Modeling feedback loops in the H-NS-mediated regulation of the *Escherichia coli bgl* operon

Nicole Radde[a],*, Jutta Gebert[a], Ulrich Faigle[a], Rainer Schrader[a], Karin Schnetz[b]

[a]*Center for Applied Computer Science, University of Cologne, Weyertal 80, 50931 Cologne, Germany*
[b]*Institute for Genetics, University of Cologne, Zülpicher Straße 47, 50674 Cologne, Germany*

Abstract

The histone-like nucleoid-associated protein H-NS is a global transcriptional repressor that controls approximately 5% of all genes in *Escherichia coli* and other enterobacteria. H-NS binds to DNA with low specificity. Nonetheless, repression of some loci is exceptionally specific. Experimental data for the *E. coli bgl* operon suggest that highly specific repression is caused by regulatory feedback loops. To analyze whether such feedback loops can account for the observed specificity of repression, here a model was built based on expression data. The model includes several regulatory interactions, which are synergy of repression by binding of H-NS to two regulatory elements, an inverse correlation of the rate of repression by H-NS and transcription, and a threshold for positive regulation by anti-terminator BglG, which is encoded within the operon. The latter two regulatory interactions represent feedback loops in the model. The resulting system of equations was solved for the expression level of the operon and analyzed with respect to different promoter activities. This analysis demonstrates that a small (3-fold) increase of the *bgl* promoter activity results in a strong (80-fold) enhancement of *bgl* operon expression. Thus, the parameters included into the model are sufficient to simulate specific repression by H-NS.
© 2007 Elsevier Ltd. All rights reserved.

1. Introduction

Understanding the dynamics of biochemical networks in a cell is one of the long-term goals in systems biology. A classical bottom-up approach focuses on small regulatory subsystems (see, for example, Yildirim and Mackey, 2003; Roeder and Glauche, 2006; Radde et al., 2006; Gebert and Radde, 2006). In this paper, we show that with the interpretation of very few experiments a model for the exceptional specificity of repression of the *Escherichia coli bgl* operon by the nucleoid-associated protein H-NS can be set up. H-NS is an abundant global repressor that binds to DNA with low specificity and affects many genes and cellular processes in enterobacteria (Dorman, 2004). H-NS has also been shown to repress genes acquired by horizontal gene transfer and to increase bacterial fitness

(Navarre et al., 2006; Lucchini et al., 2006; Dorman, 2006). H-NS binds preferentially to AT-rich and curved DNA. Binding of H-NS dimers to such AT-rich and curved 'nucleation sites' and subsequent oligomerization along the DNA results in the formation of extended nucleoprotein complexes (Dorman, 2004). Formation of the nucleoprotein complex may involve DNA loop formation, since H-NS can bind two DNA double helixes (Dame et al., 2006; Dorman, 2006). Thus, binding of H-NS and formation of a nucleoprotein complex close to a promoter causes repression by trapping of RNA polymerase at the promoter or by excluding RNA polymerase from the promoter (Dorman, 2004, 2006).

Although the DNA-binding specificity of H-NS is low, some loci are very specifically repressed by H-NS, and up to date it is an open question how this is achieved. The best studied examples for highly specific repression by H-NS are the *proU* operon present in *E. coli* and *Salmonella typhimurium* and the *bgl* operon of *E. coli*. The *proU* operon encodes a high-affinity uptake system for

*Corresponding author. Tel.: +49 221 4706024; fax: +49 221 4705160.
E-mail addresses: radde@zpr.uni-koeln.de (N. Radde), schnetz@uni-koeln.de (K. Schnetz).
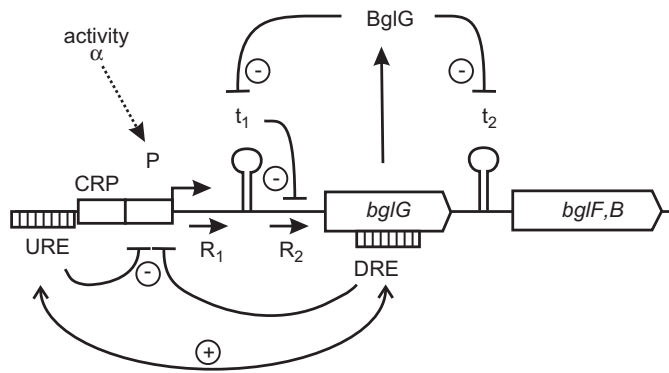
Fig. 1. Regulation of the *Escherichia coli bgl* operon. Expression of the *bgl* operon is directed by the CRP-dependent promoter (P). Transcription initiation at the promoter is repressed by binding of H-NS to an upstream regulatory element (URE) and a downstream regulatory element (DRE). Repression by H-NS through the URE and DRE is synergistic. Gene *bglG* encodes an anti-terminator BglG, which prevents transcription termination at two terminators t_1 and t_2. Genes *bglF* and *bglB* encode an enzyme II permease and a phospho-β-glucosidase. Not shown is regulation of BglG activity by the permease BglF and by the PTS protein HPr (Deutscher et al., 2006; Görke, 2003). α is the promoter activity. R_1 and R_2 indicate transcription rates next to the promoter and downstream of t_1, respectively.

osmoprotectants (Csonka, 1982; Csonka and Epstein, 1996). At low osmolarity it is repressed by H-NS. At high osmolarity repression by H-NS is relieved and expression is approximately 200-fold induced. Highly specific repression of *proU* has remained enigmatic (Jordi and Higgins, 2000; Bouffartigues et al., 2007). The *bgl* operon encodes the gene products for the uptake and fermentation of aryl-β, D-glucosidic sugars (Fig. 1) (Schaefler and Maas, 1967; Prasad and Schaefler, 1974). It is repressed approximately 100-fold by H-NS (Higgins et al., 1988; Schnetz, 1995; Dole et al., 2004), while its repression by H-NS can be completely relieved by a merely 3-fold increase in the promoter activity (Schnetz, 2002; Dole et al., 2002). Interestingly, repression of both operons involves a downstream regulatory element (DRE) in addition to an upstream regulatory element (URE), to which H-NS binds (Dattananda et al., 1991; Overdier and Csonka, 1992; Owen-Hughes et al., 1992; Schnetz, 1995; Dole et al., 2004), and repression of both loci occurs at an early step of transcription, prior to open complex formation by RNA polymerase at the promoter (Jordi and Higgins, 2000; Nagarajavel et al., 2007). Further, in both loci repression by binding of H-NS to the URE and the DRE is synergistic, and repression by H-NS through the DRE correlates inversely to the promoter activity (Nagarajavel et al., 2007). This inverse correlation suggests that repression is overcome by high rates of productive transcription initiation or by high rates of transcription elongation across the DRE (Nagarajavel et al., 2007). This inverse correlation of repression and transcription represents a regulatory feedback loop, that presumably is important for highly specific repression of the *proU* and *bgl* loci by H-NS (Nagarajavel et al., 2007). In case of the

bgl operon (Fig. 1), an additional level of regulation that may further increase specificity of repression by H-NS has been defined (Dole et al., 2002). This additional level of regulation involves positive auto-regulation by the operon encoded transcriptional anti-terminator BglG (Mahadevan and Wright, 1987; Schnetz et al., 1987; Schnetz and Rak, 1988) (Fig. 1). BglG prevents termination of transcription at two terminators, t_1 and t_2, located within the operon by binding and concomitant stabilization of an RNA hairpin (called RAT) (Aymerich and Steinmetz, 1992; Schnetz and Rak, 1988). BglG-assisted formation of the RAT hairpin prevents formation of the terminator hairpin and thus causes 'anti'-termination. Since terminator t_1 precedes the *bglG* gene, BglG is positively auto-regulated. Furthermore, anti-termination by BglG does not occur at low expression levels, but only if a threshold expression level is exceeded (Dole et al., 2002). Auto-regulation by BglG and the threshold requirement resembles another feedback loop in *bgl* operon regulation presumably important for highly specific repression by H-NS (Dole et al., 2002).

Here, we present a data-driven model for the role of regulatory feedback loops in specific repression of the *bgl* operon by H-NS. Parameters of our model are estimated using measurements of the expression level of *lacZ* fusions in the wild type and *hns* mutant. The model includes two regulatory feedback loops, which are the inverse correlation of transcription and H-NS-mediated repression through the DRE, and the auto-regulation by anti-terminator BglG. In the first part, parameterized functions for the description of the network are set up. In the second part, the model is solved for the expression level of the operon, and results are shown for different promoter activities. The model shows that a small enhancement of the promoter activity results in a manifold increased expression due to the orchestrated regulatory interactions provided by the regulatory feedback loops.

2. A model for regulation of the *bgl* operon

In order to build a data-driven model for regulation of the *bgl* operon, two simplified regulatory networks were deduced from the known regulatory interactions (Fig. 2A and B). In both simplified regulatory networks the intrinsic activity of the cAMP receptor protein (CRP)-dependent *bgl* promoter is included as an external variable α. α is one variable that determines the rate of productive transcription initiation R_1. The transcription initiation rate R_1 is negatively affected by binding of H-NS to the URE and DRE (indicated by arrows in Fig. 2). Synergy of repression by H-NS through the URE and DRE (Nagarajavel et al., 2007) is indicated by a double arrow (Fig. 2). Furthermore, the inverse correlation of repression by H-NS through the DRE and the promoter activity (Nagarajavel et al., 2007) was built into the simplified regulatory network in two alternative ways. In model A (Fig. 2A) it is assumed that repression through the DRE inversely correlates to the transcription initiation rate R_1, while in model B (Fig. 2B)
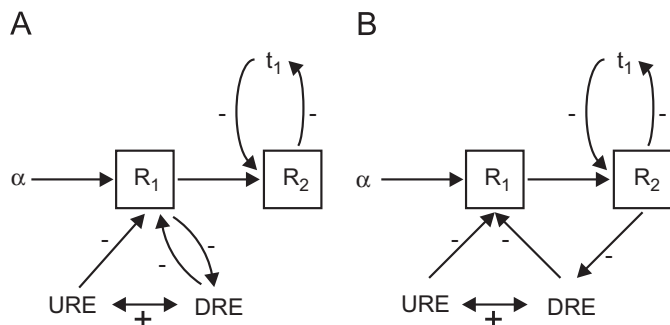
A

B

Fig. 2. Simplified regulatory network for regulation of *bgl*. The transcription initiation rate R_1 is affected by the intrinsic promoter activity α and repression by H-NS through URE and DRE (upstream and downstream regulatory elements). Transcription termination at terminator t_1 negatively regulates R_2. The active protein BglG, whose amount is approximated by R_2, acts as an anti-terminator and negatively regulates t_1. Model A: Repression by H-NS through the DRE inversely correlates to transcription rate R_1. Hence, the corresponding feedback loop includes only R_1 and the DRE. Model B: The repression by H-NS through the DRE inversely correlates to the transcription elongation rate R_2. Thus, the network contains a feedback loop including R_1, R_2, and DRE.

it is assumed that repression through the DRE inversely correlates to the transcription elongation rate R_2. Thus models A and B both contain a positive feedback loop. The feedback loop in model A includes R_1, which determines repression through DRE, and DRE, which determines R_1. The feedback loop in model B includes R_1, which determines R_2, R_2, which determines repression through DRE, and DRE, which determines R_1.

The additional level of regulation of the *bgl* operon based on transcriptional termination and on BglG-mediated specific anti-termination (Schnetz and Rak, 1988; Mahadevan and Wright, 1987) is included as follows. The rate of termination at t_1 located in the leader of the operon (Fig. 1) determines the ratio of R_2 to R_1, which is indicated by an arrow from t_1 to R_2 (Fig. 2). Anti-termination depends on the synthesis of sufficient BglG protein. In the simplified model, the transcription rate R_2 is a measure for the rate of *bglG* transcription and BglG synthesis, and thus the amount of active BglG. Therefore, BglG-mediated anti-termination at t_1 is directly correlated to the transcription rate R_2, which is indicated by an arrow from R_2 to t_1 (Fig. 2). Summarizing both negative regulations results in positive auto-regulation of R_2. To account for a threshold behavior of anti-termination by BglG (Dole et al., 2002), the auto-regulation of R_2 is modeled as a Boolean function. The corresponding equations for all of these regulatory interactions and the estimation of parameters are given below.

2.1. Regulation of transcription initiation rate R_1

The transcription initiation rate R_1 depends on the promoter activity α and repression through the upstream and downstream silencers (URE and DRE), as shown in Fig. 2. We describe these dependencies as a function that is linearly increasing with α. The proportionality factor

depends on repression by H-NS through the regulatory elements. Furthermore, we assume that this proportionality factor is given as a product of different factors c_{ure}, $c_{dre}(R_{1,2})$, and c_{syn}. Here, c_{ure} depends solely on repression through the URE, $c_{dre}(R_{1,2})$ depends on repression through the downstream silencer, which is itself regulated by the transcription rate, and c_{syn} accounts for the synergy of repression through both the URE and DRE. According to this parameterization, the transcription rate R_1 is described by

$$R_1(\alpha, ure, dre(R_{1,2}))$$

$$= \begin{cases} \alpha & \text{no H-NS,} \\ c_{ure}\alpha & \text{URE but no DRE,} \\ c_{dre}(R_{1,2})\alpha & \text{DRE but no URE,} \\ c_{ure}c_{dre}(R_{1,2})c_{syn}\alpha & \text{URE and DRE} \end{cases} \quad (1)$$

with factors $c_{ure}, c_{dre}(R_{1,2}), c_{syn} \in [0,1]$. The inverse correlation of c_{dre} and the transcription rate R_1 or R_2, respectively, is described in the following subsection.

To estimate these factors, we use experimental data on the H-NS-mediated repression of a *lacUV5* promoter (P_{UV5}) that is flanked by both the URE and DRE or by the URE or the DRE alone (Nagarajavel et al., 2007), summarized in Fig. 3. The repression of these *lacZ* fusions by H-NS was determined by measuring the expression level in the wild type and in an *hns* null mutant, in which the complete *hns* gene was deleted (Nagarajavel et al., 2007).

According to Eq. (1), c_{ure} is given by the ratio of the transcription rates R_1 of a reporter fusion containing the upstream silencer (URE$-P_{UV5}$), which are $R_1 = c_{ure}\alpha$ in

	β-galactosidase activity	
	wt	Δ*hns*

A

| | 376 | 405 |

B

| | 239 | 860 |

C

| | 39 | 763 |

Fig. 3. Synergy in repression by H-NS through upstream and downstream regulatory elements (URE and DRE) (data from Nagarajavel et al. (2007). (A) Repression of the *lacUV5* promoter by H-NS through the URE only, (B) through the DRE, and (C) in the presence of the URE and DRE. The expression values in the wild type (wt) and an *hns* null mutant (from Nagarajavel et al. (2007)) were used to estimate c_{ure}, c_{dre}, and c_{syn}. Expression is independent of BglG due to mutation t_1-RAT.

the wild type and $R_1 = \alpha$ in a strain lacking H-NS. This ratio is given by the β-galactosidase activities measured in the wild type and the *hns* mutant (Fig. 3A):

$$\frac{\text{URE}-P_{UV5} \text{ in wt}}{\text{URE}-P_{UV5} \text{ in } hns \text{ mutant}} = c_{ure} = \frac{376}{405} = 0.9. \tag{2}$$

Expression data of a *lacZ* reporter construct, $P_{UV5}-$DRE, which carries the *lacUV5* promoter followed by the DRE (Nagarajavel et al., 2007), were used to obtain a value $c_{dre}(R_{1,2})$ for fixed $R_{1,2}$. Expression of this reporter construct is independent of BglG-mediated anti-termination, due to a mutation of terminator t_1. The t_1 RAT mutation stabilizes the RAT hairpin, which thus forms without binding of BglG and constitutively prevents termination. According to Eq. (1), c_{dre} is given by the ratio of the transcription rate $R_1 = c_{dre}(R_{1,2})\alpha$ in the wild type and $R_1 = \alpha$ in the *hns* mutant (Fig. 3B):

$$\frac{P_{UV5}- \text{ DRE in wt}}{P_{UV5}- \text{ DRE in } hns \text{ mutant}} = c_{dre} = \frac{239}{860} = 0.3. \tag{3}$$

Experimental data to estimate the repression factor c_{syn} were also taken from Nagarajavel et al. (2007). A *lacZ* reporter construct, $\text{URE}-P_{UV5}-$DRE, that contains the URE, the *lacUV5* promoter, the terminator mutant t_1-RAT, and the DRE directed 39 units of β-galactosidase activity in the wild type and 763 units in the *hns* mutant (Fig. 3C). As before, the β-galactosidase activities in the wild type and *hns* mutant were used to estimate c_{syn}, yielding in Eq. (1):

$$\frac{\text{URE}-P_{UV5}-\text{DRE in wt}}{\text{URE}-P_{UV5}-\text{DRE in } hns \text{ mutant}} = c_{ure}c_{dre}(R_{1,2})c_{syn} = \frac{39}{763}, \tag{4}$$

which leads to a factor

$$c_{syn} = \frac{39}{739c_{ure}c_{dre}(R_{1,2})} = \frac{39}{739 \times 0.9 \times 0.3} = 0.2. \tag{5}$$

2.2. Positive feedback loop in repression by H-NS binding to the downstream silencer

Repression by H-NS through the DRE inversely correlates with the promoter activity, which is included in the simplified network by a positive feedback loop from R_1 to DRE (Fig. 2A) or from R_2 to DRE (Fig. 2B). This inverse correlation is described by an $R_{1,2}$-dependent repression factor $c_{dre}(R_{1,2})$ in Eq. (1). Experimental data to estimate $c_{dre}(R_{1,2})$ were taken from expression data obtained with *lacZ* reporter fusions, in which repression of promoters of different activities (P_{lacI}, P_{UV5}, and P_{tac}) by H-NS through the DRE (in the absence of the URE) was determined (Nagarajavel et al., 2007). The terminator t_1 is missing in these reporter fusions, and thus R_1 equals R_2. The experimental data are plotted in Fig. 4 as the ratio of
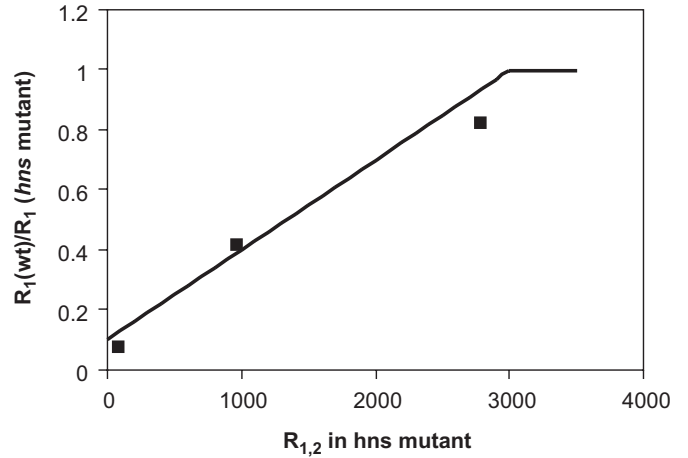

Fig. 4. Increasing transcription rates lower the efficiency of repression by H-NS. Fusion of the downstream silencer to the promoters P_{lacI}, P_{UV5}, and P_{tac} of low, medium, and high activity demonstrated an inverse correlation of the promoter activity and of the efficiency of repression by H-NS through the DRE (Nagarajavel et al., 2007). This is described by a linear relation between $R_{1,2}$ in *hns* mutants and the repression factor $c_{dre}(R_{1,2})$, which saturates at 1.

the expression in the wild type and the *hns* mutant (*y*-axis) versus the unrepressed expression level of the reporter fusion determined in the *hns* mutant (*x*-axis). We use a linear function with a biologically plausible upper limit to describe these data:

$$c_{ure}(R_{1,2}) = \begin{cases} aR_{1,2} + b & \text{wild type and } R_{1,2} \leqslant \dfrac{1-b}{a} := R_{1,2}^*, \\ 1 & hns \text{ mutant or } R_{1,2} > R_{1,2}^*. \end{cases} \tag{6}$$

The regression parameters a and b were estimated using the least squares method, which minimizes the sum of squared errors between experimental repression factors and model predictions. Estimated values are given by $a = 0.3 \times 10^{-3}$ and $b = 0.1$ and hence $R_{1,2}^* = 3 \times 10^3$. The coefficient of determination is 0.98. For transcription rates $R_{1,2}$ exceeding the upper limit $R_{1,2}^*$, no repression through the DRE occurs. Thus, the ratio of R_1 in the wild type and in the *hns* mutant was assumed to be 1 in this range. The regression function is also shown in Fig. 4.

2.3. Regulation of R_2 by BglG-mediated anti-termination at terminator t_1

The transcription rate R_2 is a function that depends on R_1 and the terminator t_1. Termination at t_1 in turn is regulated by the amount of active BglG, which is approximated by R_2. If terminator t_1 is missing, there is no difference between R_1 and R_2, thus $R_1 = R_2$. If terminator t_1 is present R_2 is assumed to be proportional to R_1 with a proportionality factor c_{t_1} that depends on termination of transcription at t_1 and anti-termination by BglG, i.e. the amount of active BglG. We describe c_{t_1} as a

Boolean function with a threshold value R_2^{th}. Hence R_2 is determined by

$$R_2(R_1, t_1(R_2)) = \begin{cases} R_1 & \text{no } t_1, \\ c_{t_1}^{min} R_1 & \text{with } t_1 \text{ and } R_2 \leqslant R_2^{th}, \\ c_{t_1}^{max} R_1 & \text{with } t_1 \text{ and } R_2 > R_2^{th}. \end{cases} \quad (7)$$

The values for $c_{t_1}^{min}$ and $c_{t_1}^{max}$ were determined experimentally (Fig. 5). To this end, terminator t_1 was inserted between the constitutive *lacUV5* promoter and the reporter gene *lacZ*. The β-galactosidase activity directed by this construct was used as a measure for R_2. The construct contains neither the upstream nor the downstream silencer and thus is independent of H-NS, as confirmed by expression analysis in an *hns* mutant (data not shown). The expression level of this reporter construct was determined in the absence of BglG and upon expression of BglG encoded by plasmid pKESK10 provided *in trans* (Dole et al., 2002). Cells were grown in minimal M9 glycerol medium to OD600 = 0.5 and the β-galactosidase activity was determined as described (Miller, 1992; Dole et al., 2002). In the absence of BglG 1490 Miller units of β-galactosidase activity were detected, while the expression increased to 4385 units in the presence of BglG (Fig. 5). In order to normalize the rate of termination and anti-termination, an additional construct was used in which terminator t_1 was inactivated by the t_1-RAT mutation that renders expression independent of BglG (Nagarajavel et al., 2007). Expression of this construct is constitutive, and it directs the expression of 4775 units of β-galactosidase activity (Fig. 5). Taken together these data show that read-through of transcription at terminator t_1 is approximately 30% in the absence of BglG, and 90% in the presence of BglG (Fig. 5). Hence, the parameter $c_{t_1}^{max}$ was set to $c_{t_1}^{max} = 0.9$, while parameter $c_{t_1}^{min}$ was set to $c_{t_1}^{min} = 0.3$.
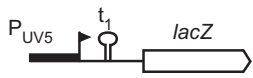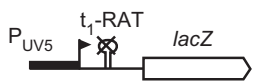

Fig. 5. Rate of termination and anti-termination at t_1. A *lacUV5* promoter, t_1, *lacZ* fusion was integrated into the chromosomal *attB*-site of strain S541 ($\Delta bgl\Delta lacZ$) yielding strain S1697 (= S541 attB::p-KESK35), and the expression of the *lacZ* fusion was determined of cells grown in minimal M9 glycerol medium supplemented with casaminoacids and B1 as described (Dole et al., 2002). BglG was provided in trans using plasmid pKESK18, and *bglG* expression was induced with 1 mM IPTG (Dole et al., 2002). As a control a *lacUV5* promoter, t_1-RAT, *lacZ* fusion was used, in which terminator t_1 was mutated by stabilization of the BglG binding motif RAT (Nagarajavel et al., 2007). This *lacZ* fusion was likewise integrated into the chromosomal *attB*-site yielding strain S1704 (= S541 attB::pKESK47). β-Galactosidase activities were determined in at least three independent assays and standard deviations were less than 10%.

The experiment gives no information about the threshold value R_2^{th}, which was therefore set manually.

2.4. Final model

Summarizing all influences, the transcription rate R_1 in the wild type with upstream and downstream silencers can now be written as

$$R_1 = \begin{cases} c_{ure}(aR_{1,2} + b)c_{syn}\alpha & \text{for } R_{1,2} \leqslant R_{1,2}^*, \\ c_{ure}c_{syn}\alpha & \text{for } R_{1,2} > R_{1,2}^*. \end{cases} \quad (8)$$

R_2 is given by Eq. (7). Inserting the estimated values reads

$$R_2 = \begin{cases} 0.3R_1 & \text{for } R_2 \leqslant R_2^{th}, \\ 0.9R_1 & \text{for } R_2 > R_2^{th}. \end{cases} \quad (9)$$

Model A: In model A, transcription rate R_1, in the following denoted by R_1^A, is independent of transcription rate R_2. This can also be seen when inserting $R_{1,2} = R_1$ and $R_{1,2}^* = R_1^*$ in Eq. (8). Thus, we can directly resolve Eq. (8) for R_1^A, which is given by

$$R_1^A(\alpha) = \begin{cases} \dfrac{c_{ure}c_{syn}b\alpha}{1 - c_{ure}c_{syn}a\alpha} & \text{for } R_1^A \leqslant R_1^*, \\ c_{ure}c_{syn}\alpha & \text{for } R_1^A > R_1^*. \end{cases} \quad (10)$$

Inserting all estimated values, the transcription rate R_1^A reads

$$R_1^A(\alpha) = \begin{cases} \dfrac{1.8 \times 10^{-2}\alpha}{1 - 5.4 \times 10^{-5}\alpha} & \text{for } R_1^A \leqslant 3000, \\ 0.18\alpha & \text{for } R_1^A > 3000. \end{cases} \quad (11)$$

Model B: In model B (Fig. 2B) $R_{1,2} = R_2$ and $R_{1,2}^* = R_2^*$ and the two feedback loops are interlocked. In order to resolve for the transcription rate R_1, which we denote by R_1^B, in this case, we insert Eq. (7) into (8), which leads to

$$R_1^B = \begin{cases} c_{ure}(ac_{t_1}^{min}R_1^B + b)c_{syn}\alpha & \text{for } R_2 \leqslant R_2^* \text{ and } R_2 \leqslant R_2^{th}, \\ c_{ure}(ac_{t_1}^{max}R_1^B + b)c_{syn}\alpha & \text{for } R_2 \leqslant R_2^* \text{ and } R_2 > R_2^{th}, \\ c_{ure}c_{syn}\alpha & \text{for } R_2 > R_2^*. \end{cases} \quad (12)$$

This can be resolved for R_1^B:

$$R_1^B(\alpha) = \begin{cases} \dfrac{c_{ure}c_{syn}b\alpha}{1 - c_{ure}c_{syn}ac_{t_1}^{min}\alpha} & \text{for } R_2^B \leqslant R_2^* \text{ and } R_2^B \leqslant R_2^{th}, \\ \dfrac{c_{ure}c_{syn}b\alpha}{1 - c_{ure}c_{syn}ac_{t_1}^{max}\alpha} & \text{for } R_2^B \leqslant R_2^* \text{ and } R_2^B > R_2^{th}, \\ c_{ure}c_{syn}\alpha & \text{for } R_2^B > R_2^*. \end{cases} \quad (13)$$

Inserting the estimated values, transcription rate R_1^B finally reads

$$R_1^B(\alpha) = \begin{cases} \dfrac{1.8 \times 10^{-2}\alpha}{1 - 1.6 \times 10^{-5}\alpha} & \text{for } R_2^B \leqslant 3000 \text{ and } R_2^B \leqslant R_2^{th}, \\ \dfrac{1.8 \times 10^{-2}\alpha}{1 - 4.9 \times 10^{-5}\alpha} & \text{for } R_2^B \leqslant 3000 \text{ and } R_2^B > R_2^{th}, \\ 0.18\alpha & \text{for } R_2^B > 3000. \end{cases}$$

(14)

3. Results

3.1. Simulation of the expression rate R_2 with increasing promoter activity α

To analyze the role of the positive regulatory feedback loops in the exceptional specificity of repression by H-NS, we simulated the behavior of the transcription rate R_2 with increasing promoter activity α (Fig. 6). First, we analyzed the behavior of R_2 when including repression by H-NS through the URE and DRE, but when omitting the regulatory feedback loop based on termination/anti-termination. In this case, transcription rates R_1 and R_2 are equal, and thus the curves of models A and B are identical. R_2 non-linearly increases with increasing α (Fig. 6a). This non-

linearity is caused by the feedback loop based on the inverse correlation of the transcription rate and of repression by H-NS through the DRE. Above the upper limit $R_{1,2}^* = 3000$, repression by H-NS through the DRE is relieved, and the relation between R_2 and α is linear (Fig. 6a).

Next, the regulation based on termination/anti-termination was included in the simulation of R_2 with increasing α (Figs. 6b–d). In this case, according to Eq. (7), the ratio of R_2 and R_1 is either 0.3 (for $R_2 \leqslant R_2^{th}$) or 0.9 (for $R_2 > R_2^{th}$). We analyzed a range of threshold values R_2^{th} for anti-termination by BglG, and results are shown for $R_2^{th} = 100$, 200, and 500, respectively (Fig. 6b–d). Here, models A and B (Fig. 2A and B, respectively) show different behaviors in the range of promoter activities in which the DRE is active. In case of model B, the simulation of R_2 with increasing promoter activity α shows a low transcription rate R_2 over a wide range of the promoter activity α, followed by a very rapid increase of R_2 when R_2 exceeds the threshold R_2^{th} for anti-termination (dashed lines in Fig. 6b–d). Further, with $R_2^{th} = 200$ or 500, R_2, when exceeding the threshold R_2^{th}, simultaneously exceeds the upper limit for repression R_2^*. In case of model A, in which the two feedback loops are not interconnected, the simulation shows that R_2 increases more gradually, although within a narrow range of α (Fig. 6b–d). The course of R_2 is identical in both models for large promoter
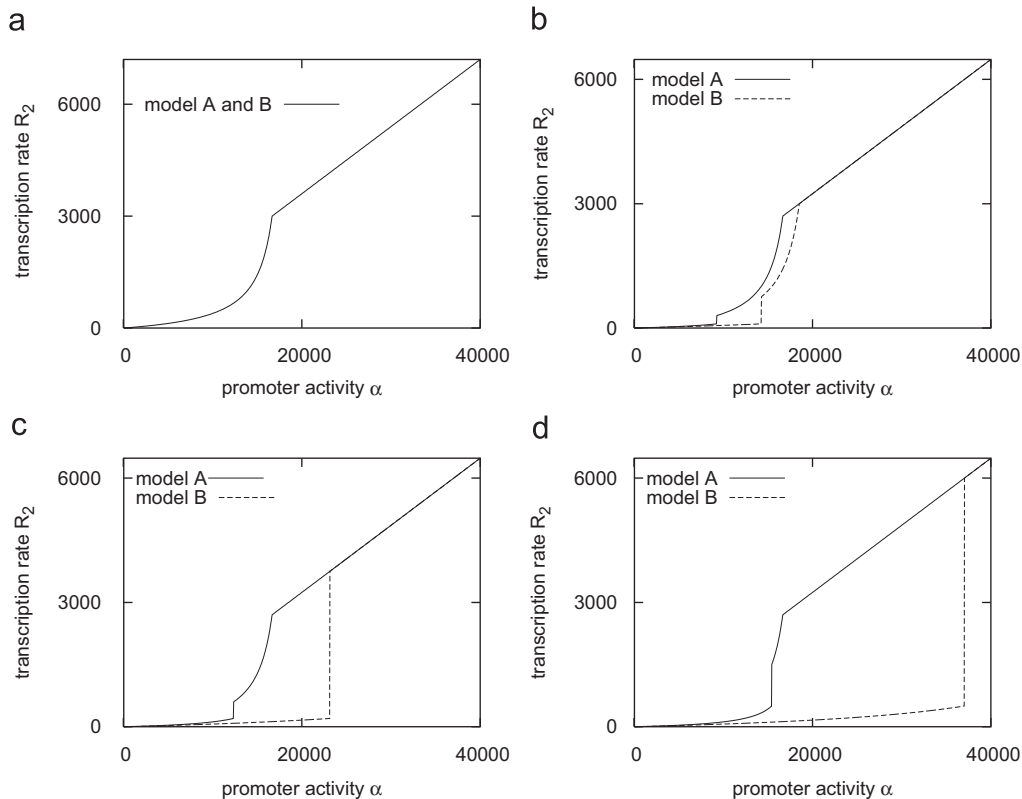

Fig. 6. Prediction of the transcription rate R_2 for increasing promoter activity α. Shown are simulations using model A (*solid line*) and model B (*dashed line*) without regulation by termination/anti-termination (a) and with regulation by termination/anti-termination using different threshold values (b)–(d). The non-linear increase of the transcription rate R_2 for low promoter activities α is caused by the feedback loop that is based on the inverse correlation of transcription and repression by H-NS through the DRE. R_2 increases linearly above a transcription rate higher than the upper limit for repression through the DRE, i.e. when $R_{1,2} > R_{1,2}^* = 3000$. (a) Without terminator, (b) $R_2^{th} = 100$, (c) $R_2^{th} = 200$, and (d) $R_2^{th} = 500$.

activities α, i.e. when the feedback loop including DRE is inoperative.

Further, the threshold regulation of termination/anti-termination causes the model to show hysteresis. This is illustrated in Fig. 7 for model B and a threshold value $R_2^{th} = 100$. In a small range of promoter activities, there exist
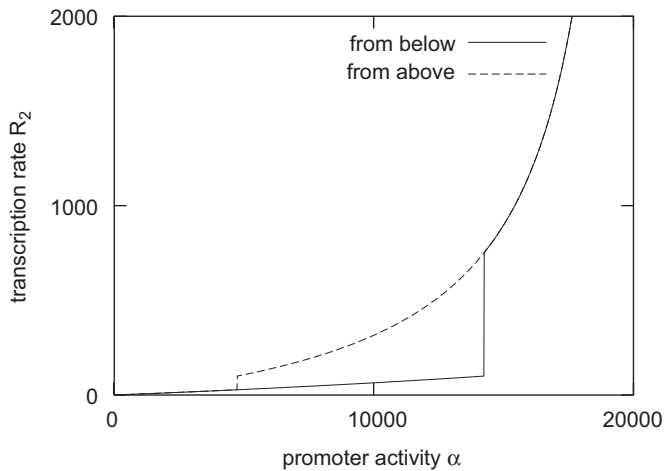

Fig. 7. The model shows hysteresis. Change in transcription rate R_2 for increasing (*black dotted line*) and decreasing (*gray solid line*) promoter activity α. In a small range of promoter activities, the system has two stable steady states and shows hysteresis. The plot is shown for a wild-type simulation with model B and threshold value $R_2^{th} = 100$.

two solutions for the two transcription rates R_1 and R_2. Starting with a small initial promoter activity α and increasing α, R_2 escalates at a promoter activity α^*, when the threshold R_2^{th} is reached. Coming, however, from above, R_2 remains high beyond this α^* and only drops to the smaller solution at a promoter activity $\alpha^{**} < \alpha^*$. This is a typical phenomenon observed in regulatory networks with positive non-linear feedback loops (Gouzé, 1998; Thomas, 1981).

3.2. Simulation of the fold increase of the expression rate R_2 upon a 3-fold increase of the promoter activity α

According to experimental data, a moderate 3-fold increase in the *bgl* promoter activity is sufficient to overcome repression by H-NS and to lead to an approximately 100-fold increase in the expression (Dole et al., 2002; Schnetz, 2002). To test whether our model is able to explain this phenomenon for a promoter activity α, we calculated the ratio of the transcription rates $R_2(3\alpha)$ and $R_2(\alpha)$. The results for $R_2(3\alpha)/R_2(\alpha)$ for increasing α are shown in Fig. 8. The result obtained when omitting the regulatory feedback loop based on termination/anti-termination is shown in Fig. 8a. In this case, the ratio $R_2(3\alpha)/R_2(\alpha)$ non-linearly increases and reaches a maximal ratio that is about 20. At this point $R_{1,2}(3\alpha) = 3000$, which is the upper limit for repression by H-NS through the DRE. Then, the ratio linearly decreases, since the positive
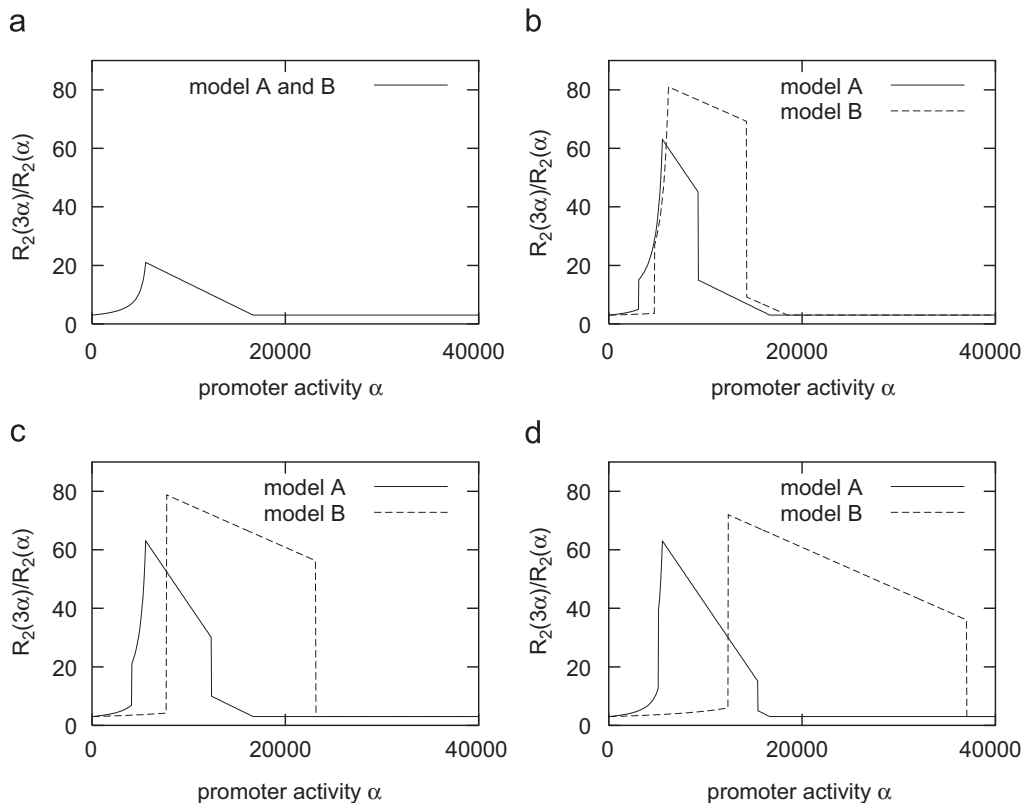

Fig. 8. Specific repression predicted by the model. Ratio of transcription rates $R_2(3\alpha)$ and $R_2(\alpha)$ as a function of increasing promoter activity α. Shown are results without regulation by termination/anti-termination (a) and with regulation by termination/anti-termination using different threshold values R_2^{th} (b)–(d). Solid lines refer to model A, dashed lines to model B. (a) Without terminator, (b) $R_2^{th} = 100$, (c) $R_2^{th} = 200$, and (d) $R_2^{th} = 500$.

feedback loop based on repression through the DRE remains only active in $R_2(\alpha)$, but not in $R_2(3\alpha)$. When $R_2(\alpha)$ also reaches this upper limit for repression, both transcription rates increase linearly with α, and thus the ratio is at a constant value 3.

The results when including the feedback loop based on termination/anti-termination are shown in Fig. 8b–d. Again simulations are shown for threshold values $R_2^{th} = 100, 200$, and 500, respectively. In case of model A, in which repression by H-NS through the DRE inversely correlates to R_1, the ratio increases over a narrow range of α to maximally 60 (solid lines in Fig. 8b–d). At this point $R_2(3\alpha)$ exceeds the upper limit $R_1^* = 3000$ for repression by H-NS. Then the ratio linearly decreases until $R_2(\alpha)$ reaches the threshold for anti-termination, where the ratio drops, and further decreases linearly. When $R_2(\alpha)$ exceeds the upper limit for repression by H-NS, the ratio is at the constant value 3 (solid lines in Fig. 8b–d).

The results for model B, in which the two feedback loops are interconnected, show some significant differences (dashed lines in Fig. 8b–d). Here, for all thresholds R_2^{th}, the ratio increases only marginally for low promoter activities. When $R_2(3\alpha)$ reaches the threshold value R_2^{th}, the ratio $R_2(3\alpha)/R_2(\alpha)$ rapidly increases and instantly reaches a maximal value that is approximately 80 (dashed lines in Fig. 8b–d). At this point $R_2(3\alpha) \geqslant R_2^* = 3000$. Then the ratio of $R_2(3\alpha)$ to $R_2(\alpha)$ is high over a wide range of α. When $R_2(\alpha)$ also reaches the threshold R_2^{th} the ratio drops, and then decreases linearly to the constant value 3.

Qualitatively similar results were obtained for all threshold values R_2^{th}. However, the threshold R_2^{th} for anti-termination affects the range of promoter activities α, for which a 3-fold increase (to 3α) causes high levels of expression (R_2). In addition, the induction ratio, which is 80 for $R_2^{th} = 100$, decreases somewhat with higher threshold values. Also the shown changes in the ratio of $R_2(3\alpha)$ to $R_2(\alpha)$ are only observed when the threshold R_2^{th} for antitermination and the upper limit $R_{1,2}^*$ for repression by binding of H-NS to the DRE are exceeded in the following order: first, transcription rate $R_2(3\alpha)$ exceeds R_2^{th}. Second, $R_{1,2}(3\alpha)$ reaches the upper limit $R_{1,2}^*$, such that silencing through the DRE is relieved. In this range, the ratio $R_2(3\alpha)/R_2(\alpha)$ is exceptionally high (Fig. 8b–d). The ratio $R_2(3\alpha)/R_2(\alpha)$ decreases immediately when also transcription rate $R_2(\alpha)$ exceeds the threshold R_2^{th}, and it is constant when $R_{1,2}(\alpha)$ reaches the upper limit $R_{1,2}^*$, i.e. when the DRE is inoperative for both transcription rates. This order defines limits for the range of R_2^{th}. A lower limit $R_{2,min}^{th}$ of R_2^{th} is defined by simultaneously exceeding the upper limit $R_{1,2}^*$ by transcription rate $R_2(3\alpha)$ and the threshold value R_2^{th} by transcription rate $R_2(\alpha)$, and is approximately $R_{2,min}^{th} = 43$ for model A and $R_{2,min}^{th} = 37$ for model B. Similarly, synergistic enhancement by the two feedback loops is also not observed if the threshold R_2^{th} for anti-termination is larger than the upper limit $R_{1,2}^*$ of repression through the DRE.

4. Discussion

We have introduced a model for the repression of the *E. coli bgl* operon by the protein H-NS. Relatively simple, i.e. constant or linear, functions were used to parameterize our model, which enables a parameter estimation with only few experimental data. In spite of its simplicity, the model is able to explain the exceptional specificity of repression by H-NS, which is reflected in a non-linear relation between the promoter activity and the expression rate of the *bgl* operon. In particular, the model shows that a 3-fold variation in the promoter activity α can result in a 60–80-fold change in expression. The analysis of the model indicates that this phenomenon is caused by the interplay of two positive feedback loops. One feedback loop includes repression of transcription initiation by binding of H-NS to the downstream silencer. Efficiency of this repression decreases with increasing promoter activity (Nagarajavel et al., 2007). The second loop is based on the positive auto-regulation of *bglG* by anti-termination, and on the limitation of BglG at low expression rates (Dole et al., 2002).

Two simplified regulatory networks, models A and B (Fig. 2), were deduced from the experimental data. In model A it was assumed that repression by H-NS when binding to the DRE inversely correlates to the rate of productive transcription initiation. In model B it was assumed that repression by H-NS through the DRE inversely correlates to the rate of transcription elongation across the DRE to which H-NS binds. In this model, the two regulatory feedback loops are interconnected. Simulations using these models revealed some interesting differences. In particular, with model B, a higher rate of induction was observed than with model A. In addition, model B demonstrates a rapid switch from the repressed state to a plateau with an approximately 80-fold increase in expression, while, with model A, the expression rate gradually changes, although in a narrow range. Comparing models A and B, model B reflects experimental data more closely. This predicts that repression by H-NS by binding to the DRE inversely correlates to the transcription elongation rate across the DRE (model B), rather than to the rate of productive transcription initiation (model A), suggesting that RNA polymerase engaged in elongation can disrupt the repressing nucleoprotein complex formed by H-NS.

The model could be further improved. Presently, the modeling is simplified for repression by binding of H-NS to the upstream silencer and for synergy of repression through both silencers. Both were modeled with constant repression factors c_{ure} and c_{syn} independent of the promoter activity α and independent of the rate of repression through the downstream silencer. A more realistic description of c_{ure} could be a decreasing function with respect to the promoter activity α. The same holds for the repression factor c_{syn}, which actually depends on the activity of both silencers. In addition, R_2 was taken as a measure for BglG activity.

However, the BglG activity is in addition regulated by phosphorylation depending on the availability of the specific sugar substrate and other sugars (Deutscher et al., 2006). An inclusion of these parameters into our model could further increase the ratio $R_2(3\alpha)/R_2(\alpha)$.

The protein H-NS is abundant nucleoid-associated protein that functions as a global repressor in *E. coli* and other enterobacteria. However, up to date it has remained a puzzle how highly specific regulation of some H-NS repressed loci is achieved. The results presented here suggest that synergy of repression by binding of H-NS to two regulatory elements and presumably remodeling of the H-NS–DNA nucleoprotein complex by RNA polymerase may be important to enhance specificity. The specificity of repression by H-NS may be further amplified by loci-specific additional regulatory feedback loops, as shown here for *bgl*.

Another conclusion from this work is that general model approaches that restrict their variables to a constraint set of components like mRNA or protein concentrations and use general parameterized functions will fail in cases where lots of different components contribute to the behavior of a system. In our study, these components are of very different type and their influences are modeled directly from the experiments with no predefined set of functions as a basis. Nonetheless, this modeling approach allowed to explain the observed phenomenon in spite of only few experiments with a relatively simple model.

References

Aymerich, S., Steinmetz, M., 1992. Proc. Natl Acad. Sci. 89, 10410.
Bouffartigues, E., Buckle, M., Badaut, C., Travers, A., Rimsky, S., 2007. Nat Struct. Mol. Biol. 14 (5), 441.
Csonka, L.N., 1982. J. Bacteriol. 151, 1433.
Csonka, L.N., Epstein, W., 1996. In: Neidhardt, F., Curtiss, R., Ingraham, J., Lin, E., Low, K., Magasanik, B., Reznikoff, W., Riley, M., Schaechter, M., Umbarger, H.E. (Eds.), *Escherichia coli* and *Salmonella*. Cellular and Molecular Biology, Second ed. ASM Press, Washington, DC, pp. 1210–1223.
Dame, R.T., Noom, M.C., Wuite, G.J.L., 2006. Nature 444 (7117), 387.
Dattananda, C.S., Rajkumari, K., Gowrishankar, J., 1991. J. Bacteriol. 173, 7481.
Deutscher, J., Francke, C., Postma, P.W., 2006. Microbiol. Mol. Biol. Rev. 70, 939.
Dole, S., Kühn, S., Schnetz, K., 2002. Mol. Microbiol. 43 (1), 217.
Dole, S., Nagarajavel, V., Schnetz, K., 2004. Mol. Microbiol. 52 (2), 589.
Dorman, C.J., 2004. Nat Rev. Microbiol. 2, 391.
Dorman, C.J., 2006. Nat. Rev. Microbiol. 5, 157.
Gebert, J., Radde, N., 2006. In: 7th International Conference on Computing Anticipatory Systems (CASYS05), AIP Conference Proceedings, Liege, Belgium: pp. 526–533.
Görke, B., 2003. J. Biol. Chem. 278, 46219.
Gouzé, J.L., 1998. J. Biol. Syst. 6 (21), 11.
Higgins, C.F., Dorman, C.J., Stirling, D.A., Waddell, L., Booth, I.R., May, G., Bremer, E., 1988. Cell 52, 569.
Jordi, B.J.A.M., Higgins, C.F., 2000. J Biol. Chem. 275, 12123.
Lucchini, S., Rowley, G., Goldberg, M.D., Hurd, D., Harrison, M., Hinton, J.C.D., 2006. PLoS Pathog. 2, e81.
Mahadevan, S., Wright, A., 1987. Cell 50, 485.
Miller, J.H., 1992. A Short Course in Bacterial Genetics. Cold Spring Harbor Laboratory Press.
Nagarajavel, V., Madhusudan, S., Dole, S., Rahmouni, A.R., Schnetz, K., 2007. J. Biol. Chem. 282 (32), 23622–23630.
Navarre, W.W., Porwollik, S., Wang, Y., McClelland, M., Rosen, H., Libby, S.J., Fang, F.C., 2006. Science 313 (5784), doi:10.1126/science1128794.
Overdier, D.G., Csonka, L.N., 1992. Proc. Natl Acad. Sci. 89, 3140.
Owen-Hughes, T.A., Pavitt, G.D., Santos, D.S., Sidebotham, J.M., Hulton, C.S.J., Hinton, J.C.D., Higgins, C.F., 1992. Cell 71, 255.
Prasad, I., Schaefler, S., 1974. J. Bacteriol. 120, 638.
Radde, N., Gebert, J., Forst, C.V., 2006. Bioinformatics 22 (21), 2674.
Roeder, I., Glauche, I., 2006. J. Theor. Biol. 241, 852, doi:10.1016/j.jtbi.2006.01.021.
Schaefler, S., Maas, W.K., 1967. J. Bacteriol. 93, 264.
Schnetz, K., 1995. EMBO J. 14, 2545.
Schnetz, K., 2002. Microbiology 148, 2573.
Schnetz, K., Rak, B., 1988. EMBO J. 7, 3271.
Schnetz, K., Toloczyki, C., Rak, B., 1987. J. Bacteriol 169, 2579.
Thomas, R., 1981. In: Della-Dora, J., Demongeot, J., Lacolle, B., (Eds.), Numerical Methods in the Study of Critical Phenomena Springer Series in Synergetics, vol. 9. Springer, Berlin, pp. 180–193.
Yildirim, N., Mackey, M.C., 2003. Biophys. J. 84, 2841.