). \tag{3}$$

For a more detailed discussion of the assumptions underlying the above models, see [12]. Notice that the total concentration of reporter protein, $r_{tot}(t)$,

can be obtained by summing Eqs. 2 and 3, giving rise to $d/dt\, r_{tot}(t) = k_u\, m(t) - (d_r + \mu(t))\, r_{tot}(t)$.

In practice, it is often possible to reduce the above models, in particular when the maturation and mRNA degradation rates are quite fast with respect to the growth rate or the protein degradation rate. This allows a quasi-steady state assumption (QSSA) to be applied, implying $d/dt\, m(t) = d/dt\, r_u(t) = 0$, and reducing the three-variable model to the following ODE:

$$\frac{d}{dt} r(t) = k'_m\, a(t) - (d_r + \mu(t))\, r(t), \tag{4}$$

where k'_m (min^{-1}) is a lumped protein synthesis constant, defined as

$$k'_m = \frac{k_m\, k_u}{d_m}, \tag{5}$$

and where $k'_m\, a(t)$ defines the protein synthesis rate. Remark that in this case, the protein synthesis rate is directly proportional to the promoter activity. If the reporter protein is very stable and the bacterial cultures are growing at a non-negligible rate, we have $d_r \ll \mu(t)$ and the model can be further simplified by ignoring the term describing the physical degradation of the protein. As will be seen below, the empirical expression for promoter activity often found in the literature corresponds to the latter case.

3.2 Estimation of Gene Expression Profiles from Primary Data

Which quantities of biological interest can be derived from the optical density/absorbance and fluorescence data by means of the models introduced above? By strain construction, as explained in Sect. 2, the promoter activities of the gene of interest and the reporter gene are identical in transcriptional fusions. In [12,51] it is shown that, if the expression of the gene of interest is adequately described by the model structure of Eqs. 1–3, in particular when assuming constant translational activity (no regulation of translation) and constant half-lives of mRNA and protein (no regulation of mRNA and protein degradation), then the reporter gene model can be used to reconstruct the promoter activity $a(t)$ of the gene of interest, up to an unknown multiplicative coefficient. In the case of active regulation of translation and degradation, transcriptional fusions may not be sufficient to obtain precise estimates and the chromosomal regions involved in post-transcriptional regulation need to be integrated into the reporter construction (Sect. 2). In addition to promoter activities, other quantities of interest can be estimated from the data using the above model and some additional assumptions and extensions, *e.g.*, protein concentrations [12,51] and concentrations of transcriptional regulators [15,19]. In the remainder of this section, for the sake of simplicity, we will focus on the estimation of promoter activities using the reduced model of Eq. 4.

Figure 3A shows an example of typical time-course fluorescence and absorbance data acquired in a reporter gene experiment, derived from [43].

An *E. coli* strain carrying a low-copy reporter plasmid with a transcriptional fusion of the promoter of the gene *tar*, encoding a chemoreceptor protein playing a role in motility. The strain, carrying a deletion of the transcription regulator CpxR, was grown in a thermostated and agitated microplate in minimal medium with glucose for almost 10 h, with a first phase of exponential growth followed by growth arrest after glucose exhaustion. The reporter protein is a stable and fast-maturing GFP, called GFPmut2 [50]. The microplate reader captures fluorescence and absorbance signals at regular time intervals, leading to more than 100 measurements over the time window of the experiment.

A first step in data analysis is the detection of outliers, not necessary in the example of Fig. 3, and the subtraction of background levels of fluorescence and absorbance. The corrected absorbance signal $A(t)$ is computed as

$$A(t) = A_u(t) - A_b(t), \tag{6}$$

where $A_u(t)$ is the primary absorbance signal and $A_b(t)$ the absorbance of the growth medium (M9 in this example). As explained in Sect. 2.3, the fluorescence signal needs to be corrected for autofluorescence generated by wild-type bacteria carrying a non-functional reporter plasmid or no plasmid at all (in practice these two measures of autofluorescence usually give the same result). Contrary to the absorbance background, the autofluorescence depends on the (time-varying) population size. An obvious strategy would be to directly subtract the autofluorescence of the control culture from the fluorescence of the culture of the bacteria carrying the reporter plasmid. This does not always work though, since the two cultures may not be exactly synchronized. As an alternative, one can use a calibration procedure, such that the corrected signal $I(t)$ is defined by

$$I(t) = I_u(t) - s(A(t)), \tag{7}$$

where $I_u(t)$ is the primary fluorescence level and s a calibration function, mapping absorbance levels to autofluorescence levels [43]. The calibration function is obtained by fitting a cubic smoothing spline to the autofluorescence generated by bacteria carrying the non-functional reporter plasmid or no plasmid at all as a function of the absorbance. Splines have the advantage that they can be evaluated for any absorbance within the observed range and easily extrapolated beyond this range. Figure 3B-C gives an example of this strategy of background correction of absorbance and fluorescence data. Another strategy, in principle more powerful than the one used above, would be to measure autofluorescence at a different wavelength and use spectral unmixing to correct autofluorescence [28].

A second step is the computation of the promoter activity from the corrected absorbance and fluorescence signals. From Eq. 5 it follows that

$$k'_m\, a(t) = \frac{d}{dt} r(t) + (d_r + \mu(t))\, r(t). \tag{8}$$

The growth rate $\mu(t)$ can be directly estimated from the absorbance, that is,

$$\mu(t) = \frac{d}{dt} A(t) \frac{1}{A(t)} = \frac{d \ln A(t)}{dt}. \tag{9}$$

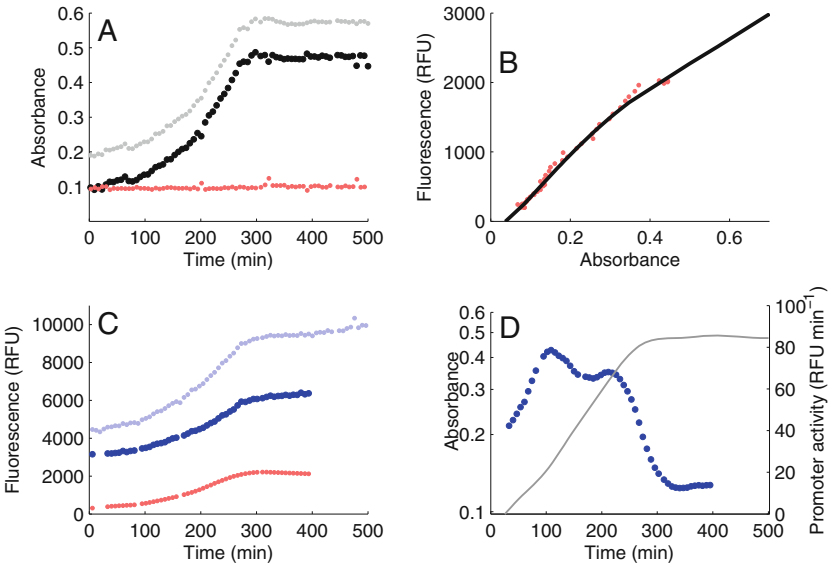

Fig. 3. Illustration of the data analysis procedures (adapted from [43]). Absorbance and fluorescence data acquired for the *ΔcpxR* mutant strain carrying a pUA66*tar-gfp* reporter plasmid, grown in M9 with glucose. *A:* Primary (uncorrected) absorbance (•, grey), background absorbance (•, red), and corrected absorbance (•, black). *B:* Calibration curve obtained by measuring the autofluorescence of the wild-type strain without plasmid. Primary fluorescence data are plotted against (corrected) absorbance data and the curve is obtained by fitting a smoothing spline. *C:* Primary fluorescence data (•, grey), and the corrected fluorescence (•, blue) obtained after subtracting the fluorescence of the background (•, red) as in Eq. 7. *D:* Promoter activity of *tar* (•, blue) computed from the corrected absorbance (-, grey) and corrected fluorescence by means of Eq. 11 (Color figure online).

The time-varying GFP concentration in the bacterial population, $r(t)$, can also be estimated from the absorbance and fluorescence, making the usual assumptions that the fluorescence is proportional to the number of GFP molecules and the absorbance proportional to the biomass (Sect. 2.3):

$$r(t) \sim \frac{I(t)}{A(t)}. \tag{10}$$

We arbitrarily set the proportionality constant in Eq. 10 to 1, thus expressing the reporter protein concentration in relative fluorescence units, RFU (and the synthesis rate in units RFU min^{-1}). Notice that this results in a relative, not an absolute quantification of promoter activity, as is common in the literature. In order to obtain an absolute quantification of promoter activity, an additional calibration step would be necessary to relate fluorescence units to the number of GFP molecules [45].

Substituting the expressions for $r(t)$ and $\mu(t)$ into Eq. 8 yields [12]:

$$k'_m\, a(t) = \frac{dI(t)}{dt}\frac{1}{A(t)} - \frac{dA(t)}{dt}\frac{I(t)}{A(t)^2} + \left(\frac{dA(t)}{dt}\frac{1}{A(t)} + d_r\right)\frac{I(t)}{A(t)}$$
$$= \frac{\frac{d}{dt}I(t)}{A(t)} + d_r\frac{I(t)}{A(t)}. \tag{11}$$

This definition is equivalent to other definitions in the literature [38] when $\mu(t) \gg d_r$. The expression can be evaluated using estimates of $A(t)$, $I(t)$, and $dI(t)/dt$ obtained by means of cubic smoothing splines, following the procedure in [12]. Moreover, the half-life of the fluorescent reporter, and thus the degradation constant d_r, can be easily measured experimentally. Figure 3D shows the promoter activity of *tar* computed from the data in panels *A* and *C*.

The approach outlined above is indirect, in the sense that it smooths the data first and reconstructs the promoter activity from the measurement model only in a second step. This results in the propagation of estimation errors that may be difficult to control. Other methods formulate a regularized data fitting problem directly in terms of the quantities to be estimated, thus solving the inference problem in a single and better controlled optimization step. Examples of the latter approach are linear inversion methods [3,51] and methods based on the use of Kalman filters [1].

3.3 Computer Tools for Analyzing Fluorescent Reporter Gene Data

Population-level experiments with reporter genes in microplate readers generate a huge amount of data, on the order of 10^4–10^5 data points per microplate. In order to treat these data in a systematic and efficient way, and make the methods available to biologists, the latter need to be implemented and packaged in user-friendly computer tools. An example of such a tool is Wellreader, a Matlab program for the analysis of fluorescent and luminescent reporter gene data based on the indirect methods outlined above [7]. Other examples are WellInverter [51] and BasyLICA [1], which employ indirect approaches for inferring gene expression profiles. WellInverter is a web server application that provides a graphical user interface allowing online access to the linear inversion methods through a web-based platform. The user can upload experimental files by means of WellInverter, remove outliers and subtract background, and launch the procedures for computing growth rates, promoter activities, and protein concentrations. The methods underlying WellInverter are also available as a stand-alone Python package (WellFARE).

4 Use of Fluorescent Reporter Genes for Studying Bacterial Regulatory Networks

Reporter gene experiments in microplate readers contain information on population-level gene expression that is highly valuable for studying the

adaptation of bacterial gene expression to changes in the environment. In this section, we provide two examples illustrating the use of such fluorescent reporter gene data sets.

4.1 Joint Control of Gene Expression Changes by Global Physiological Effects and Specific Regulatory Interactions

Regulatory networks controlling adaptation of gene expression in bacteria involve transcription factors that sense environmental and metabolic signals and specifically activate or inhibit target genes. In addition to such specific factors, gene expression also responds to changes in a variety of physiological parameters that modulate the rate of transcription and translation, such as the concentrations of (free) RNA polymerase and ribosome, gene copy number, and the size of amino acid and nucleotide pools. Contrary to specific regulators, these so-called global physiological effects affect the expression of all genes. The importance of global physiological effects, in particular the activity of the transcriptional and translational machinery in the control of gene expression, have been quantitatively assessed in recent publications [6,16,22,24].

It is to be expected that, during the transition from one growth condition to another, both global physiological effects and specific regulatory interactions play a role in reorganizing the gene expression state of the bacterial cell. However, only very few studies have investigated the relative contributions of these factors in an actual regulatory network and dynamically, during a growth transition. Here, we summarize one such study carried out in our group [6], where we considered this question in the context of a central regulatory circuit of carbon metabolism in *E. coli*. The network, shown in Fig. 4*A*, consists of two transcription regulators, Crp and Fis, that regulate the expression of a large number of genes encoding enzymes in central metabolism, including the gene *acs*. Since the latter gene is strongly expressed in the absence of glucose, it provides an excellent indicator of the transcriptional response of carbon metabolism to a change in glucose availability and the accompanying change in growth rate. The signaling metabolite cyclic AMP (cAMP), which is required to activate Crp, also responds to a change in glucose availability by a strong increase of its intracellular concentration. Moreover, the expression of all genes is controlled by the activity of the transcriptional and translational machinery.

The original question was rephrased in the context of this network as follows. How do global physiological effects and the dense pattern of transcriptional regulatory interactions in Fig. 4*A* jointly contribute to the change in promoter activity observed during a growth switch? In particular, we considered the transition of *E. coli* cells growing in minimal medium supplemented with glucose as the sole carbon source to a complete arrest of growth due to the exhaustion of glucose. The cell maintains a minimal metabolism after glucose exhaustion by utilizing the low-energy substrate acetate that has been excreted during fast growth on glucose. During this transition between fast growth on glucose and slow growth on acetate, we measured the growth rate, the concentration of cAMP in the growth medium, from which we estimated the intracellular cAMP

concentration, and the activity of the *acs*, *crp*, and *fis* promoters by means of plasmid-borne fluorescent reporter genes. In addition, in order to quantify the global physiological state we monitored the activity of a constitutively expressed promoter, that is, a promoter whose transcriptional activity only depends on the global physiological state of the cell, in particular the activity of the gene expression machinery [27]. The particular promoter used for this purpose was the p_{RM} promoter, a phage promoter that is not known to be regulated by any specific transcription factor in non-infected *E. coli* cells.

In order to interpret the data obtained in these experiments, we proposed the following simple model of promoter activity. We denote by $p(t)$ the time-varying promoter activity [μM min^{-1}] and write

$$p(t) = k\, p_1(t)\, p_2(t), \tag{12}$$

where k [μM min^{-1}] represents the maximum promoter activity. The dimensionless term $p_1(t)$, for convenience assumed to vary between 0 and 1, quantifies the modulation of the promoter activity by global physiological effects, for instance through the availability of free RNA polymerase. The dimensionless term $p_2(t)$, also varying between 0 and 1, accounts for the effect of transcription factors and other specific regulators. The unknown constant k can be eliminated by normalizing Eq. 12 with respect to a reference state at time t^0, for instance steady-state growth on glucose. We define $p^0 = p(t^0)$, $p_1^0 = p_1(t^0)$, and $p_2^0 = p_2(t^0)$, and divide Eq. 12 by $p^0 = k\, p_1^0\, p_2^0$. A logarithmic transformation of the resulting terms results in

$$\log \frac{p(t)}{p^0} = \log \frac{p_1(t)}{p_1^0} + \log \frac{p_2(t)}{p_2^0}. \tag{13}$$

How can this very simple model be used to answer the question asked above? We considered two extreme cases, namely one in which it is assumed that specific regulators do not play any role in the dynamic regulation of the promoter activity. In this case, the second term in the right-hand side of Eq. 13 drops out of the equation and the model reduces to a simple equality of two terms. For instance, if we are interested in the promoter activity of the gene *crp*, we obtain $\log(p_{crp}(t)/p_{crp}^0) = \log(p_{RM}(t)/p_{RM}^0)$, where as stated above, the activity of the p_{RM} promoter is used as a proxy for the quantification of global physiological effects. This hypothesis can be directly tested with the experimental data by plotting the normalized activities of the *crp* and p_{RM} promoters in a scatter plot (Fig. 4*B*). If the assumption is correct and global physiological effects dominate regulation of *crp* activity, then the data points are expected to cluster around the diagonal, which is indeed seen to be the case ($R^2 = 0.96$). The same conclusion was reached for the *fis* promoter (data not shown), but in the case of the *acs* promoter the hypothesis clearly breaks down (Fig. 4*C*). However, we showed that a slightly adapted model fits the data well. In particular, when assuming that the specific regulatory effect is mostly determined by the concentration of cAMP, we obtain the following expression for the *acs* promoter from Eq. 13: $\log(p_{acs}(t)/p_{acs}^0) = \log(p_{RM}(t)/p_{RM}^0) + \log(c(t)/c^0)$, where c represents the intracellular cAMP concentration. Since this concentration has also been measured, the model can be directly tested [6].

We have shown by means of the above analysis, and similar analyses performed in different growth conditions, with mutants strains, and for an additional target gene coding for a transcriptional regulator (RpoS), that the gene expression profiles can be adequately captured by the much simpler regulatory network shown in Fig. 4*B*. It can be directly seen that in comparison with the original network in panel *A*, apart from the transcriptional regulation of *acs* by the complex of Crp and cAMP, all specific regulatory interactions have disappeared. In other words, the assumption that global physiological effects dominate during the adaptation of *E. coli* cells to the exhaustion of glucose is sufficient to account for the observed gene expression profiles. While the control of gene expression during growth transitions is shared between global physiological effects and specific transcription factors, our results question the central role often attributed to transcriptional regulatory networks in controlling genome-wide expression changes during physiological transitions. It may be more appropriate to regard transcriptional regulators as complementing and finetuning the global control exerted by the physiological state of the cell.

Several other recent studies have used fluorescent reporter genes to study the shared control of gene expression in microorganisms by transcription factors and global cell physiology. For instance, Gerosa *et al.* have developed quantitative models to dissect global and specific regulation of *E. coli* genes involved in arginine biosynthesis [16]. Keren *et al.* have measured activities of 900 *S. cerevisiae* and 1800 *E. coli* promoters using fluorescent reporters. They showed that in both organisms 60 to 90 % of promoters change their expression between conditions by a constant global scaling factor that depends only on the conditions and not on the specific promoter considered, thus also suggesting that global physiological effects are an important driver for the adaptation of gene expression during growth transitions.

4.2 Importance of Protein Stability for Inference and Analysis of Transcriptional Regulatory Networks

The synthesis of flagella and the chemotaxis sensing system, enabling *E. coli* bacteria to orient themselves along gradients of certain chemicals in their environment, is under the control of a complex regulatory network. The more than 60 genes responsible for motility in bacteria are structured in a transcriptional hierarchy of three operon classes [9,21,31]. Several studies have used fluorescent reporter genes to better understand the functional organization of this hierarchy. For instance, Kalir et al. [21] found a detailed temporal ordering of the activation of the different promoters in the hierarchy, suggesting that flagella proteins are synthesized just-in-time, that is, not earlier than needed. Dudin et al. [14] embedded the flagellar hierarchy in the broader context of global regulatory networks in *E. coli*, notably by showing that two well-known transcription regulators, CpxR and CsgD, control the expression of flagellar genes.

In a study carried out by our group, we have focused on a central motif in this transcriptional hierarchy, consisting of the FliA and FlgM transcription factors and their targets (Fig. 5). FliA or σ^{28} is a sigma factor that directs

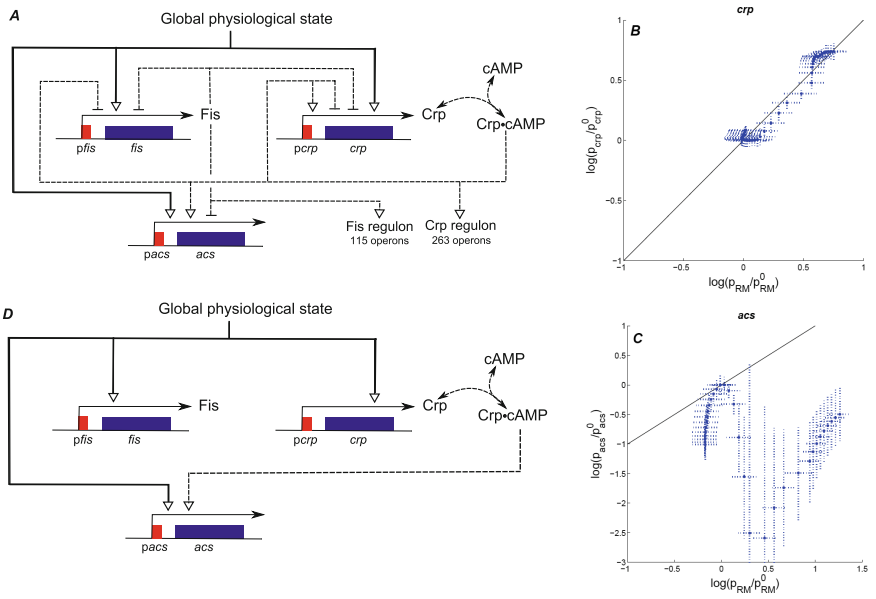

Fig. 4. Shared control of gene expression in bacteria by transcription factors and global physiology of the cell (adapted from [6]). *A:* Central regulatory circuit involved in the control of *E. coli* carbon metabolism, consisting of the two pleiotropic transcription factors Crp and Fis, the signaling metabolite cAMP, a target gene *acs*, and their mutual regulatory interactions. The global physiological state affects the expression of all genes in the network. *B:* Predicted and observed control of *crp* promoter activity by global physiological effetcs. Predicted (–, black) and measured (•, blue) relative activity of the *crp* promoter ($\log(p_{crp}(t)/p^0_{crp})$) as a function of the relative activity of the p_{RM} promoter ($\log(p_{RM}(t)/p^0_{RM})$). *C:* Idem for *acs*. *D:* Reduced regulatory network, including the interactions that were found to dominate the transcriptional response of the network in *A*: the activation of all genes by the physiological state of the cell and the activation of *acs* by Crp·cAMP (Color figure online).

RNA polymerase to operons coding for the flagellar filament and the chemotaxis sensing system controlling the flagellar motor. The effect of FliA is counteracted by the anti-sigma factor FlgM. As a typical example of a FliA-dependent gene we study *tar*, which encodes a chemoreceptor protein Tar that responds to a decrease of the aspartate concentration in the medium. This signal is transmitted to downstream regulators of the flagellar motor [9]. The stability of FliA and FlgM are actively regulated, forming a check-point in the transcriptional hierarchy [4].

An often used, implicit assumption for inferring regulatory interactions from gene expression data is that it is sufficient to focus on the transcriptional level, that is, on measurements of mRNA concentrations or promoter activities. However, this brushes aside the fact that the active regulator of a gene is the protein translated from the mRNA. While the concentration of the two may be correlated to some extent at steady state [30,45], this is certainly not the case

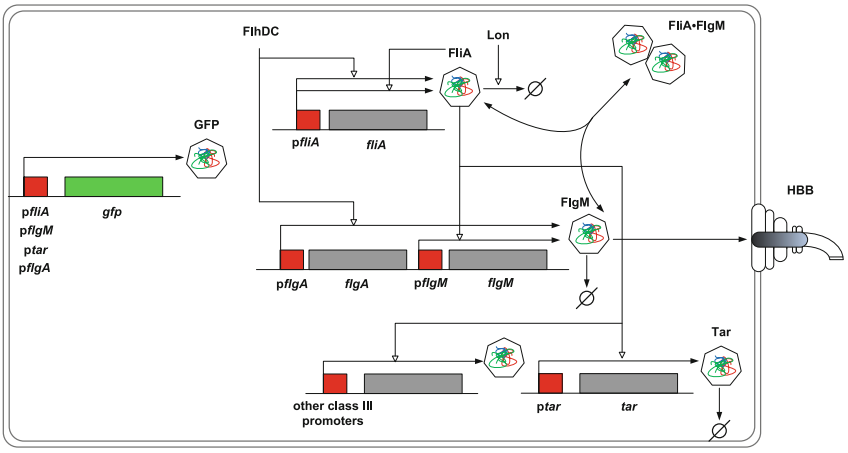

Fig. 5. FliA-FlgM module controlling the expression of motility genes in *E. coli* (adapted from [43]). The regulatory circuit composed of the flagellar-specific transcription factor FliA, a sigma factor also known as σ^{28}, and the anti-sigma factor FlgM forms a check-point in the transcriptional hierarchy of the motility genes in *E. coli*. FliA binds to RNA polymerase core enzyme and directs transcription from a large number of promoters [23], including p*tar* and p*flgM*. When bound to FlgM, FliA cannot activate transcription. When the hook basal-body (HBB) structure is in place, however, FlgM is exported from the cell, thus releasing FliA from the inactive complex. FliA is subject to proteolysis by Lon, but FlgM-binding protects FliA from degradation. The *fliA* promoter is auto-regulated by FliA and by a number of other regulators, most importantly the motility master regulator FlhDC. The activity of the genes in the figure is measured by fusion of their promoters to a *gfp* reporter gene on a low-copy plasmid. Genes are shown in grey or green and their promoter regions in red. Regulatory interactions are represented by open arrows, association and dissociation of FliA and FlgM as well as degradation and export by filled arrows (Color figure online).

in time-course experiments, where the time-varying concentrations of mRNA and translated protein are expected to diverge, due to the different half-lives of mRNA and protein. Typically, the half-live of mRNA is on the order of a few minutes, whereas the half-live of most proteins is longer than 10 h. We therefore asked the question to which extent the use of data on the mRNA level, instead of the protein level, biases the inference process and how this bias could be mitigated, either by adding experimental controls or by strengthening the data analysis procedures. The FliA-FlgM module provides a good test case for this study, since the regulatory interactions have been well-studied and protein turnover, an effect that is not visible on the transcriptional level, plays an important role in its functioning.

In order to address the above question, we measured the time-varying transcription of *fliA*, *flgM*, and *tar* by means of fluorescent reporter systems, consisting of transcriptional fusions of a *gfp*mut2 reporter gene to the promoters of the target genes, carried on a low-copy number plasmid [50]. Five different time-series data

sets were generated, involving different strains and growth media. Specifically, we asked the question if the known model of the *tar* promoter, the expression of which is controlled by FliA and thus FlgM, can be fitted to the data when measured promoter activities are used in the identification process, instead of protein concentrations. The results are shown in Fig. 6*A* and even visual inspection shows that

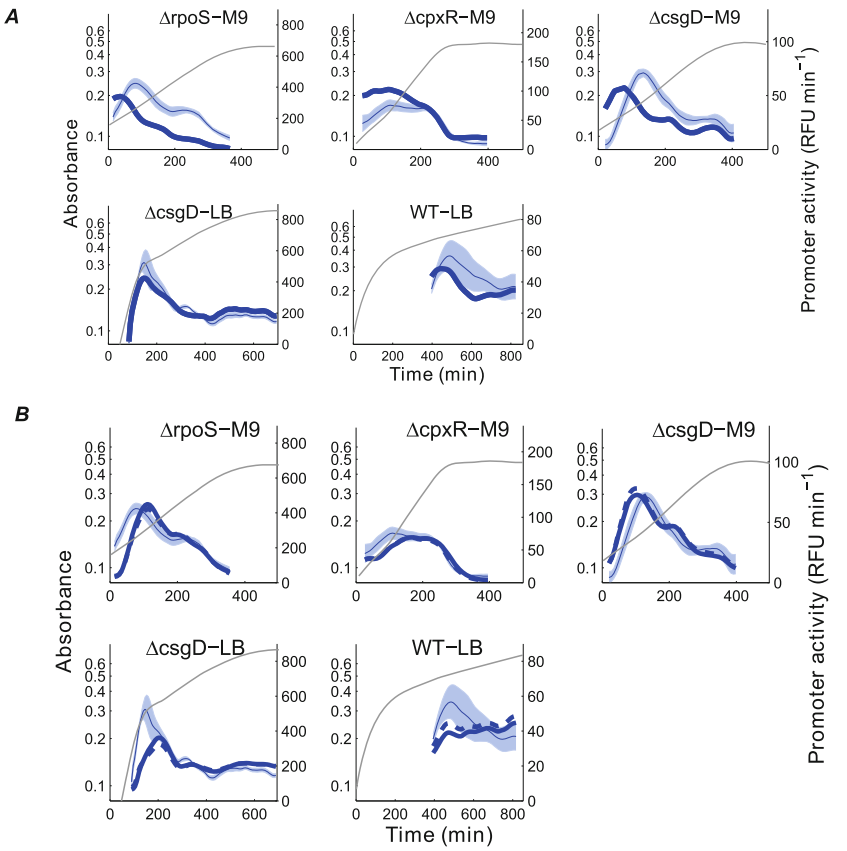

Fig. 6. Promoter activity of *tar* fitted to fluorescent reporter gene data (figure adapted from [43]). Different models of the activity of the *tar* promoter, as a function of the concentrations of FliA and FlgM and (possibly) global physiological effects, were fitted against the measured promoter activity of *tar*. The measurements were carried out in the following experimental conditions: *ΔrpoS* strain grown in M9 (*ΔrpoS*-M9), *ΔcpxR* strain grown in M9 (*ΔcpxR*-M9), *ΔcsgD* strain grown in M9 (*ΔcsgD*-M9), *ΔcsgD* strain grown in LB (*ΔcsgD*-LB), and wild-type strain grown in LB (WT-LB). *A:* Fit of a model of *tar* promoter activity when replacing the concentrations of FliA and FlgM by the measured promoter activities of *fliA* and *flgM*, respectively. Model predictions are in dark blue (thick solid line), *tar* reporter gene data are in light blue (thin solid line and shaded area). *B:* Idem, but when using protein concentrations computed from the measurement model in Sect. 3 and quantification of global physiological effects by means of a constitutive promoter.

the fit is at best of modest quality. While the fit is reasonable in two conditions, it is not satisfactory for the three others [43].

In order to test if the bad fit originates in a bias introduced by ignoring the difference in temporal expression profiles of promoter activities and protein concentrations, we developed computational procedures for reconstructing protein concentrations from promoter activities, given the (approximately) known half-lives of FliA and FlgM in the experimental conditions. These procedures build upon the measurement model of Eq. 4, which allows the promoter activity to be computed from reporter gene data and, through forward integration for a given protein half-life, the protein concentration to be predicted. In addition, following the results of the study in the previous section, we integrated global physiological effects by quantifying the activity of a constitutive promoter. The extended model, fitted against the same data, led to much improved results, shown in Fig. 6*B*. Additional tests showed that both extensions, reconstruction of protein concentrations and integration of global physiological effects, were necessary for obtaining the improved fits.

The results of this study suggest that, in order to reconstruct quantitative models of regulatory networks from time-series gene expression data, it is critical not to neglect the distinction between mRNA and protein. The different time-scales on which the mRNA and protein concentrations evolve causes their temporal profiles to decorrelate and makes mRNA unsuitable as a proxy for transcription factors and other regulatory proteins. While this effect has been illustrated for a single, well-studied module of the motility network, additional simulation studies have shown that the conclusions hold more generally. Due to the fact that the effective half-lives of FliA and FlgM are rather short, in comparison with most of the bacterial proteome, the consequences of using promoter activities rather than protein concentrations in the inference process are actually milder in the case of the FliA-FlgM module than in other situations [43]. While it currently remains difficult to quantify protein concentrations directly, we provided an easy-to-apply procedure for reconstructing protein concentrations from promoter activities, using the measurement models discussed in Sect. 3.1.

From a biological point of view, the results demonstrate the important role played by the active regulation of FliA and FlgM half-lives, through proteolysis and export from the cell, in shaping the dynamics of FliA-dependent promoters.

5 Conclusions

The above examples clearly show that reporter gene data can be used for revealing new connections in gene regulatory networks and for characterizing the dynamical adaptation of bacteria to a changing environment. In order to exploit the tool of dynamical measurements using reporter genes we need to combine an experimental strategy tailored to the specific scientific question with a generic, but flexible measurement model in order to extract the biologically relevant quantities from the data. Here, we have shown that such a measurement model can be developed and implemented in a computer tool to reliably estimate promoter activities and even protein concentrations from the primary data.

The model takes into account inherent difficulties of using fluorescent reporter genes, such as the maturation time of the fluorophore and the background fluorescence of the cell.

We have applied these tools to quantify the contribution of the global physiological state of the cell to changes in gene expression during growth transitions. Surprisingly, the response of individual genes in the network of Fig. 4 is dominated by global regulatory effects instead of specific transcription factors. This result spurs new interest in the study of the dynamics of such global phenomena. Reporter gene experiments can again provide valuable new information that would be difficult to obtain by any other technique because the phenomena we want to study are inherently dynamical. Future experiments will probably focus more on translational fusions of reporter genes to key components of the cell. This strategy will provide quantitative and dynamical measures of the "global physiological state" of the cell. The cellular components to be tagged include RNA polymerase, ribosomes, components of the DNA replication machinery, possibly chaperones, and many more. As mentioned in Sect. 2, translational fusions are more difficult to construct and generally yield signals of lower intensity. However, new and improved fluorescent reporter genes can be expected to emerge to meet these challenges.

A second application of the reporter genes shows the limits of using promoter activity, which corresponds to the concentration of mRNA, as a proxy for the activity of regulatory proteins. We show that the measurement models can be used and adapted to overcome this difficulty. Modifying the experiment may further improve the results. For example, replacing the transcriptional fusion with translational fusions will give us a more direct access to protein concentration. One way to avoid complications of long protein half-lives is to limit experiments to exponential growth where the degradation rate is dominated by growth dilution. However, such a requirement would limit the kinds of biological phenomena that can be studied.

Protein half-lives are an important parameter even for reporter genes. The measurement model developed above is independent of the absolute value of the half-life of the reporter gene. However, since promoter activity is closely related to the derivative of the reporter gene concentration, the accuracy of estimation of this variable may depend on the value of the degradation constant of the reporter gene. This parameter can be adjusted by adding a degradation tag to the end of the protein. Shorter-lived reporter genes will allow the dynamics to be estimated with greater precision, however at the cost of lower signal intensity. Moreover, the concentration of proteases involved in degradation of the reporter may be growth-rate dependent.

Another future improvement of the use of reporter genes concerns the experimental conditions for measuring the dynamics of their expression. The typical experiment involves growth transitions involving the depletion of a nutrient source. This type of experiment is easily carried out in an automated microplate reader as described here. However, in many applications we would like to change environmental conditions in a better controlled and more varied ways. For this

purpose, we will need to develop online measurements of fluorescence and absorbance signals of a bacterial culture grown in a continuous fashion, for example in a chemostat.

Reporter gene measurements, combined with the appropriate measurement models and computational methods, are probably the best tool for studying and understanding the dynamics of gene regulatory networks. Future developments will further enhance the power of this tool.

Acknowledgements. This work was supported by the Investissements d'avenir Bioinformatique programme under project Reset (ANR-11-BINF-0005).

References

1. Aïchaoui, L., Jules, M., Le Chat, L., Aymerich, S., Fromion, V., Goelzer, A.: BasyLiCA: a tool for automatic processing of a Bacterial Live Cell Array. Bioinf. **28**(20), 2705–2706 (2012)
2. Alon, U.: An Introduction to Systems Biology: Design Principles of Biological Circuits. Chapman & Hall/CRC, Boca Raton (2007)
3. Bansal, L., Chu, Y., Laird, C., Hahn, J.: Determining transcription factor profiles from fluorescent reporter systems involving regularization of inverse problems. In: Proceedings of the 2012 American Control Conference (ACC 2012), pp. 2725–30 (2012)
4. Barembruch, C., Hengge, R.: Cellular levels and activity of the flagellar sigma factor FliA of Escherichia coli are controlled by FlgM-modulated proteolysis. Mol. Microbiol. **65**(1), 76–89 (2007)
5. Bernstein, J., Khodursky, A., Lin, P.H., Lin-Chao, S., Cohen, S.: Global analysis of mRNA decay and abundance in Escherichia coli at single-gene resolution using two-color fluorescent DNA microarrays. Proc. Natl. Acad. Sci. USA **99**(15), 9697–9702 (2002)
6. Berthoumieux, S., de Jong, H., Baptist, G., Pinel, C., Ranquet, C., Ropers, D., Geiselmann, J.: Shared control of gene expression in bacteria by transcription factors and global physiology of the cell. Mol. Syst. Biol. **9**(634), 634 (2013)
7. Boyer, F., Besson, B., Baptist, G., Izard, J., Pinel, C., Ropers, D., Geiselmann, J., de Jong, H.: WellReader: a MATLAB program for the analysis of fluorescence and luminescence reporter gene data. Bioinf. **26**(9), 1262–1263 (2010)
8. Browning, D.F., Busby, S.J.W.: The regulation of bacterial transcription initiation. Nat. Rev. Microbiol. **2**(1), 57–65 (2004)
9. Chevance, F., Hughes, K.: Coordinating assembly of a bacterial macromolecular machine. Nat. Rev. Microbiol. **6**, 455–465 (2008)
10. Chudakov, D., Matz, M., Lukyanov, S., Lukyanov, K.: Fluorescent proteins and their applications in imaging living cells and tissues. Physiol. Rev. **90**(3), 1103–1163 (2010)
11. Cox, J., Mann, M.: Quantitative, high-resolution proteomics for data-driven systems biology. Curr. Opin. Biotechnol. **80**, 273–299 (2011)
12. de Jong, H., Ranquet, C., Ropers, D., Pinel, C., Geiselmann, J.: Experimental and computational validation of models of fluorescent and luminescent reporter genes in bacteria. BMC Syst. Biol. **4**(1), 55 (2010)

13. Dharmadi, Y., Gonzalez, R.: DNA microarrays: experimental issues, data analysis, and application to bacterial systems. Biotechnol. Prog. **20**(5), 1309–1324 (2004)
14. Dudin, O., Geiselmann, J., Oqasawara, H., Ishihama, A., Lacour, S.: Repression of flagellar genes in exponential phase by CsgD and CpxR, two crucial modulators of Escherichia coli biofilm formation. J. Bacteriol. **196**(3), 707–715 (2014)
15. Finkenstädt, B., Heron, E., Komorowski, M., Edwards, K., Tang, S., Harper, C., Davis, J., White, M., Millar, A., Rand, D.: Reconstruction of transcriptional dynamics from gene reporter data using differential equations. Bioinf. **24**(24), 2901–2907 (2008)
16. Gerosa, L., Kochanowski, K., Heinemann, M., Sauer, U.: Dissecting specific and global transcriptional regulation of bacterial gene expression. Mol. Syst. Biol. **9**, 658 (2013)
17. Giepmans, B., Adams, S., Ellisman, M., Tsien, R.: The fluorescent toolbox for assessing protein location and function. Sci. **312**(5771), 217–224 (2006)
18. Hebisch, E., Knebel, J., Landsberg, J., Frey, E., Leisner, M.: High variation of fluorescence protein maturation times in closely related Escherichia coli strains. PLoS ONE **8**(10), e75991 (2013)
19. Huang, Z., Senocak, F., Jayaraman, A., Hahn, J.: Integrated modeling and experimental approach for determining transcription factor profiles from fluorescent reporter data. BMC Syst. Biol. **2**, 64 (2008)
20. Ingolia, N.: Ribosome profiling: new views of translation, from single codons to genome scale. Nat. Rev. Genet. **15**(3), 205–213 (2014)
21. Kalir, S., McClure, J., Pabbaraju, K., Southward, C., Ronen, M., et al.: Ordering genes in a flagella pathway by analysis of expression kinetics from living bacteria. Sci. **292**(5524), 2080–2083 (2001)
22. Keren, L., Zackay, O., Lotan-Pompan, M., Barenholz, U., Dekel, E., et al.: Promoters maintain their relative activity levels under different growth conditions. Mol. Syst. Biol. **9**, 701 (2013)
23. Keseler, I., Collado-Vides, J., Santos-Zavaleta, A., Peralta-Gi, M., Gama-Castro, S., et al.: EcoCyc: a comprehensive database of Escherichia coli biology. Nucleic Acids Res. **39**, D583–D590 (2011)
24. Klumpp, S., Zhang, Z., Hwa, T.: Growth rate-dependent global effects on gene expression in bacteria. Cell **139**(7), 1366–1375 (2009)
25. Koga, K., Harada, T., Shimizu, H., Tanaka, K.: Bacterial luciferase activity and the intracellular redox pool in Escherichia coli. Mol. Genet. Genom. **274**(2), 180–188 (2005)
26. Leveau, J., Lindow, S.: Predictive and interpretive simulation of green fluorescent protein expression in reporter bacteria. J. Bacteriol. **183**(23), 6752–6762 (2001)
27. Liang, S., Bipatnath, M., Xu, Y., Chen, S., Dennis, P., Ehrenberg, M., Bremer, H.: Activities of constitutive promoters in Escherichia coli. J. Mol. Biol. **292**(1), 19–37 (1999)
28. Lichten, C., White, R., Clark, I., Swain, P.: Unmixing of fluorescence spectra to resolve quantitative time-series measurements of gene expression in plate readers. BMC Biotechnol. **14**, 11 (2014)
29. Longo, D., Hasty, J.: Dynamics of single-cell gene expression. Mol. Syst. Biol. **2**, 64 (2006)
30. Lu, P., Vogel, C., Wang, R., Yao, X., Marcotte, E.: Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nat. Biotechnol. **25**(1), 117–124 (2007)

31. Macnab, R.: Flagella and motility. In: Neidhardt, F., Curtiss III, R., Ingraham, J., Lin, E., Low, K., Magasanik, B., Reznikoff, W., Riley, M., Schaechter, M., Umbarger, H. (eds.) Escherichia coli and Salmonella: Cellular and Molecular Biology, pp. 123–45. ASM Press, Washington, DC, 2nd edn. (1996)
32. Mäder, U., Nicolas, P., Richard, H., Bessières, P., Aymerich, S.: Comprehensive identification and quantification of microbial transcriptomes by genome-wide unbiased methods. Curr. Opin. Biotechnol. **22**(1), 32–41 (2011)
33. Manen, D., Pougeon, M., Damay, P., Geiselmann, J.: A sensitive reporter gene system using bacterial luciferase based on a series of plasmid cloning vectors compatible with derivatives of pBR322. Gene **186**(2), 197–200 (1997)
34. Van Dyk, T., Wei, Y., Hanafey, M., Dolan, M., Reeve, M., Rafalski, J., Rothman-Denes, L., LaRossa, R.: A genomic approach to gene fusion technology. Proc. Natl. Acad. Sci. USA **98**(5), 2555–2560 (2001)
35. Meighen, E.A.: Bacterial bioluminescence: organization, regulation, and application of the lux genes. FASEB J. **7**(11), 1016–1022 (1993)
36. Muzzey, D., van Oudenaarden, A.: Quantitative time-lapse fluorescence microscopy in single cells. Annu. Rev. Cell. Dev. Biol. **25**, 301–327 (2009)
37. Porreca, R., Cinquemani, E., Lygeros, J., Ferrari-Trecate, G.: Structural identification of unate-like genetic network models from time-lapse protein concentration measurements. In: Proceedings of 49th IEEE Conference on Decision and Control (CDC 2010), pp. 2529–2534 (2010)
38. Ronen, M., Rosenberg, R., Shraiman, B., Alon, U.: Assigning numbers to the arrows: Parameterizing a gene regulation network by using accurate expression kinetics. Proc. Natl. Acad. Sci. USA **99**(16), 10555–10560 (2002)
39. Rowe, L., Dikici, E., Daunert, S.: Engineering bioluminescent proteins: Expanding their analytical potential. Anal. Chem. **81**(21), 8662–8668 (2009)
40. Schaechter, M., Ingraham, J., Neidhardt, F.: Microbe. ASM Press, Washington DC (2006)
41. Sharan, S.K., Thomason, L.C., Kuznetsov, S.G., Court, D.L.: Recombineering: a homologous recombination-based method of genetic engineering. Nat. Protoc. **4**(2), 206–223 (2009)
42. Silhavy, T.: Gene fusions. J. Bacteriol. **182**(21), 5935–5938 (2000)
43. Stefan, D., Pinel, C., Pinhal, S., Cinquemani, E., Geiselmann, J., de Jong, H.: Inference of quantitative models of bacterial promoters from time-series reporter gene data. PLoS Comput. Biol. **11**(1), e1004028 (2015)
44. Süel, G.: Use of fluorescence microscopy to analyze genetic circuit dynamics. Methods Enzymol. **497**, 275–293 (2011)
45. Taniguchi, Y., Choi, P., Li, G.W., Chen, H., Babu, M., Hearn, J., Emili, A., Xie, X.: Quantifying E. coli proteome and transcriptome with single-molecule sensitivity in single cells. Sci. **329**(5991), 533–539 (2010)
46. Pédelacq, J.D., Cabantous, S., Tran, T., Terwilliger, T.C., Waldo, G.S.: Engineering and characterization of a superfolder green fluorescent protein. Nat. Biotechnol. **24**(1), 79–88 (2006)
47. Tsien, R.Y.: The green fluorescent protein. Annu. Rev. Biochem. **67**, 509–544 (1998)
48. Wang, X., Errede, B., Elston, T.: Mathematical analysis and quantification of fluorescent proteins as transcriptional reporters. Biophys. J. **94**(6), 2017–2026 (2008)

49. Young, J., Locke, J., Altinok, A., Rosenfeld, N., Bacarian, T., Swain, P., Mjolsness, E., Elowitz, M.: Measuring single-cell gene expression dynamics in bacteria using fluorescence time-lapse microscopy. Nat. Protoc. **7**(1), 80–88 (2011)
50. Zaslaver, A., Bren, A., Ronen, M., Itzkovitz, S., Kikoin, I., Shavit, S., Liebermeister, W., Surette, M., Alon, U.: A comprehensive library of fluorescent transcriptional reporters for Escherichia coli. Nat. Methods **3**(8), 623–628 (2006)
51. Zulkower, V., Page, M., Ropers, D., Geiselmann, J., de Jong, H.: Robust reconstruction of gene expression profiles from reporter gene data using linear inversion. Bioinf. **31**(12), i71–i79 (2015)

http://www.springer.com/978-3-319-27655-7

Hybrid Systems Biology
Second International Workshop, HSB 2013, Taormina, Italy, September 2, 2013 and Third International Workshop, HSB 2014, Vienna, Austria, July 23-24, 2014, Revised Selected Papers
Maler, O.; Halasz, A.; Dang, T.; Piazza, C. (Eds.)
2015, XI, 175 p. 51 illus. in color., Softcover
ISBN: 978-3-319-27655-7