Noise in transcription negative feedback loops: simulation and experimental analysis

Yann Dublanche1, Konstantinos Michalodimitrakis1, Nico Kümmerer, Mathilde Foglierini and Luis Serrano*

European Molecular Biology Laboratory (EMBL), Heidelberg, Germany
* Corresponding author. European Molecular Biology Laboratory (EMBL), SCB, Meyerhofstrasse 1, 69117 Heildelberg, Germany.
Tel.: + 49 6 221387320; Fax: + 49 6 221387306; E-mail: serrano@embl.de
1 These authors contributed equally to this work

Received 11.1.06; accepted 14.6.06

Negative feedback loops have been invoked as a way to control and decrease transcriptional noise. Here, we have built three circuits to test the effect of negative feedback loops on transcriptional noise of an autoregulated gene encoding a transcription factor (TF) and a downstream gene (DG), regulated by this TF. Experimental analysis shows that self-repression decreases noise compared to expression from a non-regulated promoter. Interestingly enough, we find that noise minimization by negative feedback loop is optimal within a range of repression strength. Repression values outside this range result in noise increase producing a U-shaped behaviour. This behaviour is the result of external noise probably arising from plasmid fluctuations as shown by simulation of the network. Regarding the target gene of a self-repressed TF (sTF), we find a strong decrease of noise when repression by the sTF is strong and a higher degree of noise anti-correlation between sTF and its target. Simulations of the circuits indicate that the main source of noise in these circuits could come from plasmid variation and therefore that negative feedback loops play an important role in suppressing both external and internal noise. An important observation is that DG expression without negative feedback exhibits bimodality at intermediate TF repression values. This bimodal behaviour seems to be the result of external noise as it can only be found in those simulations that include plasmid variation.

Molecular Systems Biology 1 August 2006; doi:10.1038/msb4100081
Subject Categories: synthetic biology; simulation and data analysis
Keywords: autoregulation; negative feedback loop; noise; SmartCell; TetR

Introduction

The genetic programme of a cell and/or an organism is determined by a complex web of gene and protein networks. Taking into account the small number of some of the components in the networks (especially in transcription processes), it is surprising that in general cells and organisms manage to carry out this programme in a reproducible way. How the cell copes with the noise, or even in some cases uses it for its advantage (Chen et al., 2005), is a fascinating topic that has prompted many groups to analyse it theoretically (Kepler and Elston, 2001; Swain et al, 2002; Paulsson, 2004; Austin et al, 2006) and experimentally (Becskei and Serrano, 2000; Elowitz et al, 2002; Ozbudak et al, 2002; Swain et al, 2002; Blake et al, 2003; Raser and O’Shea, 2004; Chen et al, 2005; Hooshangi et al, 2005; Pedraza and van Oudenaarden, 2005; Austin et al, 2006). In principle, there are two sources of noise that could affect transcriptional networks: intrinsic and extrinsic (Swain et al, 2002; Paulsson, 2004). Although the definition of both is somehow relative (Paulsson, 2004), we could in principle define intrinsic noise as that arising directly from the circuit and external as that due to changes in the surrounding environment. For example, intrinsic noise could be due to the randomness in the binding of the polymerase to the promoter of one of the network genes and external could be due to fluctuations in the activity of the polymerase owing to changes in the metabolic status of the cell. Elegant experiments have been carried out by different groups to try to separate one type of noise from the other (Swain et al, 2002), to determine the sources of noise (Ozbudak et al, 2002; Blake et al, 2003), as well as to study cascade noise amplification in a network (Hooshangi et al, 2005; Pedraza and van Oudenaarden, 2005). One way in which cells could suppress noise is by using negative feedback loops (Savageau, 1974). Experimental study of a designed negative feedback loop in Escherichia coli showed how introduction of negative feedback decreased noise level (Becskei and Serrano, 2000; Austin et al, 2006), explaining its high frequency in prokaryotic organisms (Thieffry et al, 1998).

Recently, Paulsson (2004) published a thorough theoretical analysis of noise in gene networks and presented an equation that decomposes the intrinsic and extrinsic noise contributions, simplifying its analysis. Based on this equation, the author analysed the experimental negative feedback circuit described by Becskei and Serrano (2000). The results suggested that the noise in the circuit analysed by the authors
should mainly come from fluctuations in plasmid numbers and the suppression of the noise by the negative feedback loop should be mainly owing to elimination of the fluctuations introduced by changes in the plasmid number. This would suggest that negative feedback loops do not suppress intrinsic noise, but rather eliminate external noise that might arise for example from plasmid variation, or other external sources of noise like ribosome variation (Austin et al., 2006). In a recent work, it has also been shown that negative feedback shifts noise to higher frequencies characteristic of intrinsic noise, while filtering external noise (Austin et al., 2006).

To see if this is the case, we have reproduced the original experiment of Becskei and Serrano (2000) using a more sensitive approach to quantify GFP expression (FACS sorting versus microscopy used in the previous work). Also we have created new constructs in which the reporter protein (GFP) is no more fused to the Tet repressor (TetR) but rather regulated by TetR. In this way, we can analyse the effect of noise suppression of a negative feedback loop on a downstream gene and compare it with that found in a gene transcription cascade with no feedback (Hooshangi et al., 2005; Pedraza and van Oudenaarden, 2005). Three constructions have been analysed: (a) TetR fused to GFP and repressing itself (TG-nf); (b) TetR repressing itself and GFP, with the latter being expressed from a different promoter (T + G-nf); and (c) TetR expressed by a constitutive promoter and repressing GFP (T + G) (Figure 1).

To reproduce the normal situation in a prokaryotic cell, we have used a low-copy plasmid for TetR (around four copies) and a medium-copy plasmid (around 60 copies) for the GFP reporter. In this way, we reproduce the fact that many negatively regulated transcription factors in E. coli regulate several downstream genes. In parallel, we have carried out a detailed stochastic simulation of the networks taking into account cell division, plasmid variation and competition of other promoters for the polymerase. Simulations have been carried out with the SmartCell software (http://smartcell.embl.de/) (Ander et al., 2004) that uses the Gillespie (1977) algorithm as modified by Gibson and Bruck (2000) and Stundzia and Lumsden (1996) to include space and diffusion. This new version of SmartCell uses the modification of Elf and co-workers (Hattne et al., 2005) to accelerate calculations (http://smartcell.embl.de/).

**Results**

**Experimental analysis of GFP expression in the different circuits**

The three different circuits shown in Figure 1 were analysed under different concentrations of anhydrotetracycline hydrochloride (aTc) ranging from 0 to 100 ng/ml. This range was chosen on the following basis: the amount of TetR produced by the constitutive promoter is smaller than that produced in DH5αZ1 cells and 100 ng/ml of aTc has been shown to fully titrate TetR in the latter case (Lutz and Bujard, 1997). Specifically, the amount of TetR produced by the constitutive promoter was estimated by Western blotting (as explained in Materials and methods) to be approximately 2500 molecules when using DH5αZ1 cells as reference, which have approxi-
responsive to the increase of aTc. In the case of GFP under the control of TetR, we find similar final amounts of fluorescence for the case in which TetR represses itself and in that in which TetR is not involved in a negative feedback loop (Figures 2D and 3A; it should be noted that the two figures correspond to different experiments). This appears to be in qualitative agreement with the simulations (Figure 3B). However, the response to aTc is very different in the two cases (Figure 2D). In the negative feedback loop circuit, we observe a very homogeneous distribution of GFP in the cell population that gradually moves from a low fluorescence value to higher ones with increasing aTc concentration. In the other case, we see a complicated behaviour in which at intermediate aTc values fluorescence of cells spans several orders of magnitude, whereas some cells seem not to express GFP at all. Only at high and very low aTc values, we see a single population. Another interesting feature is the ‘ease’ with which the two circuits reach the state corresponding to saturation: although the negative feedback circuit starts with a barely detectable level of TetR, it requires a higher dose of aTc to reach saturation, compared to the non-regulated circuit, which reaches GFP saturation very quickly. This can be explained when considering the fact that in the case of the negative feedback loop, addition of aTc could be subjected to a type of ‘buffering’ effect: aTc can relieve temporarily autorepression, leading to higher expression of TetR until TetR levels are such...

Figure 2  FACS analysis and simulation-derived histograms of the three circuits under different aTc concentrations. (A–C) Simulation-derived histograms for the circuits for different numbers of aTc molecules with constant plasmid copy number (A), variable polymerase levels (B) and variable plasmid copy number (C). In all three cases, the values for circuit T+G for 750, 1000, 1500 and 2500 aTc molecules have been plotted on a secondary Y-axis for clarity and the respective legend entries are at the bottom of the legend to separate them from the others. (D) FACS analysis of the three circuits for different aTc concentrations (in ng/ml). Cells with no plasmids (Top10) were used as a control for the level of autofluorescence.
that the loop can be restored in the presence of the given aTc concentration.

Noise analysis in the experimental circuits

To compare the level of noise in the three circuits, we used the \( V_c \) (coefficient of variance determined by dividing the standard deviation by the mean; see Materials and methods) values. This value gives an indication of the spread of the population with respect to the mean. The higher the \( V_c \) value is, the more noisy the system. In Figure 4D, we show the plots of the change in \( V_c \) with respect to aTc concentration for the three circuits for four different experiments; three were conducted for 2.5 h and one for 3 h. Comparing the results obtained with the three circuits, the first obvious conclusion is that all of them arrive to similar \( V_c \) values (around 0.5 ± 0.1) at high aTc concentrations where there is no repression by TetR. Second, we can observe that the \( V_c \) values are very high for the T + G circuit (Figure 4D, right) at low aTc concentrations where TetR still represses the GFP promoter. For the negative feedback (T + G-nf), we find intermediate values at low aTc values (Figure 4D, centre). In the case of the TG-nf circuit (Figure 4D, left), we find that noise is slightly larger at very low aTc concentration, then decreases when increasing aTc and finally goes up again at high aTc concentration (the large \( V_c \) value at 0 aTc could be partly due to intrinsic autofluorescence of the E. coli cells; see below). This behaviour is different from previous published work (Becskei and Serrano, 2000), as it was indicated that the noise of the TG-nf circuit simply increased with the amount of aTc. Here, we have found that at very low aTc, the noise is as high as at very high aTc and it decreases at slightly higher aTc concentrations. This discrepancy can be explained if we consider that light microscopy was used to analyse the noise in the original paper and thus very low values of aTc could not be explored. Here, by using FACS sorting, we have been able to explore very low aTc values. Another reason for the discrepancy could lie in the sampling of cells for calculating \( V_c \) values. In the absence of aTc for the TG-nf circuit, the level of GFP expression is so low that the distribution greatly overlaps with the autofluorescence distribution of Top10 cells. Although this means that the \( V_c \) value will not represent accurately the levels of GFP, it will still reflect the profile of the entire population, as it is including even the non-expressing or very low-expressing cells.

Simulation analysis of the three circuits

To get a better understanding of the mechanisms behind the different behaviour of the three circuits with respect to GFP expression and \( V_c \) values, we have made a detailed model of the three circuits (see Materials and methods). In this model, we consider fluctuations owing to binding of the polymerase to competing promoters in the cell, the possibility of having more than one polymerase per gene/operon at the same time and binding of aTc to the free and bound TetR at the two binding sites (see Figure 5 for all reactions being considered). The values used for the simulation have been obtained from experimental data when available (Table I). As simulations covered time periods (1 day = 86 400 s) much longer than that of the E. coli cell cycle compared to the 30 min cell division time experimentally measured for the given growth conditions, it was necessary to limit the lifetime of the species used in the simulation in order to capture the real changes in their concentration owing to cell division. Thus, a pseudo-degradation rate was given to such components, so that they have an apparent half-life corresponding to the measured cell division time. The circuits have been simulated using the software SmartCell (Ander et al, 2004).

The first runs, using values previously published in literature for half-lives, promoter strengths, etc., resulted in too high values for the final concentrations of TetR and GFP. Thus, we slightly altered the degradation rates of TetR and GFP in order to have approximately the same final numbers of molecules as found in this study (similar results were obtained if instead we slightly modified the production rates). Regarding aTc concentration, it is very difficult to make equivalence between experimental values and the ones used for simulation. The reason is that in the case of the negative feedback loop, the aTc concentration inside an E. coli cell could be much higher than in the medium, due to a sink effect of TetR binding to aTc. Specifically, owing to the strong binding of aTc to TetR (Degenkolb et al, 1991), molecules of aTc that enter the cell will readily bind to TetR, resulting, on the one hand, in de-repression and production of more TetR and, on the other, depletion of free aTc molecules inside the cell. The latter would be the driving force for more aTc molecules to diffuse from the medium into the cell, increasing the total amount of aTc (free and bound) inside the cell compared to that of the medium.
Comparison between experimental and simulated data

We have run three different types of simulations (see Figure 5 for the reactions considered) to see the effect of external and internal noise. In the first case, we simulated the three circuits with a fixed number of plasmids and polymerase. In the second case, we allowed for variation in the number of polymerase molecules by changing degradation and production rates within reasonable values (see Materials and methods). Finally, we considered the possibility of variation in the number of plasmids per cell using the values described in the literature.

Comparison between the experimental and simulated values for the expression of GFP shows that the behaviour of TG-nf and T+G-nf is well captured by the simulation (Figure 2) independently of the conditions considered (polymerase variation, plasmid number variation or only internal noise). However, for the circuit without negative feedback (T+G), only in the case of plasmid variation (Figure 2C) we see a good fitting to the experimental data (Figure 2D), with a bimodal distribution at intermediate aTc concentrations.

In Figure 4, we show the $V_c$ values as calculated from the simulations of the three circuits (no external noise (Figure 4A), varying polymerase concentration (Figure 4B) and varying plasmid number (Figure 4C), the experimental data (Figure 4D)). In this case, we could see that only when considering plasmid variation we can observe a good fitting to the experimental data for the TG-nf circuit. Regarding the T+G-nf circuit, we could see a marginally better reproduction.
We also examined the importance of having a medium-copy plasmid with the reporter gene by repeating the simulations with low-copy plasmids and allowing for plasmid variation. Analysis of the relationship between $V_c$ and $\alpha_{TC}$ shows exactly the same behaviour as above (data not shown).

**Correlated noise and negative feedback loop**

To quantify the expression fluctuations and the degree of correlation between TetR and GFP genes in the T + G-nf and in the T + G circuits, we computed the correlation parameter $C_{ij}$ as described by Pedraza and van Oudenaarden (2005), using the computer simulation data considering plasmid variation (Figure 6A). In the case of the T + G circuit, we see a behaviour similar to that described by Pedraza and van Oudenaarden (2005) for genes 1 and 2 in their gene cascade. Essentially at low $\alpha_{TC}$, the noise in GFP expression does not correlate with the noise in TetR. At higher $\alpha_{TC}$ concentrations, there is a negative correlation between both protein products, so that noise in GFP expression decreases with the amount of active TetR product. Finally, at high $\alpha_{TC}$, we see again an uncoupling of the noise in both proteins. However, the behaviour for the T + G-nf circuit is different. In this case, we see a larger negative correlation from the beginning, indicating a much higher coupling between the two systems. At intermediate $\alpha_{TC}$ concentrations, this coupling is maximal and more than double than that seen in the T + G circuit.
In a recent paper, it was shown that negative feedback regulation in a gene results in a shift of the noise frequency towards high frequencies when compared to a non-regulated gene (Austin et al., 2006). To see if this was the case in our simulated circuits, we calculated the noise frequency (see Materials and methods). We did this for the TG-nf and T + G-nf
network at intermediate αTc concentrations. This bimodality could only be reproduced in our simulations when considering plasmid variation. Bimodal behaviour has been described in positive feedback loops (Becskei et al., 2001) and in switches (Gardner et al., 2000). However, to our knowledge, no bimodal behaviour has been attributed to the combination of a particular circuit (in this case repression by a transcription factor) and a source of external noise. Thus, our result suggests that noise contribution to circuit behaviour could be very important and should be taken into account. Also, this result supports the hypothesis that plasmid variation is one of the major sources of external noise in our circuits.

Noise behaviour and repression

We observed that for all cases the presence of a negative feedback loop decreases significantly the noise in the production of the GFP protein compared to the non-regulated circuit, at αTc concentrations that still allow repression. However for TetR, it places the optimum level of noise suppression at intermediate αTc concentrations. Under conditions where repression is too tight, the TetR noise level is similar to the totally unregulated circuit. A simpler explanation for this behaviour is the following: although the negative feedback loop suppresses noise, at low αTc concentrations the base noise level is higher under these conditions owing to low TetR-GFP fusion protein population. Thus, a random increase in plasmid number has a good chance of allowing one or more rounds of transcription before the TetR level goes high enough to repress these extra copies. Increasing the αTc concentration reduces the base noise level by increasing expression levels (as seen in the T + G circuit). However, at the same time it disrupts the feedback loop, leading to an increase in noise or decrease of the circuit’s noise-suppressing ability. The final result of these two opposing phenomena is a U-shaped behaviour. This means that noise is minimized at intermediate levels of repression. At higher levels of repression, noise increases owing to low protein population, whereas the negative feedback suppression of noise decreases at lower levels of repression.

Regarding the non-repressed circuit, we see an immediate decrease of noise the moment some αTc is added to the medium. This contradicts the behaviour observed by other groups for non-regulated networks. In the case of Pedraza and van Oudenaarden (2005), they observed a plateau followed by a decrease in noise at high IPTG concentrations. Other groups like Elowitz et al. (2002) or Hooshangi et al. (2005) observed a hump in the noise at medium IPTG (Elowitz et al., 2002) or αTc (Hooshangi et al., 2005) concentrations and smaller Vc values (maximum around 1, whereas we obtained values around 3–4). In the case of Hooshangi et al. (2005), this hump was not predicted to happen in the simulation of their circuit for a one-stage cascade circuit. In fact their simulation showed a very similar behaviour to the one observed in our work. At this point, it is difficult to determine the reason behind these differences. The fact that simulations using Gillespie-based approximations showed similar behaviours for one-stage cascade (Hooshangi et al., 2005) could suggest that discrepancies could be due to sensitivity problems in the low expression range of fluorescent proteins. In fact when looking at Figure 2 circuits at the αTc concentration where Vc of TetR was lower and at high αTc where repression does not take place anymore. The results of this analysis are shown in Figure 6B and C. As previously found, negative feedback shifts noise frequency to higher values when compared to a non-regulated gene in both cases. However, the frequency shift is larger for the T-G-nf circuit (Figure 6B) than for the T + G-nf (Figure 6C).

Discussion

In this work, we have analysed experimentally and computationally the noise in a negative feedback-regulated transcription factor, as well as the effect the negative feedback loop has on a gene repressed by the same transcription factor.

Noise and bimodality

One of the interesting results we have obtained is the existence of a bimodal behaviour in the expression of GFP in the T + G network at intermediate αTc concentrations. This bimodality could only be reproduced in our simulations when considering plasmid variation. Bimodal behaviour has been described in positive feedback loops (Becskei et al., 2001) and in switches (Gardner et al., 2000). However, to our knowledge, no bimodal behaviour has been attributed to the combination of a particular circuit (in this case repression by a transcription factor) and a source of external noise. Thus, our result suggests that noise contribution to circuit behaviour could be very important and should be taken into account. Also, this result supports the hypothesis that plasmid variation is one of the major sources of external noise in our circuits.

Noise behaviour and repression

We observed that for all cases the presence of a negative feedback loop decreases significantly the noise in the production of the GFP protein compared to the non-regulated circuit, at αTc concentrations that still allow repression. However for TetR, it places the optimum level of noise suppression at intermediate αTc concentrations. Under conditions where repression is too tight, the TetR noise level is similar to the totally unregulated circuit. A simpler explanation for this behaviour is the following: although the negative feedback loop suppresses noise, at low αTc concentrations the base noise level is higher under these conditions owing to low TetR-GFP fusion protein population. Thus, a random increase in plasmid number has a good chance of allowing one or more rounds of transcription before the TetR level goes high enough to repress these extra copies. Increasing the αTc concentration reduces the base noise level by increasing expression levels (as seen in the T + G circuit). However, at the same time it disrupts the feedback loop, leading to an increase in noise or decrease of the circuit’s noise-suppressing ability. The final result of these two opposing phenomena is a U-shaped behaviour. This means that noise is minimized at intermediate levels of repression. At higher levels of repression, noise increases owing to low protein population, whereas the negative feedback suppression of noise decreases at lower levels of repression.

Regarding the non-repressed circuit, we see an immediate decrease of noise the moment some αTc is added to the medium. This contradicts the behaviour observed by other groups for non-regulated networks. In the case of Pedraza and van Oudenaarden (2005), they observed a plateau followed by a decrease in noise at high IPTG concentrations. Other groups like Elowitz et al. (2002) or Hooshangi et al. (2005) observed a hump in the noise at medium IPTG (Elowitz et al., 2002) or αTc (Hooshangi et al., 2005) concentrations and smaller Vc values (maximum around 1, whereas we obtained values around 3–4). In the case of Hooshangi et al. (2005), this hump was not predicted to happen in the simulation of their circuit for a one-stage cascade circuit. In fact their simulation showed a very similar behaviour to the one observed in our work. At this point, it is difficult to determine the reason behind these differences. The fact that simulations using Gillespie-based approximations showed similar behaviours for one-stage cascade (Hooshangi et al., 2005) could suggest that discrepancies could be due to sensitivity problems in the low expression range of fluorescent proteins. In fact when looking at Figure 2
it is clear that for aTc in the TG-nf and T + G circuits, there is a significant overlap between cell autofluorescence and GFP expression. One could argue that a correction for this factor should be introduced so that Vc reflects only the GFP-expressing cells. However, we would argue that as long as the overlapping is not 100%, no correction should be made. Otherwise, we will eliminate the noise component due to cells that do not express GFP when others do, as seen in the bimodal distribution for the T + G circuit. Thus, as long as the overlapping between autofluorescent cells and cells induced with aTc is not 100%, we should consider the Vc values. In any case, even when eliminating the first two points for the TG-nf circuit, we will still have a U-shaped behaviour.

**Negative feedback loop and transfer of information**

Calculation of the cross-correlation parameter Cxy between the TetR and the downstream GFP-regulated gene shows a stronger negative coupling between TetR levels and GFP levels when there is a negative feedback loop (around four times). This means that a negative feedback loop not only decreases the noise of the transcription factor auto-repressing itself, but also increases significantly the transfer of information to the downstream gene, which ultimately will be the desired biological trait of any circuit.

In a recent work (Simpson et al., 2003; Austin et al., 2006), it has been proposed that negative feedback loops shift noise frequency from low to high values when compared with non-regulated circuits. The authors proposed that this frequency shift may have biological relevance, as higher-frequency noise is more easily filtered out by downstream gene circuits in a regulatory cascade, and therefore has little regulatory impact. Analysis of noise frequency on our simulation data considering external source coming from plasmid variation shows this kind of behaviour for GFP when comparing the T + G-nf circuit at high aTc concentration (no regulation) and at medium aTc concentration (strong self-regulation and low Vc) (Figure 6C). Similar results are also found for the TG-nf circuit (Figure 6B), indicating that the shift in frequency takes place at the level of the self-repressed gene as well as its target. The only difference is that the shift is larger in the case of the TG-nf circuit.

**Noise origins**

Simulation of the three circuits, using a model of the transcription regulation in *E. coli*, as well as empirical parameters, shows a good agreement between experimental observations and the predicted behaviour when plasmid variation is taken into account. Thus, for the TG-nf circuit, we could really see the U-shaped behaviour for the noise. For the T + G-nf and T + G circuits, simulations show higher noise for the latter. Regarding GFP expression, we saw how simulations reproduced the experimental data showing some kind of bistability for the non-regulated T + G circuit and a smooth transition from low to high expression for the other two circuits. These results are independent of plasmid number, as similar results are obtained with medium and low copy numbers for the reporter gene (data not shown).

The advantage of the simulation is that it allows one to test different sources of noise in the system, which when compared with the experimental data allows to discriminate the more probable ones. Our results clearly indicate that as suggested by Paulsson (2004) plasmid variation upon cell division could be the most probable source of noise on the circuits analysed here. However, we cannot rule out other sources of external noise contributing as well. Polymerase variation could be one of those but our data indicate that at least within what one could expect to be the natural variation this will not be a likely cause. Other mechanisms like metabolic status, variation in ribosome numbers (Austin et al., 2006), access to the promoter, etc. could also contribute. In a recent paper (Austin et al., 2006) analysing noise in a negative feedback circuit, the authors found that aTc alone could have a strong influence on the noise of a non-regulated circuit. They proposed that binding of aTc to ribosomes and inhibition of translation could be an important external noise source. However, previous literature indicated that aTc has some toxic effects in *E. coli* at very high concentrations (>1 µg/ml and not ng/ml) through interference with the cell membrane but not through ribosome binding (Oliva et al., 1992). In any case, it seems that intrinsic noise of the system alone cannot explain the experimental observations and that external noise produced for example by plasmid variation is filtered by the negative feedback loop. Although it is not the case here, prokaryotic cells could use other mechanisms to diminish noise than negative feedback loops. One of them is DNA looping (Vilar and Saiz, 2005). In DNA looping, a repressor or activator binds to DNA at the promoter site and at the same time to a distal site on DNA. As a result, a repressor could release the binding site near the promoter allowing transcription while still bound to the distal site and this in turn allows faster re-binding and repression compared to a free repressor. The outcome then is equivalent to a faster koff for the repressor and therefore to noise decrease.

**Materials and methods**

**Experimental analysis**

**Materials**

Ampicillin, kanamycin and aTc were purchased from Sigma. For Western blotting, the primary antibody was a mixture of two monoclonal mouse antibodies for TetR from MoBiTec and the secondary antibody was peroxidase-conjugated donkey anti-mouse, from Jackson ImmunoResearch Laboratories. Bands were visualized by the ECL Western blotting analysis kit from Amersham Biosciences.

**Bacterial strain and plasmids**

The strain used for cloning and experiments is Top10 from Invitrogen (F<sup>–</sup> mcrA A(mrr-hsdRMS-mcrBC) φ80lacZAM15 ΔlacX74 deoR recA1 araD139 A(ara-leu)7697 galU galK rpsL (Str<sup>R</sup>) endA1 nupG). Only for the construction of pZEmG (see below), XL10-Gold cells, from Stratagene, were used. Plasmids pZE1-MCS-1, pZE12-Luc and pZS<sup>+</sup>24-MCS-1 (Lutz and Bujard, 1997) were used as starting material to construct vectors carrying the constructs. A monomeric form of GFP was cloned into pZE1-MCS-1 with KpnI–BamHI, resulting in plasmid pZEmG. This GFP variant is GFPmut3.1 with the A206K mutation (Chalfie et al., 1994). For Western blotting, the primary antibody was a mixture of two monoclonal mouse antibodies for TetR from MoBiTec and the secondary antibody was peroxidase-conjugated donkey anti-mouse, from Jackson ImmunoResearch Laboratories. Bands were visualized by the ECL Western blotting analysis kit from Amersham Biosciences.
the PLamO3 promoter of pZS14 MCS-1 by digestion with Aattl and KpnI, resulting in plasmid pZS*11 MCS-1. TetR from Tn10 was cloned into pZS*14 MCS-1 and pZS**11 MCS-1, yielding plasmids pZS*14T (non-regulated TetR expression) and pZS**11T (TetR negative feedback loop), respectively. A BamHI site was placed before the stop codon of the tetR gene. This site was used to introduce the gene of the monomeric GFP in pZS*11T, resulting in plasmid pZS*11TG, which expresses the fusion protein TetR-GFP.

Bacterial growth
Bacteria were grown at 37°C, in LB containing the appropriate antibiotics to ensure plasmid propagation, inoculating directly from glycerol stocks. The overnight cultures were used to inoculate 1:100 fresh LB medium containing antibiotics and the desired amount of acTc and grown for 2.5 or 3 h at 37°C. Comparison of the behaviour of the three circuits analysed at these two growth times did not show any significant difference, indicating that the system has reached steady state.

Flow cytometry analysis
One millilitre of cells, prepared as described above, was harvested and re-suspended in 2 ml of filtered PBS (0.22 μm filter). Cells without GFP were always used as a control to establish the level of autofluorescence. The analysis was performed on a DAKO MoFlo Flow Cytometer (DakoCytomation GmbH, Hamburger Strasse 181, 22083, Hamburg). The laser power used was 1.2 W of 488 nm light. This has an added benefit of near saturating the GFP on the sorter. Approximate measurement of fluorochrome saturation was performed by increasing the light until no more increase in fluorescent signal was seen.

Oscilloscope set-up
Channel 1 on the oscilloscope had a BNC T splitter fitted. On the DOT Plot board, the signal sent to channel 1 of the oscilloscope was set to 2 (SSC—see above). The T splitter feeds the signal from the Dot Plot board to the input channel 1 of the oscilloscope and, via an additional length of BNC cable, to the trigger board. The SSC amplifier was switched to LOG mode to enable LOG signal triggering. The SSC signal was thresholded while the sample was running by adjusting the threshold rotary knob until the bacterial population was revealed on the scatter plots. The sample rate was approximately 2000–5000 events/s. The differential pressure was low to confine the bacteria to the centre of the co-axial flow. The data were analysed using DOKOCytomation Summit software.

Quantification of TetR
Cells were grown as described above, harvested, re-suspended in water, lysed by boiling with Laemmli buffer for 10 min and analysed by SDS–PAGE, followed by Western blotting. After transfer, all proteins on the membranes were visualized by Ponceau staining, and then the membranes were scanned with Agra Duoscan 410 scanner. The stain was subsequently removed by washing with PBS. The membranes were blocked at room temperature for 1 h with 5% milk in PBST (PBS containing 0.05% Tween 20) and then exposed to the primary antibody for 1 h at room temperature and to the secondary antibody for 45 min. Each step with an antibody was followed by two 10 min washes with PBST. After the final wash, TetR bands were visualized with ECL and the exposed films were scanned. Using the program IQMac, the number of pixels for each band on the film was determined and then divided by the number of pixels from the total protein stain of the corresponding lane. Both values were first corrected for background (by subtracting the number of pixels of membrane/film corresponding to the same area as the lane/band) to adjust for differences in the amount of protein transferred in each lane. The value for the bands of TetR from DH5αZ1 cells was used as a reference to calculate the concentration of TetR in cells carrying pZS*14T and pZS**11T.

Simulation
Basis of SmartCell
SmartCell, a software written in C++, is designed for modelling biological processes occurring in a cell (Ander et al., 2004). The simulation environment is divided into elementary divisions, called ‘voxels’, to localize events and species. Two groups of voxels can be defined: (1) the compartment, limited by a membrane, is used to define an area with particular properties, and (2) the region is only used to localize reactions or initial amounts. In each voxel, the species can be represented by the concentration or the number of particles. Two types of events exist in SmartCell, diffusion events, representing the movements of species in one compartment or between two compartments, and chemical reactions. At the end of the simulation, two types of output presenting the evolution of species during time are created, the SUM files, representing the evolution in a compartment, and the VOXEL files, representing the evolution in each voxel.

There are two easy-to-use interfaces, available to use SmartCell. First, the graphic user interface facilitates the writing and design of the biological model. Second, an output analysis tool gives an easier and faster way to analyse a huge amount of outputs.

The executable version of SmartCell is freely available on the web page of SmartCell project, held at EMBL: http://smartcell.embl.de.

Improvements of SmartCell algorithm
The version of SmartCell presented in the SmartCell paper (Ander et al., 2004) was based on the next event algorithm, using Gibson and Bruck’s (2000) optimization of the Gillespie algorithm. The most important aspect of this algorithm is the use of one event queue to sort the events that can happen during the simulation. All events are duplicated as many times as the number of voxels where it may occur. The next subvolume method (Elf et al., 2003; Elf and Ehrenberg, 2004) is an alternative to the next event algorithm. For this method, the program uses only one queue, with a size equal to the number of voxels. This queue is sorted with the time when the next event will occur in each voxel. As explained in MesosRD papers (Elf et al., 2003; Elf and Ehrenberg, 2004), the next reaction algorithm could be complementary to that of the next event and, consequently, SmartCell now proposes both algorithms.

Noise determination
We have run each simulation 200 times for each circuit and condition. As there is no correlation between two time points separated by 10 000 s, we have extracted several values from each run. For the Vc, we have taken values at times 40 000, 60 000 and 80 000 s. For the histograms of GFP, we have taken values at times 20 000, 30 000, 40 000, 50 000, 60 000, 70 000 and 80 000 s. For the networks with low copy number, we have made 100 runs and we have taken values at times 40 000, 60 000 and 80 000 s.

The noise frequency analysis was performed using the normalized autocorrelation functions as described by Austin et al. (2006), using a sampling interval Tt of 10 s. The noise frequency range Ft was found using Ft=F0/τ1/2, where τ1/2 was the value of τ where the normalized autocorrelation function reached a value of 1/2. The normalized autocorrelation function is defined by

$$ACF_n(τ_t) = \frac{\sum_{m=1}^{N_p} X_m(τ_t) X_{m+1}(τ_t+τ)}{\sum_{m=1}^{N_p} X_m(τ) X_{m+1}(τ)} - \frac{X_0[n]}{X_0[n]}$$

where Xn is the copy number function, [Xn] the mean value of the Xn function and Tt the sampling interval (here 10 s). A graph of normalized probability of noise frequency was finally made using a binning of 0.5 × 10^-3.

Description of the networks
Three related networks involving TetR and GFP in an E. coli cell are modelled. In all networks, competition of RNA polymerase binding to the promoter in the plasmids for chromosomal promoters is modelled assuming that all chromosomal promoters have the same properties.
and are thus represented by a single species (Z). The number of Z was assumed to be 870 based on the fact that this is approximately the number of genes/operons expressed simultaneously in E. coli (Selinger et al., 2000). The networks are schematically shown in Figure 1 and the reactions describing them in Figure 5.

Simulations are made in an E. coli cell of \(0.8 \times 10^{-3} \text{m}^3\). This cell is represented by a single voxel with a lattice length of 0.8974 \(\mu\text{m}\). There are 10 main species involved in the networks:

- RNA polymerase (P). In all simulations, the initial value for P is 3600 molecules (Link et al., 1997; Shepherd et al., 2001; Mooney and Landick, 2003).
- Plasmid expressing TetR, with the DNA segment of interest divided in two parts: the ‘activator’ part (A), which corresponds to the promoter (including any regulatory region), and the tetR gene (B).
- Plasmid expressing GFP, with the DNA segment of interest divided in two parts: the ‘activator’ part (C), which corresponds to the promoter (including any regulatory region), and the gfp gene (D).
- Chromosomal E. coli promoters, which are considered to be the same for simplicity (Y).
- Chromosomal E. coli genes/operons, which are considered to be the same for simplicity (Z).
- TetR mRNA (M).
- GFP mRNA (N).
- TetR protein. TetR is biologically active as a dimer (Hillen and Berens, 1994). As TetR has an exceptionally high dimerization equilibrium constant (Backes et al., 1997), it is safe to assume that it will be only in the dimer form, even at low concentrations, and thus dimerization is not explicitly modelled. Therefore, for simplicity, it is considered that the product of the RNA transcript M is the biologically active species.
- GFP protein.
- aTc is modelled to bind with the same affinity to both free TetR and TetR–DNA complexes, resulting in preventing DNA binding or releasing DNA, respectively. As two molecules of aTc can bind to a TetR dimer, the stepwise binding of the aTc molecules to either free TetR (reactions (22) and (23)) or DNA-bound TetR (reactions (12), (13), (27) and (28)) has been explicitly modelled (see Figure 5).

For the sake of simplicity, the following assumptions were made:

- The whole process starting from the closed promoter–polymerase complex to actual transcription of the gene is modelled as one reaction with a rate constant \(k_{\text{on}}\), which is considered to be the same for all promoters (reactions (2), (7) and (18)) (see Figure 5).
- The rate constants for the dissociation of promoter–polymerase complex are the same for all promoters. Thus, as \(k_{\text{on}}\) is also the same, differences in promoter strengths are modelled only by changing the \(k_{\text{off}}\) for the formation of the promoter–polymerase complex.
- As the length of the tetR and gfp genes is approximately the same, the rate constants of transcription, translation and RNA degradation are the same for them.
- As the half-life of TetR (Becskei and Serrano, 2000) and GFP (Andersen et al., 1998) exceed by far the time of cell division, it can be assumed that the degradation rate constants are equal and correspond to a half-life equal to the time of cell division as measured from the experiments in this study (degT<sub>1</sub> rate=3.85e–4, degG<sub>1</sub> rate=9.8e–4).
- Degradation of TetR is the same regardless of the number of aTc molecules bound to it.
- The rates of M/N degradation and TetR/GFP production ensure an RNA lifetime of around 3 min and that one copy of RNA produce around 20 copies of TetR/GFP protein before it is degraded.
- The rate constants for the dissociation of promoter–polymerase complex to actual transcription of the gene is modelled as one reaction with a rate constant \(k_{\text{on}}\), which is considered to be the same for all promoters (reactions (2), (7) and (18)) (see Figure 5).
- The rate constants for the dissociation of promoter–polymerase complex are the same for all promoters. Thus, as \(k_{\text{on}}\) is also the same, differences in promoter strengths are modelled only by changing the \(k_{\text{off}}\) for the formation of the promoter–polymerase complex.
- As the length of the tetR and gfp genes is approximately the same, the rate constants of transcription, translation and RNA degradation are the same for them.
- As the half-life of TetR (Becskei and Serrano, 2000) and GFP (Andersen et al., 1998) exceed by far the time of cell division, it can be assumed that the degradation rate constants are equal and correspond to a half-life equal to the time of cell division as measured from the experiments in this study (degT<sub>1</sub> rate=3.85e–4, degG<sub>1</sub> rate=9.8e–4).
- Degradation of TetR is the same regardless of the number of aTc molecules bound to it.
- The rates of M/N degradation and TetR/GFP production ensure an RNA lifetime of around 3 min and that one copy of RNA produce around 20 copies of TetR/GFP protein before it is degraded.
- Formation of complex with one molecule of aTc reduces the repressor’s binding rate constant to A from \(k_{\text{off}}\) \(=10^{-3} \text{s}^{-1}\) to \(k_{\text{off}}\) \(=0.001 \text{s}^{-1}\) and with two molecules to \(k_{\text{off}}\) \(=100 \text{s}^{-1}\) based on the respective equilibrium constants (Lederer et al., 1995) and assuming that \(k_{\text{on}}\) is the same in all cases.
- When considering that plasmid numbers can fluctuate, we add or eliminate plasmids in a random fashion. Molecules attached to existing plasmids are released when the plasmid disappears.

All simulations followed the evolution of the different species for different aTc concentrations for 1 day (86 400 s).

**Network 1**

In the E. coli cell, there is only one type of plasmid, the one expressing TetR-GFP, with copy number 4 (thus A=4). TetR-GFP has the ability to repress its own production by competing with the RNA polymerase for binding to promoter A.

**Network 2**

In the E. coli cell, there are two types of plasmids: the one as in network 1 with A=4 and a plasmid that expresses GFP from the TetR-regulated promoter C, with a copy number 60 (thus C=60). The regulated promoter in both plasmids is the same and is repressed by TetR; so apart from the copy number, the values for the rate constants are the same for reactions involving A and C (Table I).

**Network 3**

It is the same as network 2 with the difference that TetR cannot repress its own transcription. Therefore, in this model, the rate constants for reactions involving A and C are different (Table I).

**Acknowledgements**

We thank Professor Bujard for providing plasmids pZE21-MCS-1, pZEl2-luc and pZS*24-MCS-1 and strain DH5αZ1. We also thank Andy Riddell at the EMBL core facilities for his help with the flow cytometry experiments. This project was partly supported by EC grants LSHG-CT-2004-503568 and LSHG-CT-2003-505520.

**References**

Ander M, Beltrao P, Di Ventura B, Ferkinghoff-Borg J, Foglierini M, Kaplan A, Lemerie C, Tomas-Oliveira I, Serrano L (2004) SmartCell, a framework to simulate cellular processes that combines stochastic approximation with diffusion and localisation: analysis of simple networks. Systems Biol 1: 129–138

Andersen JB, Sternberg C, Poulsen LK, Bjorn SP, Givskov M, Molin S (1998) New unstable variants of green fluorescent protein for studies of transient gene expression in bacteria. Appl Environ Microbiol 64: 2240–2246

Austin DW, Allen MS, McCollum JM, Dar RD, Wilgus JR, Sayler GS, Samatova NF, Cox CD, Simpson ML (2006) Gene network shaping of inherent noise spectra. Nature 439: 608–611

Backes H, Berens C, Helbl V, Walter S, Schmid FX, Hillen W (1997) Combinations of the alpha-helix-turn-alpha-helix motif of TetR with respective residues from LacI or 434Cro: DNA recognition, inducer binding, and urea-dependent denaturation. Biochemistry 36: 5311–5322

Becskei A, Serrano L (2000) Engineering stability in gene networks by feedback control. Nature 405: 590–593

Becskei A, Seraphin B, Serrano L (2001) Positive feedback in eukaryotic gene networks: cell differentiation by graded to binary response conversion. EMBO J 20: 2528–2535

Blake WJ, Kaern M, Cantor CR, Collins JJ (2003) Noise in eukaryotic gene networks: cell differentiation by graded to binary response conversion. EMBO J 22: 2722–2729

Degenkolb J, Takahashi M, Elledstad GA, Hillen W (1991) Structural requirements of tetracycline–Tet repressor interaction: determination of equilibrium binding constants for tetracycline analogs with the Tet repressor. Antimicrob Agents Chemother 35: 1591–1595
Elf J, Doncic A, Ehrenberg M (2003) Mesoscopic reaction–diffusion in intracellular signaling. SPIE: Fluctuations Noise Biol Biophys Biomed Systems 5110: 114–124
Elf J, Ehrenberg M (2004) Spontaneous separation of bi-stable biochemical systems into spatial domains of opposite phases. Systems Biol 1: 230–236
Elowitz MB, Levine AJ, Siggia ED, Swain PS (2002) Stochastic gene expression in a single cell. Science 297: 1183–1186
Gardner TS, Cantor CR, Collins JJ (2000) Construction of a genetic toggle switch in Escherichia coli. Nature 403: 339–342
Gibson MA, Bruck J (2000) Efficient exact stochastic simulation of chemical systems with many species and many channels. J Phys Chem A 104: 1876–1889
Gillespie DT (1977) Exact stochastic simulation of coupled chemical reactions. J Phys Chem 81: 2340–2361
Hattne J, Fange D, Elf J (2005) Stochastic reaction–diffusion simulation with MesoRD. Bioinformatics 21: 2923–2924
Hillen W , Berens C (1994) Mechanisms underlying expression of Tn10 encoded tetracycline resistance. Annu Rev Microbiol 48: 345–369
Hooshangi S, Thiberge S, Weiss R (2005) Ultrasensitivity and noise propagation in a synthetic transcriptional cascade. Proc Natl Acad Sci USA 102: 3581–3586
Kepler TB, Elston TC (2001) Stochasticity in transcriptional regulation: origins, consequences, and mathematical representations. Biophys J 81: 3116–3136
Lederer T, Takahashi M, Hillen W (1995) Thermodynamic analysis of tetracycline-mediated induction of Tet repressor by a quantitative methylation protection assay. Annu Biochem 232: 190–196
Link AJ, Robison K, Church GM (1997) Comparing the predicted and observed properties of proteins encoded in the genome of Escherichia coli K-12. Electrophoresis 18: 1259–1313
Lutz R, Bujard H (1997) Independent and tight regulation of transcriptional units in Escherichia coli via the LacR/O, the TetR/O and AraC/II-12 regulatory elements. Nucleic Acids Res 25: 1203–1210
Mooney RA, Landick R (2003) Tethering sigma70 to RNA polymerase reveals high in vivo activity of sigma factors and sigma70-dependent pausing at promoter-distal locations. Genes Dev 17: 2839–2851
Oliva B, Gordon G, McNicholas P, Ellestad G, Chopra I (1992) Evidence that tetracycline analogs whose primary target is not the bacterial ribosome cause lysis of Escherichia coli. Antimicrob Agents Chemother 36: 913–919
Ozbudak EM, Thattai M, Kurtser I, Grossman AD, van Oudenaarden A (2002) Regulation of noise in the expression of a single gene. Nat Genet 31: 69–73
Paulsson J (2004) Summing up the noise in gene networks. Nature 427: 415–418
Pedraza JM, van Oudenaarden A (2005) Noise propagation in gene networks. Science 307: 1965–1969
Raser JM, O’Shea EK (2004) Control of stochasticity in eukaryotic gene expression. Science 304: 1811–1814
Savageau MA (1974) Comparison of classical and autogenous systems of regulation in inducible operons. Nature 252: 546–549
Selinger DW, Cheung KJ, Mei R, Johansson EM, Richmond CS, Blattner FR, Lockhart DJ, Church GM (2000) RNA expression analysis using a 30 base pair resolution Escherichia coli genome array. Nat Biotechnol 18: 1262–1268
Shepherd N, Dennis P, Bremer H (2001) Cytoplasmic RNA polymerase in Escherichia coli. J Bacteriol 183: 2527–2534
Simpson ML, Cox CD, Sayler GS (2003) Frequency domain analysis of noise in autoregulated gene circuits. Proc Natl Acad Sci USA 100: 4551–4556
Stundzia AB, Lumsden CJ (1996) Stochastic simulation of coupled reaction–diffusion processes. J Comput Phys 127: 196–207
Swain PS, Elowitz MB, Siggia ED (2002) Intrinsic and extrinsic contributions to stochasticity in gene expression. Proc Natl Acad Sci USA 99: 12795–12800
Thieffry D, Huerta AM, Perez-Rueda E, Collado-Vides J (1998) From specific gene regulation to genomic networks: a global analysis of transcriptional regulation in Escherichia coli. BioEssays 20: 433–440
Vilar JM, Saiz L (2005) DNA looping in gene regulation: from the assembly of macromolecular complexes to the control of transcriptional noise. Curr Opin Genet Dev 15: 136–144
Zacharias DA, Violin JD, Newton AC, Tsien RY (2002) Partitioning of lipid-modified monomeric GFPs into membrane microdomains of live cells. Science 296: 913–916

Noise suppression in transcription
Y Dublanche et al