Model-based robustness and bistability analysis for methylation-based, epigenetic memory systems

Viviane Klingel1, Jakob Kirch1,*, Timo Ullrich2,†, Sara Weirich2, Albert Jeltsch2 and Nicole E. Radde1

1 Institute for Systems Theory and Automatic Control, University of Stuttgart, Germany
2 Institute of Biochemistry and Technical Biochemistry, University of Stuttgart, Germany

Keywords
bifurcation analysis; bistability; epigenetic memory systems; positive feedback; quasi-bistability

Correspondence
N. E. Radde, Institute for Systems Theory and Automatic Control, University of Stuttgart, Pfaffenwaldring 9, 70569 Stuttgart, Germany
Tel: +49 711 685 6772
E-mail: nicole.radde@ist.uni-stuttgart.de

Present address
* Insilico Biotechnology AG, Stuttgart, Germany
† Friedrich Miescher Laboratory, Tuebingen, Germany

(Received 16 October 2020, revised 12 March 2021, accepted 23 March 2021)
doi:10.1111/febs.15838

Introduction
Epigenetic systems are able to store information in form of DNA methylation patterns in a durable but reversible manner [1,2]. Memory function is usually realized by a positive feedback loop. Figure 1 shows the functionality of a minimal motif for a memory system. It consists of a state x that is regulated by a positive feedback circuit and an input signal u(t) (Fig. 1A). The feedback corresponds in the simplest case to a positive autoregulation, but can also contain more than one component [3–6]. It is described by a term r(x), which has to be nonlinear (Fig. 1B) [7,8]. Often, Hill-type functions are used to model r(x), especially in the context of gene regulation systems [3], but also to describe post-transcriptional regulation such as multisite phosphorylation levels [9,10]. Consequently, the system can operate in two different states, the ON state, in which regulation by the positive feedback circuit operates near its maximum, and the OFF state, where r(x) is close to zero. To generate simulation results in Fig. 1, we have used the system

\[ f(x) = a + r(x) - dx + u(t), \quad r(x) = c \frac{x^4}{x^4 + 8}, \]  

Abbreviations
AI, after induction; BI, before induction; CcrM, cell cycle-regulated methyltransferase; MTase, methyltransferase; SNB, saddle-node bifurcation; ZnF4, (multimerized engineered) zinc finger.

In recent years, epigenetic memory systems have been developed based on DNA methylation and positive feedback systems. Achieving a robust design for these systems is generally a challenging and multifactorial task. We developed and validated a novel mathematical model to describe methylation-based epigenetic memory systems that capture switching dynamics of methylation levels and methyltransferase amounts induced by different inputs. A bifurcation analysis shows that the system operates in the bistable range, but in its current setup is not robust to changes in parameters. An expansion of the model captures heterogeneity of cell populations by accounting for distributed cell division rates. Simulations predict that the system is highly sensitive to variations in temperature, which affects cell division and the efficiency of the zinc finger repressor. A moderate decrease in temperature leads to a highly heterogeneous response to input signals and bistability on a single-cell level. The predictions of our model were confirmed by flow cytometry experiments conducted in this study. Overall, the results of our study give insights into the functional mechanisms of methylation-based memory systems and demonstrate that the switching dynamics can be highly sensitive to experimental conditions.
where $a$ describes the basal production rate, $dx$ a linear degradation, and $u(t)$ a time-dependent input.

Memory systems usually operate in a bistable range, as shown in Fig. 1C. If the input $u$ is used as a bifurcation parameter, the system behavior can be partitioned into three regimes: (a) For small (negative) input values, the system is monostable, with a low steady state, in which the feedback regulation is turned off, and (b) for an intermediate range of input values, which is bounded by two saddle-node bifurcations (SNB), the system is bistable and can operate in the ON or OFF state, depending on the history of the system. In this range, the input $u$ is sufficient to maintain the ON state, but too small to induce a switch to the ON state, and (c) for large input values, the system again becomes monostable, with a high steady state in which the feedback regulation is near its maximal value.

The response of a system operating in the bistable range for $u(0)$ is shown in Fig. 1D. Starting in the OFF state, a transient increase in the input $u(t)$ induces a switch to the ON state, which is then constantly maintained even after resetting $u$. We note here that strong ultrasensitivity in the response curve, which corresponds in our example to large Hill coefficients, is often favorable to obtain bistability or oscillating behavior. Multiple minimal models require unrealistically large Hill coefficients [11,12]. Strong ultrasensitivity can be caused by cooperative interactions between regulatory subunits such as transcription factors, by interacting binding sites or by spatial regulation [9,13–16].

Here, we consider synthetic epigenetic memory systems based on DNA methylation, which controls the interaction of DNA-binding proteins with their target sites [17]. Experimental studies with these bistable systems have shown that transient input information can cause a switch in the steady state of the systems that can be memorized for many cell generations [17]. A scheme of the two systems, which are considered in this study, is shown in Fig. 2. The bistable basic module of both systems consists of a multimerized engineered zinc finger (ZnF4), an artificial methylation-dependent DNA-binding protein that is autoregulated and hence stably expressed at low levels, and the cell cycle-regulated methyltransferase (CcrM). This module can exist in two states [17]: In the initial OFF state, the ZnF4 repressor is cooperatively bound to four binding sites in the promoter region of the ccrM gene and represses its expression. Once binding of the ZnF4 repressor is lost, a reporter maintenance operon including an EGFP reporter and the CcrM MTase is expressed and the methyltransferase can methylate its own promoter region. This methylation weakens ZnF4 binding, thus resulting in a positive feedback loop and a switch to the ON state, in which CcrM is stably expressed. This state can be inherited throughout DNA replication and cell division. Since during cell division the DNA is duplicated and thereby the epigenetic marks are diluted, the methylation pattern can only stably be inherited to the daughter cells if the system operates in a bistable regime [18].
It has been shown that different trigger systems can be coupled to this positive feedback device [17]. In system I (Fig. 2A), temperature is used as a physical parameter. An overnight shift of the cultivation temperature of *E. coli* cells from 30 to 37 °C weakens the repressive effect of ZnF4 and induces a switch into the ON state, which is stably maintained for at least 72 h after induction (Fig. 2B,C). System II (Fig. 2D) uses a trigger plasmid to detect and memorize transient chemical signals, here arabinose supplementation. In the initial OFF state, trigger CcrM expression is repressed by glucose. Arabinose induces CcrM and mCherry expression from the trigger plasmid, resulting in a switch into the ON state, which is stably maintained after removal of arabinose, as can be seen in the EGFP and mCherry signal courses (Fig. 2E,F).

Recently, this memory system has also been connected to a tetracycline-responsive trigger system [19].

A main difficulty in the design of synthetic epigenetic memory systems is the fact that the behavior of even simple circuits is hard to predict, because often a given network topology can show qualitatively different behaviors depending on its quantitative features [20], such as mono- and bistability in our case. For the design of a functional memory system, a bistable regime has to be found. Whether such a regime exists at all in a positive feedback system depends on many design parameters, and if it exists, it is often necessary to optimize several parameters to reach this regime. This is usually only possible by a trial-and-error procedure and a good intuition. For a reliable functioning, the design has to a certain extent be robust against internal...
fluctuations, variations of environmental factors, experimen-
tal procedures, and other perturbations.

Besides the challenge to develop parts that do not interfere with the physiology of the host organism, for the functionali-
ty of the systems under study it was for instance crucial to select the artificial repressor for binding of their target sequence and its sensitivity on the methylation level, such that the preference for unmethylated binding sites was much larger than the preference for methylated sites. Furthermore, tetramer-
ization of the repressor and a specific arrangement of pairwise ZnF4-binding sites on the DNA were necessary to increase the cooperativity and binding affinity for effective repression and avoiding spontaneous acti-
vation in the OFF state, as well as a tight autoregulation of ZnF4 expression, to keep low repressor levels, were essential for the design [17].

When describing memory systems with a mathemati-
cal model, all these optimization steps directly or indi-
rectly affect model parameters. Thus, if a good model is available, it can be used to support the design pro-
cess by exploring the search space and providing edu-
cated guesses to improve the design.

In this study, we introduce models for the two epi-
getic memory systems described above. Our modeling approach is able to capture courses of CcrM expression and methylation levels and to recapitulate the dynamics of an experiment, which investigated cyclic switches between ON and OFF states [17]. Robust-
ness of the memory function is investigated by bifurca-
tion analyses, which reveals that the bistable range is not too robust with respect to parameter vari-
ations. We particularly focus on the temperature at which the experiment is conducted, which affects cell division and ZnF4 dissociation, to support the design of the experimental procedure.

Results

Mathematical model for methylation-based epigenetic memory systems

We set up a differential equation model for the two epigenetic memory systems described in Maier et al. [17] and shown in Fig. 2. Therefore, we made the simplifying assumption that the ensemble of cells which were used for the measurements can be described by a closed and well-stirred chemical reaction system with a constant volume, which implies that the total number of DNA molecules, DNAtot, as well as the total number of methylation sites, MStot, is conserved quantities. Since we do not have data regarding the cellular concentrations of ZnF4, and it anyway is tightly autoregulated, we describe it as a constant parameter.

The dynamics of the system are described by the fol-
lowing differential equations:

\[
\frac{dZF4BS}{dt} = c_v ZF4BS + d_1 u_1(t) ZF4BS \rho^{MS} - a_1 ZF4BS,
\]

\[
\frac{dZF4BS}{dt} = -\frac{dZF4BS}{dt},
\]

\[
\frac{dMS}{dt} = \Delta t ZF4BS + MS_{CcrM_{maint} + CcrM_{trigger}} - c_v MS,
\]

\[
\frac{dMS}{dt} = -\frac{dMS}{dt},
\]

\[
\frac{dCcrM_{maint}}{dt} = k_v ZF4BS - d CcrM_{maint},
\]

\[
\frac{dCcrM_{trigger}}{dt} = k_v u_2(t) - d CcrM_{trigger}.
\]

The number of binding sites not bound by ZnF4 (ZF4BSb) changes via association and dissociation of ZnF4, which are both modeled using first-order kinetics with respect to the number of binding sites available for the respective reaction, with rate constants \(a_1\) and \(c_v\), respectively. Dissociation of ZnF4 can be split into two parts: First, new DNA strands with free binding sites are generated during each cell division process; hence, the rate constant \(a_1\) corresponds to the cell division rate. Second, also cell-cycle-independent dissociation can happen, which consists of a parameter for the basal rate, \(d_1\), a heat-dependent increase in dissociation for system I (input \(u_1\)) and the influence of methylation in the binding site on binding. The heat-
dependent increase in dissociation may be an indirect effect of ZnF4 unfolding.

This is captured by the term \(\rho^{MS}\), which can be derived as follows: Be \(\Delta G_u = -4U\) and \(\Delta G_m = -4M\) the changes in the free energies for the ZnF4-binding reaction to unmethylated and fully methylated DNA, respectively, and \(d_{U-M} := U - M\). The factor 4 appears here because there are in total four ZnF-binding sites in our design. In an ensemble, the average change \(\Delta G(x_2)\) can be expressed as a function of the fraction \(x_2\) of methylated sites,

\[
\begin{align*}
\Delta G_u &= -4U, \\
\Delta G_m &= -4M = -4(U - d_{U-M}), \\
\Delta G(x_2) &= -4U + 4d_{U-M}x_2.
\end{align*}
\]

By using the Arrhenius equation for the dissociation rate constant \(k_v = d(x_2)/a_1\) of ZnF4 binding, where the dissociation rate \(d\) depends on the methylation pattern, we get
\[ k_d(x_2) = \frac{d(x_2)}{a_1} = \exp \left( \frac{c \Delta G(x_2)}{RT} \right) \]

\[ = \frac{1}{a_1} \exp^{-4U(x_2)} c^{d(x_2)-w} \]

\[ = \frac{d_1}{a_1} p^{x_2}. \]

Since the total number of DNA molecules and therefore the number of ZnF4-binding sites are conserved quantities, the temporal derivative of the number of binding sites that are bound by ZnF4 (ZF4BS) equals that of ZF4BSd, with an opposite sign.

The number of methylated sites (MSn) changes upon methylation via CcrM, which is only possible when ZnF4 is not bound. Methylation can be done by wild-type CcrM (s1 = 1) and trigger CcrM. The inactive CcrM mutant in Fig. 2 is mimicked by s1 = 0. Since the trigger plasmid is only present in system II, s2 = 0 for system I and s2 = 1 for system II. Similar to the dissociation of ZnF4 from the DNA, demethylation is coupled to cell division. Also here, the number of unmethylated sites (MSd) is described by a sign change in the right-hand side of the differential equation.

Maintenance CcrM (CcrM\text{maint}) is expressed if ZnF4 is not bound to its own promotor, and it is degraded with a rate \(d_c\). Finally, the dynamics of the trigger CcrM in system II (CcrM\text{trigger}) is described by an input-dependent production term and a first-order degradation term.

The conservation laws allow us to normalize the number of DNA molecules that are not bound by ZnF4 and the number of methylated sites to DNA\text{tot} and MS\text{tot}, respectively. Such a normalization is necessary since we do have fold measurements rather than absolute amounts. This leads to a rescaling of variables and model parameters as follows (for details, see Appendix S1, Section 1),

\[ x_1 := \frac{ZF4BS_d}{DNA_{tot}} \quad \Rightarrow \quad \frac{ZF4BS_d}{DNA_{tot}} = 1 - x_1 \]

\[ x_2 := \frac{MS_d}{MS_{tot}} \quad \Rightarrow \quad \frac{MS_d}{MS_{tot}} = 1 - x_2 \]

\[ x_3 := \frac{c_{\text{maint}} CcrM_{\text{maint}}}{c_{\text{trigger}} CcrM_{\text{trigger}}} \quad x_4 := \frac{c_{\text{trigger}} CcrM_{\text{trigger}}}{c_{\text{maint}}} \]

\[ k_3 := \tilde{k}_3 DNA_{tot} CcrM_{\text{trigger}} \quad k_4 := \tilde{k}_4 CcrM_{\text{trigger}} \]

\[ p := \frac{p_{MS\text{tot}}}{a_2 := \tilde{a}_2 DNA_{tot} CcrM_{\text{trigger}}} \]

\[ r := \frac{c_{\text{trigger}}}{c_{\text{maint}}} \]

where \(c_{\text{maint}}\) and \(c_{\text{trigger}}\) account for the different proportionality factors of EGFP and mCherry signals. Therefore, variables \(x_3\) and \(x_4\) describe quantities that are proportional to the amounts of maintenance and trigger CcrM, respectively.

The normalized model reads

\[ \dot{x}_1 = c_s (1 - x_1) + d_1 u_1(t) (1 - x_1)p^{x_2} - a_1 x_1 \]

\[ \dot{x}_2 = a_2 x_1 (1 - x_2) (x_1 x_3 + x_2 x_4) - c_r x_2 \]

\[ \dot{x}_3 = \frac{k_3}{r} x_1 - d_3 x_3 \]

\[ \dot{x}_4 = k_4 u_2(t) - d_3 x_4, \]

with inputs

\[ u_{i,2}(t) = \begin{cases} 1 & t < AI - 12 \text{ h} \\
0 & t \in [AI - 12 \text{ h}, AI] \\
1 & t < AI \end{cases} \]

The model variables, their dependency structure, the respective interaction graph, and the feedback function are shown in Fig. 3. The interaction graph in Fig. 3B, which illustrates the dependency structure in Fig. 3A, shows two positive feedback loops. First, the feedback between variables \(x_1\) and \(x_2\) describes that on the one hand methylation is only possible if the repressor ZnF4 is not bound, and on the other hand, the binding preference of ZnF4 to an unmethylated site is much higher than to a methylated one. It is responsible for stabilizing the ON and the OFF states. The second positive feedback loop comprises the variables \(x_1\), \(x_2\), \(x_3\), and \(x_4\) and describes methylation by CcrM, which is responsible for a sustained response and the functioning as a memory module. Readouts of this system are \(x_2\), \(x_3\), and \(x_4\), as indicated in red. The positive feedback on \(x_2\) is of a nonlinear nature to enable saddle-node bifurcations, but not monotonically increasing due to the conservation relations in the system (Fig. 3C). The intersections of the positive feedback functions \(r^{(1,2)}(x_2)\), which describe the methylation rates caused by the feedback term in dependence of \(x_2\), with the demethylation terms, correspond to the fixed points of the system.

A correct matching of measured and actual methylation levels is necessary to capture experimental data

During the modeling process, it turned out that a detailed description of the methylation measurement procedure and a mapping from measured to actual methylation levels, which we denote methylation matching (MM), improved our model fit significantly.
As illustrated in Fig. 4A, the DNA has two methylation sites per DNA strand. In the restriction/protection-based DNA methylation detection system applied by Maier et al. [17], the DNA was cleaved and hence recognized as unmethylated \( y_2 = 0 \) if at least one of the two methylation site pairs is unmethylated in both DNA strands. Thus, DNA molecules with a single methylated site are cleaved and hence the measured methylation level underestimates the actual level in this case. In contrast, if at least one DNA strand of each methylation site is methylated, the DNA is protected from cleavage and counts as methylated \( y_2 = 1 \). Thus, a DNA with, for example, three out of four methylated sites counts as completely methylated and the measured methylation level \( y_2 \) in a cell population is obtained from the actual methylation level \( x_2 \) by considering the probability that both restriction sites are protected, \[ y_2(x_2) = P(\text{Both restriction sites protected}) = x_2^4 + 4 \cdot x_2^3 \cdot (1 - x_2) + 4 \cdot x_2^2 \cdot (1 - x_2)^2. \]

This nonlinear relation between measured and actual methylation is illustrated in Fig. 4B. It shows that the measured methylation (a) underestimates the actual level at low overall methylation levels in the population, and (b) it overestimates the actual level at large overall methylation levels.

For both model variants, with and without MM, parameter values

\[
\theta = [a_1 \ d_1 \ k_3 \ k_4 \ d_e \ r \ p \ u_{1m} \ u_{2m}]^T \in \mathbb{R}^{10}_+
\]

were estimated from experimental data by maximum-likelihood estimation, for which we defined an error model, as explained in Appendix S1, Section 2. The parameters of both models are well identifiable, as illustrated in Fig. S1 in Appendix S1, where the distributions and 2D scatter plots of the best estimated
parameters from multiple optimization runs are shown for both model variants.

Figure 4C shows a comparison of model fits with and without MM. Without matching, the high methylation levels in the experiment 48 h after induction cannot be captured. This is due to the fact that CcrM amounts are higher directly after induction than several hours thereafter. This decrease must come along with a similar decrease in methylation levels, which is, however, not visible in the measured data. MM accounts for this by allowing the actual methylation to be smaller than the measured one. Along with this, CcrM amounts are slightly overestimated without matching as a compromise to adjust to the high
methylation levels. This is not the case with matching for system I (Fig. 4C (right)), and the effect is at least reduced in system II by taking the matching into account. Trigger CcrM is not affected by MM.

A performance plot, which depicts the objective function values from the optimization sorted according to their goodness of fit (the lower the value, the better the fit), also shows that better overall fits are obtained with MM (Fig. 4D). According to the Akaike information criterion as a model comparison measure, the model variant with MM is significantly more likely than the one without MM (Appendix S1, Section 2). In summary, the model including MM captures CcrM and methylation dynamics of both systems quite well.

The model is able to describe cycles of ON/OFF switching

In order to evaluate the predictive power of our model, we used an experiment in which the state of the epigenetic memory system was reset after induction [17], as illustrated in Fig. 5A.

Here, the signal was terminated via a synthetic protein degradation system to selectively degrade the maintenance CcrM and thus stop further methylation. Therefore, a protein degradation tag was fused to the maintenance CcrM and an arabinose-inducible mF-Lon gene was cloned into a second plasmid. The ON state was induced by overnight heat induction as in system I, and a reset to the OFF state was induced by arabinose induction, causing expression of mF-Lon and subsequent selective degradation of CcrM-deg. Data (from ref. [17]) and model predictions are shown in Fig. 5B. The amount of maintenance CcrM was measured over time and shows that the system switches stably between ON and OFF states for two cycles (green data points). Untagged CcrM was used as a control, which, as expected, was not affected by arabinose-induced degradation (yellow data points). In order to capture these experiments, we had to introduce mF-Lon as an additional state variable (x₂). The respective dependency structure and interaction graph are depicted in Fig. 5C, D, respectively. In order to describe the experiment, we adapted mF-Lon degradation and production rates (see Appendix S1, Section 3). The untagged CcrM is not affected by these new parameters and therefore completely reflects the original, estimated parameter vector. Predictions of the model are in good agreement with experimental data and capture dynamics of the tagged and the untagged CcrM quite well (green and yellow curves in Fig. 5B, respectively).

To summarize, these simulations show that our model does not only describe the data of systems I and II that have been used for parameter estimation, but is also able to recapitulate the behavior of a combined epigenetic memory system with cycles of ON/OFF switching output. Overall, we are equipped with a reliable model that can be used for further analysis of the system.

The systems operate in a bistable range, which is not too robust to changes in parameters

In order to investigate robustness of the memory function with respect to parameter variations, we conducted a bifurcation analysis with each model parameter (Appendix S1, Section 4). Exemplary bifurcation diagrams are depicted in Fig. 6, where the ZnF4 association constant, parameter a₁, the parameter p that describes the influence of methylation on ZnF4 DNA binding, and the cell division rate, parameter c₁, have been used as bifurcation parameters. All parameters were varied around the maximum-likelihood estimate θ ML from θ ML − 25% to θ ML + 25%. The bifurcation diagrams (upper row) show steady-state methylation rates, and systems I and II are depicted in red and green, respectively. Respective objective function values are shown below. As expected (Fig. 1C), the systems operate in a bistable range, which is bounded by SNBs.

A reduction in the parameter a₁ leads to a decreased binding affinity of ZnF4 to the DNA, which ultimately results in a monostable system with a globally stable ON state (Fig. 6A). The SNB appears at about 20% reduction for system I and about 10% reduction for system II. Contrary, if a₁ is increased, the system becomes monostable with a low methylation rate (OFF state), which is the case for a small increase of about 5% for system I and of about 10% for system II. Thus, a₁ is a critical parameter for the robustness of the memory system. This is in accordance with the procedure described in ref. [17], where a selection for an affinity of ZnF4 that is strong enough to prevent spontaneous switches to the ON state without input was important for the design process. Our analysis suggests in addition that the binding affinity must also not be too high, since this would prevent the permanent switch to the ON state upon induction.

A reduction in the parameter p, which describes the impact of methylation on the binding affinity of the repressor, leads to a monostable system, with SNBs at 5% reduction for system I and 10% reduction for system II, while an increase in p within 25% has no effect on the bistability (Fig. 6B). This demonstrates that methylation must indeed have a sufficiently large impact on ZnF4 binding for the memory system to be functional, which is also in accordance with Ref. [17].
The system remains in a bistable regime upon a reduction in the cell division rate $c_r$, as can be seen in Fig. 6C. Cell division results in demethylation and in dissociation of ZnF4 from the DNA. If both reactions are slowed down by decreasing the cell division rate, this stabilizes both ON and OFF states. An increase in $c_r$ eventually leads to a monostable system in the OFF state, which happens at different thresholds for systems I and II. Again, system I is more sensitive to changes in this parameter and the SNB appears at a smaller value. These results suggest that the memory function is not robust to increasing cell division rates. For all parameters, objective function values increase rapidly but continuously near the borders of the bistable ranges, as can be seen in the lower row in Fig. 6.

Bifurcation diagrams of the remaining parameters are shown in Appendix S1, Figs S2 and S3. They can be divided into parameters, which have an impact on bistability in the investigated range (Appendix S1, Fig. S2), which are the temperature-dependent ZnF4 dissociation rate (parameter $d_1$), the maintenance CcrM expression rate (parameter $k_3$), the DNA methylation rate (parameter $a_2$) and the CcrM degradation rate (parameter $d_4$), and parameters that do hardly affect system dynamics (Appendix S1, Fig. S3). The bistability in both systems is not sensitive to changes in the expression rate of trigger CcrM (parameter $k_4$), the ratio of proportionality factors for trigger and maintenance CcrM (parameter $r$), and the input parameters $u_{m1}$ and $u_{m2}$, which describe transient inputs of systems I and II, respectively.

Summarizing, the bistable regime in the parameter space is small, and hence, the memory function of both systems is relatively sensitive to changes in individual model parameters, which offers a retrospective explanation of the design of the memory system as described in ref. [17]. Furthermore, we anticipate that the cell division rate is an important parameter for the reliable functioning of the memory device.

**Population model captures loss of methylation upon a moderate temperature decrease**

Based on the outcome of the bifurcation analysis, we decided to extend our model to capture heterogeneity...
among cells by distributed cell division rates. The cell division rate is a macroscopic parameter that can be quantified and can easily be influenced by factors such as availability of nutrients or temperature. Thus, it is sensitive to the experimental protocol, for example, to the cell density at which cells are split. Growing to higher cell densities leads to lower average cell division rates due to nutrient availability and inhibition of cell division at high cell density. Temperature changes affect both cell division rate and ZnF4 binding, with dual effects on the system. On the one hand, higher temperatures increase the cell division rate and therefore also the fraction of demethylated DNA, and this effect destabilizes the ON state. On the other hand, the fraction ZnF4 bound to the DNA is decreased directly via an increased dissociation rate at higher temperatures and indirectly via an increased cell division rate, and this effect destabilizes the OFF state. This was used in system I, where the ON state was induced by a switch to higher temperatures. The question arises of which effect a lower temperature has for the switching dynamics. When the cells are cultivated at lower temperature, cell division is slower, which has a stabilizing effect for both states, but at the same time ZnF4 binding is stronger and thus might destabilize the ON state in the long run. As a consequence, if the change in ZnF4 binding is the dominating effect, the memory function would suffer from lower temperatures.

To test this prediction experimentally, cells encoding the system II were cultivated at 28 °C and the switch to the ON state was triggered by addition of arabinose. Following removal of arabinose, the fraction of cells in the ON state and the OFF state was analyzed by flow cytometry over up to 3 days. In contrast to the previous experiments at 30 °C shown in Fig. 2B,C [17], which show a fairly stable ON state over up to 3 days, the response of the cells in this new experiment was highly heterogeneous (Fig. 7A). While most of the cells are in the ON state immediately after induction, this fraction decreases continuously, such that about 80% are in the OFF state 72 h after induction.

In order to simulate these experiments, we used our model for system II and assigned individual cell division rates to each cell in the population. The decrease in temperature was taken into account by slightly increasing the mean generation time from 2 to 2.1 h, and \( c_c \) was sampled from a normal distribution with a standard deviation of about 10%, \( C_c \sim N(2.1 \text{ h}, (0.2 \text{ h})^2) \). Furthermore, the ZnF4 dissociation rate \( d_l \) was decreased from \( 10^{-0.855} \) to \( 10^{-0.93} \) to account for

---

**Fig. 6.** The bistable range is not robust against parameter variations. Both systems I and II operate in the bistable range. Upper row: exemplary bifurcation diagrams for the parameters (A) \( a_1 \) (B) \( p \) and (C) \( c_c \). The bistable range is bounded by SNBs. Objective function values in dependence of the bifurcation parameter values are shown in the lower row. These show a rapid increase around the SNBs.
stronger ZnF4 binding at lower temperatures. This value was set manually after a few forward simulation runs of the population model with different $d_1$ values and a visual comparison with experiments. Simulation results are shown in Fig. 7B,C. Figure 7B shows maintenance CcrM levels. Each trajectory corresponds to a single cell with an individual division rate. As already indicated in the bifurcation diagrams in Fig. 6C, a small cell division rate stabilizes the ON state, and hence, slowly dividing cells operate in the bistable regime and remain in the ON state after removing the input stimulus. Increasing the cell division rate brings the system closer to its bifurcation point and ultimately to the monostable regime. Consequently, two subpopulations emerge on the cell population level: Cells that operate in the bistable regime and remain stably in the ON state after stimulation, and cells that are monostable and ultimately converge to the lower globally stable fixed point, the OFF state. Many of those show quasi-bistable behavior [21]: They are monostable systems, which operate close to a SNB and are therefore characterized by a long maintenance of the ON state after removing the stimulus before converging to its low steady state [21]. On the bulk level, which corresponds to the mean of all sampled trajectories, this heterogeneous behavior results in an overall decrease in CcrM after induction, as indicated with a bold line. Figure 7C shows a direct comparison with the experimental data in Fig. 7A. It was created by sorting sample trajectories into ON and OFF according to a threshold EGFP fluorescence value of 20, as defined in ref. [17]. Here, we observe that approximately 70% of the cells have returned in the OFF state after 72 h. Overall, simulations are qualitatively in good agreement with experiments. In summary, these results show that slight variations in temperature and cell division rates have severe effects on the response of the systems and might be a main source of variability and bimodality in cell populations.

Discussion and conclusions

In this paper, we have introduced a mathematical model to describe methylation-based epigenetic memory systems, which can sense and memorize different inputs (Figs 2 and 3). Both considered systems are based on a positive feedback, in which CcrM methylates the DNA and thereby prevents binding of ZnF4, which acts as a repressor and abrogates the expression of CcrM when bound to the DNA. Different inputs were used to induce a switch from an initial OFF state to the ON state, in which CcrM is stably expressed for multiple cell cycles after induction. To capture CcrM and methylation-level courses, it was important to include a nonlinear mapping from measured to actual methylation levels that takes details of the measurement procedure into account (Fig. 4). The model was used to predict the cyclic behavior of CcrM during ON/OFF switches, which was in good agreement with experimental data (Fig. 5). A bifurcation analysis reveals that the bistable range in the parameter space...
is not large, and hence, the memory function is not robust against a variation of most of the parameters within a range of ±25% (Fig. 6). In particular, from our analysis we conclude that the cell division rate is a crucial parameter, which has to be small enough for a proper memory functioning. In order to test this hypothesis, the model was extended to capture heterogeneity within a cell population by defining the cell division rate, which is dependent on temperature, as a distributed parameter. This allowed us to simulate the behavior of cell populations at different temperatures. Since a higher temperature has a dual effect on the system, the overall effect of a temperature shift on the memory function is not a priori clear. Simulations anticipate that the behavior on the bulk level emerges from the average of two subpopulations, one of which operates in the bistable regime and shows a sustained response, the other one operating in a quasi-bistable regime near the SNB (Fig. 7). The ratio between these two subpopulations is highly sensitive to small changes in the temperature at which the experiment is conducted. Hence, for a robust memory module the temperature and the cell division rates are key parameters. This result was confirmed by flow cytometry experiments in this study and is further supported by other studies, in which growth rates were also found to significantly impact memory capacity [7,20].

In system I, the effect of temperature on ZnF4 binding was used to induce a switch to the ON state, which was stably maintained after induction. We have not included the effect of this temperature shift during induction on the cell division rate, since the exact connection between temperature and cell division is unknown. However, extending the model by an additional parameter that accounts for a change in the cell division rate during induction in system I and re-estimation of parameters did not lead to a significant change in the model fits and the parameters (see Appendix S1, Section 6).

In order to capture cell-to-cell heterogeneity, we used an ensemble model by assigning a distribution to the cell division rate. A similar approach is described in Blasco et al. [18], who consider variability in cofactors that control epigenetic regulation and thereby stochastically drive phenotypic variability in terms of distinct epigenetic states in the context of cell fate reprogramming in aging. Since cells actually differ in many aspects, a comparison of our modeling approach with experiments on the single-cell level can only be done in a qualitative way. Thus, we have not conducted a new parameter estimation. From our simulations, we anticipate that the cell division rate is a critical parameter for robustness of the switching behavior, but a detailed quantification of its contribution to heterogeneity was not possible.

The model introduced here is specific for the two particular epigenetic memory systems and hence not ad hoc applicable to other biological memory systems. Our models integrate details of these systems such as the influence of methylation on ZnF4-DNA binding and have been adapted to the available experimental data, including parameter estimation and model normalization. However, we have also learned some general facts about epigenetic memory systems. Most importantly, by our feedback analysis, we were able to reveal features important for a robust design of such memory devices. This includes for example a strong mutual inhibition between ZnF4 binding and CcrM-mediated DNA methylation, which can be generalized to similarly designed systems, and proper fine tuning of experimental conditions, which might explain difficulties in designing functional motifs in general.

Our study relies on a deterministic modeling approach, which implicitly assumes that the number of molecules involved in the biochemical reactions of the system is large enough and stochastic fluctuations can be neglected. Together with the rapid development in single-cell technologies, stochastic modeling approaches are more and more frequently used to capture heterogeneity in cell populations [22] and, in the context of memory systems, to investigate robustness to intrinsic noise due to low copy numbers [8,18,23]. In [22] the dynamics of epigenetic regulation on a single-cell level is investigated, and time-lapse microscopy reveals that silencing and reactivation of gene expression by chromatin regulators occur in all-or-none events in individual cells rather than gradually. According to this, a graduate response on the bulk level is reached by varying the fractions of cells in the ON and the OFF states, which is in accordance with our modeling results.

Application of the chemical master equation (CME) and the Gillespie algorithm to generate sample traces are standard stochastic modeling approaches. Compared to reaction rate equations, stochastic modeling via the CME is often not feasible, because of the large state space. However, the Gillespie algorithm and related approaches [24,25] usually have much longer forward simulation times, which is a challenge for all analyses that require many forward simulations such as sensitivity analysis or parameter estimation. A stochastic approach on the other hand allows to define probabilistic concepts to quantify robustness such as mean first passage times [23], or steady-state probability distributions, which are not applicable in a deterministic framework. These are related to the so-called
potential landscape of a dynamic network [26,27], which can also give rise to robustness measures. Consequently, a stochastic modeling approach could provide further insights by exploiting new concepts in the future.

For simplification, we have assumed a constant cell volume over time. In fact, molecules are diluted during cell growth [7], which effectively contributes as an additional degradation term that is coupled to the cell cycle. This could in the future be considered in a refined version of the model by introducing a volume parameter.

Finally, a further future direction to achieve a robust design of synthetic memory modules is probably the coupling of different positive circuits. Functional robustness has also been investigated in relation to network topologies [28–35], with a focus on oscillations (see, e.g., refs [36–38]), decision processes [39,40], signaling pathways [41–43] and bistability [23,44,45]. Some of these studies indicate that robustness of bistable systems can be increased by interlinked feedback loops [23,44], which might also be an interesting starting point for model-based design of epigenetic memory systems. In particular, in Brandman et al. [46] it is argued that the coupling of fast and slow positive feedback loops can produce switches that can transit rapidly from the OFF state to the ON state and are resistant to noise. Our system also consists of two coupled positive feedback loops: First, competition between ZnF4 binding and methylation, which acts on a fast time scale and leads to rapid induction, and second, the feedback on the slow time scale, which involves CcrM gene expression and protein synthesis and stabilizes the ON state. As argued in Brandman et al. [46], our system is also quite robust to large changes in the input parameters, as demonstrated by our bifurcation analysis with input parameters $u^m_1$ and $u^m_2$.

A further interesting modification of this system could be the generation of oscillations. From theoretical studies, we know that adding a negative feedback to the positive feedback can lead to oscillations [47,48]. A requirement for this to work is a large enough time delay in the system [47].

In contrast to artificially designed minimal systems, evolution has produced complex regulation mechanisms that function robustly under a wide range of conditions. In the context of epigenetic memory, a paradigmatic example of a real biological process is the activation and differentiation of B cells. This process is much more complex than the synthetic systems considered here and regulated by a concerted interplay of DNA methylation and demethylation, histone modification, and noncoding RNAs [49]. In contrast to our systems, the role of DNA methylation is mainly in the repression of gene expression [49,50]. Furthermore, in addition to passive demethylation by replication, DNA is also actively demethylated by the involvement of TET enzymes, for example. Moreover, several mechanisms contribute to the inheritance of methylation pattern, such as the preference of DNMT1 as a maintenance methyltransferase to methylate CpG sites of DNA strands whose parent strand is already methylated. All these mechanisms are not considered here.

However, cell heterogeneity is relevant in many epigenetic processes including tumor development [51]. Therefore, the inclusion and analysis of cell heterogeneity represent a more general application of this study. Interestingly, we observed that already the distribution of only one critical parameter, the cell cycle rate, leads to two subpopulation with different memory behavior.

It is an interesting general question whether parts of these processes can be explained by modeling methylation patterns with a similar simplistic model as we introduce here, and how the embedding of these processes into a larger regulatory network contributes to a robust functioning in real biological processes.

**Materials and methods**

**Model simulation**

The models were simulated in MATLAB R2019b and R2020a using the IQM-Toolbox. This enables the use of the CVODE solver from SUNDIALS for conversion and simulation of the model in C. The optimizers fmincon (MATLAB) and eSS (MEIGO-Toolbox [52]) were used to obtain the optimal parameter sets. For further information and details regarding the modeling and simulation, see Appendix S1.

**ON state stability after Arabinose-induced gene expression at 28 °C**

Experiments with the arabinose-induced gene expression system were conducted at 28 °C as described [19]. XL1-Blue E. coli cells were co-transformed with the memory plasmid or the negative memory plasmid and arabinose trigger plasmid. Single colonies were used to inoculate overnight cultures in LB medium containing 25 μg·mL$^{-1}$ kanamycin and 100 μg·mL$^{-1}$ ampicillin, 10 μM ZnSO4, and 0.2% glucose cultivated. Expression of the ccrM gene in liquid culture was induced by exchange of the glucose containing to arabinose (0.02% (w/v)) containing medium. The fraction of ON cells was based on EGFP expression and determined by flow cytometry as described [19] using a MACSQuant® VYB Flow Cytometer (Miltenyi Biotec).
Data were evaluated with the FlowJo v10 software (FlowJo™ BD, Heidelberg Germany). Gates were optimized to allow the most stringent discrimination between the ON and OFF states.

Acknowledgements

This work was funded by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany’s Excellence Strategy—EXC2075-390740016 and grants RA1840/2-1 to NR and JE252/35-1 to AJ. We furthermore acknowledge valuable input from two anonymous referees.

Conflict of interest

The authors declare no conflict of interest.

Author contributions

VK and JK developed the models and implemented all mathematical analyses. TU, SW, and AJ provided experimental data and contributed to model development. NR supervised the work and wrote the manuscript with input from all authors.

References

1 Jeltsch A (2002) Beyond Watson and Crick: DNA methylation and molecular enzymology of dna methyltransferases. ChemBioChem 3, 274–293.
2 Wion D & Casadesus J (2006) N6-methyl-adenine: an epigenetic signal for DNA-protein interactions. Nat Rev Microbiol 4, 183–192.
3 (2006), An Introduction to Systems Biology – Design Principles of Biological Circuits. Mathematical and Computational Biology Series. Chapman & Hall/CRC, Mathematical and Computational Biology Series. 1st edn. London, UK.
4 Alon U (2007) Network motifs: theory and experimental approaches. Nat Rev Genet 8, 450–461.
5 Ešišing T, Waldherr S, Allgöwer F, Scheurich P & Bullinger E (2007) Response to bistability in apoptosis: roles of Bax, Bel-2, and mitochondrial permeability transition pores. Biophys J 92, 3332–3334.
6 Xiong W & Ferrell JE (2003) A positive-feedback-based bistable 'memory module' that governs cell fate decision. Nature 426, 460–465.
7 Burrill DR & Silver PA (2010) Making cellular memories. Cell 140, 13–18.
8 Dodd IB, Micheelsen MA, Sneppen K & Thon G (2007) Theoretical analysis of epigenetic cell memory by nucleosome modification. Cell 129, 813–822.
9 Dushek O, van der Merwe PA & Shahrezaei V (2011) Ultrasensitivity in multisite phosphorylation of membrane-anchored proteins. Biophys J 100, 1189–1197.
10 Gunawardena J (2005) Multisite protein phosphorylation makes a good threshold but can be a poor switch. Proc Natl Acad Sci USA 102, 14617–14622.
11 Gonze D & Abou-Jaoudé W (2013) The Goodwin model: behind the Hill function. PLoS One 8, e69573.
12 Tyson JJ & Novák B (2010) Functional motifs in biochemical reaction networks. Annu Rev Phys Chem 61, 219–240.
13 Abeliiovich H (2005) An empirical extremum principle for the Hill coefficient in ligand-protein interactions showing negative cooperativity. Biophys J 89, 76–79.
14 Ferrell Jr JE & Ha SH (2014) Ultrasensitivity part I: Michaelian responses and zero-order ultrasensitivity. Trends Biochem Sci 39, 496–503.
15 Ferrell Jr JE & Ha SH (2014) Ultrasensitivity part II: multisite phosphorylation, stoichiometric inhibitors, and positive feedback. Trends Biochem Sci 39, 556–569.
16 Ferrell Jr JE & Ha SH (2014) Ultrasensitivity part III: cascades, bistable switches, and oscillators. Trends Biochem Sci 39, 612–618.
17 Maier J, Möhrle R & Jeltsch A (2017) Design of synthetic epigenetic circuits featuring memory effects and reversible switching based on DNA methylation. Nat Commun 8, 15336. https://doi.org/10.1038/ncomms15336
18 Folguera-Blasco N, Cuyàs E, Menéndez JA & Alarcón T (2018) Epigenetic regulation of cell fate reprogramming in aging and disease: a predictive computational model. PLoS Comput Biol 14, e1006052.
19 Ulrich T, Weirich S & Jeltsch A (2020) Development of an epigenetic tetracycline sensor system based on DNA methylation. PLoS One 15, e0232701.
20 Ajo-Franklin CM, Drubin DA, Eskin JA, Gee EPS, Landgraf D, Phillips I & Silver PA (2007) Rational design of memory in eukaryotic cells. Genes Dev 21, 2271–2276.
21 Thomaseth JC & Radde N (2017) Sampling-based bayesian approaches reveal the importance of quasi-bistable behavior in cellular decision processes on the example of the MAPK signaling pathway in PC-12 cell lines. BMC Syst Biol 11.
22 Bintu L, Yong J, Antebi YE, McCue K, Kazuki Y, Uno N, Oshimura M & Elowitz MB (2016) Dynamics of epigenetic regulation at the single-cell level. Science 351, 720–724.
23 Cheng Z, Liu F, Zhang X-P & Wang W (2008) Robustness analysis of cellular memory in an autoactivating positive feedback system. FEBS Lett 582, 3776–3782.
24 Higham DJ (2008) Modeling and simulating chemical reactions. SIAM Rev 50, 347–368.
25 Wilkinson DJ (2006) Stochastic Modelling for Systems Biology. Mathematical and Computational Biology. Vol. 11, Boca Raton, FL, USA: Chapman & Hall/CRC.
26 Han B & Wang J (2007) Quantifying robustness and dissipation cost of yeast cell cycle network: the funneled energy landscape perspectives. Biophys J 92, 3755–3763.
27 Lapidus S, Han B & Wang J (2008) Intrinsic noise, dissipation cost, and robustness of cellular networks: the underlying energy landscape of MAPK signal transduction. Proc Natl Acad Sci USA 105, 6039–6044.
28 Brunner M & Káldi K (2008) Interlocked feedback loops of the circadian clock of Neospora crassa. Mol Microbiol 68, 255–262.
29 Emmert-Streib F & Dehmer M (2009) Information processing in the transcriptional regulatory network of yeast: functional robustness. BMC Syst Biol 3, 1–11.
30 Ingolia NT (2004) Topology and robustness in the Drosophila segment polarity network. PLoS Biol 2, e123.
31 Kitano H (2002) Systems biology: a brief overview. Science 295, 1662–1664.
32 Kitano H (2007) Towards a theory of biological robustness. Mol Syst Biol 3, 1–7.
33 Li F, Long T, Ying L, Ouyang Q & Tang C (2004) The yeast cell-cycle network is robustly designed. Proc Natl Acad Sci USA 101, 4781–4786.
34 Ma W, Lai L, Ouyang Q & Tang C (2006) Robustness and modular design of the Drosophila segment polarity network. Mol Syst Biol 2, 1–9.
35 Wang R, Chen L & Aihara K (2007) Detection of cellular rhythms and global stability within interlocked feedback systems. Math Biosci 209, 171–189.
36 Barkai N & Shilo B-Z (2007) Variability and robustness in biomolecular circuits. Cell 28, 755–760.
37 Caicedo-Casso H-W, Kang SL & Hong CI (2015) Robustness and period sensitivity analysis of minimal models for biochemical oscillators. Nat Sci Rep 5, 1–13.
38 Wagner A (2005) Circuit topology and the evolution of robustness in two-gene circadian oscillators. Proc Natl Acad Sci USA 102, 11775–11780.
39 Heinrich R, Geiss E-M, Kamenz J, Trautmann S, Widmer C, Knop M, Radde N, Hasenauer J & Hauf S (2013) Determinants for robustness in spindle assembly checkpoint signalling. Nat Cell Biol 15, 1328–1339.
40 Krantz M, Ahmadpour D, Ottosson L-G, Warringer J, Waltermann C, Nordlander B, Klipp E, Blomberg A, Hohmann S & Kitano H (2009) Robustness and fragility in the yeast high osmolarity glycerol (HOG) signal transduction pathway. Mol Syst Biol 5, 1–7.
41 Fritsche-Guenther R, Witzel F, Sieber A, Herr R, Schmidt N, Braun S, Brummer T, Sers C & Blüthgen N (2011) Strong negative feedback from Erk to Raf confers robustness to MAPK signalling. Mol Syst Biol 7, 1–13.
42 Dexter JP & Gunawardena J (2012) Dimerization and bifunctionality confer robustness to the isocitrate dehydrogenase regulatory system in Escherichia coli. J Biol Chem 288, 5770–5778.
43 Paul D & Radde NE (2016) Robustness and filtering properties of ubiquitous signaling network motifs. In Proc. of Foundations of Systems Biology in Engineering (Fosbe2016), Vol. 49, pp. 120–127, Magdeburg, Germany.
44 Cheng P, Yang Y & Liu Y (2001) Interlocked feedback loops contribute to the robustness of the Neurospora circadian clock. Proc Natl Acad Sci USA 98, 7408–7413.
45 Qiao L, Nachbar RB, Kevrekidis IG & Shvartsman SY (2007) Bistability and oscillations in the Huang-Ferrrell model of MAPK signaling. PLoS Comput Biol 3, 1819–1826.
46 Brandman O, Ferrell JE, Li R & Meyer T (2005) Interlinked fast and slow positive feedback loops drive reliable cell decisions. Science 310, 496–498.
47 Tyson JJ, Chen KC & Novak B (2003) Sniffers, buzzers, toggles and blinkers: dynamics of regulatory and signaling pathways in the cell. Curr Opin Cell Biol 15, 221–231.
48 Kholodenko BN (2006) Cell-signalling dynamics in time and space. Nat Rev Mol Cell Biol 7, 165–176.
49 Wu H, Deng Y, Peng Y, Long D, Ma K, Wang X, Zhao M, Lu L & Lu Q (2018) Epigenetic regulation in B-cell maturation and its dysregulation in autoimmune. Cell Mol Immunol 15, 676–684.
50 Bird A (2002) DNA methylation patterns and epigenetic memory. Genes Dev 16, 6–21.
51 Dominguez PM, Teater M & Shaknovich R (2017) The new frontier of epigenetic heterogeneity in b-cell neoplasms. Curr Opin Hematol 24, 402–408.
52 Egea JA, Henriques D, Cokelaer T, Villaverde AF, MacNamara A, Danciu DP, Banga JR & Saez-Rodriguez J (2014) Meigo: an open-source software suite based on metaheuristics for global optimization in systems biology and bioinformatics. BMC Bioinformatics 15, 1–9.

Supporting information

Additional supporting information may be found online in the Supporting Information section at the end of the article.

Fig. S1. Parameter scatter plots.
Fig. S2. Bifurcation diagrams for parameters $d_1$, $k_3$, $a_2$ and $d_r$.
Fig. S3. Bifurcation diagrams for parameters $k_a$, $r$, $u_{1m}$ and $u_{2m}$.
Fig. S4. Flow-cytometry experiments and model simulations with system II at 28 °C.
Fig. S5. Primary FACS data for the bar diagram displayed in Figure 4a and b showing the ON-state cell analysis of the arabinose memory system at 28 °C.
Fig. S6. Model version with cell-cycle rate sensitive to heat induction.
Table S1. Estimated model parameters.

Table S2. Experimental data system I, WT CcrM.
Table S3. Experimental data system I, inactive CcrM.
Table S4. Experimental data system II, WT CcrM.
Table S5. Experimental data system II, inactive CcrM.
Table S6. Experimental data, validation.
Appendix S1. Supporting Information: Model-based robustness and bistability analysis for methylation-based, epigenetic memory system.