DNA inversion is an important mechanism by which bacteria and bacteriophage switch reversibly between phenotypic states. In such switches, the orientation of a short DNA element is flipped by a site-specific recombinase enzyme. We propose a simple model for a DNA inversion switch in which recombinase production is dependent on the switch state (orientational control). Our model is inspired by the fim switch in Escherichia coli. We present an exact analytical solution of the chemical master equation for the model switch, as well as stochastic simulations. Orientational control causes the switch to deviate from Poissonian behaviour: the distribution of times in the on state shows a peak and successive flip times are correlated.

PACS numbers: 87.18.Cf, 87.16.Yc, 82.39.-k

Reversible and heritable stochastic switching between two different states of gene expression is a common phenomenon among bacteria and bacteriophage, known as phase variation. Phase variation is often linked to pathogenesis, and may help bacteria survive fluctuating environmental conditions (e.g., a host immune system). An important molecular mechanism for phase variation is site-specific DNA inversion. Here, a short piece of DNA (the “invertible element”) is excised from the genome and reinserted (strand-by-strand) in the opposite orientation by a site-specific recombinase enzyme binding to sequences at the ends of the invertible element. Different states of gene expression correspond to the two orientations of the invertible element (“switch states”). A well-known example is the fim genetic regulatory system, which controls the production of type 1 fimbriae in Escherichia coli. These fimbriae are important in uropathogenesis. The fim system, the FimE recombinase is produced more strongly when the switch is in the on state than in the off state. This phenomenon is known as orientational control.

In this Letter, we present a simple and general stochastic model for a DNA inversion switch with orientational control. We solve this model analytically, allowing us to determine the range of stochastic switching behaviour possible for this type of switch. We find that non–Poissonian behaviour occurs, resulting in a peak in the probability distribution of time spent in the on state and correlations between successive flips. Such non–Poissonian behaviour could have important effects on the population dynamics of switching microbes in changing environments. One key parameter (the concentration of the recombinase not under orientational control) controls the degree to which our model is non–Poissonian; this parameter corresponds to the main point of environmental regulation for the fim switch. In contrast, bistable genetic switches such as positive feedback loops or mutually repressing genes in general show only Poissonian behaviour (exponential waiting time distributions and uncorrelated flips).

The Model.—Our model DNA inversion switch, illustrated in Fig. contains three elements: the invertible DNA element and two types of recombinase enzyme ($R_1$ and $R_2$). The invertible element has two possible orientations (the “on” and “off” states). These correspond to alternative patterns of gene expression, leading to different phenotypic states; however, we model here only the core of the switch and not its downstream effects. The switch can be flipped between its two orientations by either of the recombinases. The concentration of recombinase $R_2$ is assumed to be fixed, while the production of $R_1$ depends on the switch state: $R_1$ is produced only in the on state. This feature of the model constitutes its orientational control and leads to its non–Poissonian behaviour. The model is represented by the following reaction scheme:

\[
\begin{align*}
R_1 & \xrightarrow{k_1} \emptyset & S_{on} & \xrightarrow{k_2} S_{on} + R_1 \\
S_{on} + R_1 & \xrightarrow{k_{on}} k_{3}^{-1} S_{off} + R_1 & S_{on} & \xrightarrow{k_{on}^{-1}} S_{off}.
\end{align*}
\]

Here, $S_{on}$ and $S_{off}$ denote the on and off states of the switch. Reactions describe the production and decay of recombinase $R_1$: $R_1$ is removed from the system with rate constant $k_1$ (due to cell growth and division, which we do not model explicitly), and is produced at rate $k_2$ only when the switch is in the on state. Reactions describe switch flipping. This may be catalysed by $R_1$ with rate constants $k_{on}$ (on to off) and $k_{off}$ (off to on). We shall mainly consider here the case $k_{off} = 0$. Recombinase $R_2$ can also catalyse switch flipping. The concentration of $R_2$ (fixed in our model) is not explicitly included in the reaction scheme, but is implicit through a dependence of the rate constants $k_{on}^{-1}$ on the $R_2$ concentration.

We note that mean-field, macroscopic rate equations corresponding to the above reaction scheme yield only one steady-state solution corresponding to the average switch state and average concentration of $R_1$. The underlying deterministic structure of the model is thus not bistable. In this sense, our model is fundamentally different from the bistable reaction networks presented in .

Our model is inspired by the fim genetic regulatory system. In analogy with fim, $R_1$ in our model represents FimE while $R_2$ represents FimB. Environmental stimuli such as nutrient conditions and temperature act on the fim switch largely through changes in the level of FimB; this would correspond to variation of our key parameters $k_{on}^{-1}$. However, our model is highly simplified in comparison to fim, in that it neglects cooperative and competitive recombinase binding...
and the effects of other DNA binding proteins. Our objective in this work is not to model the details of the \textit{fim} system, as other authors have done \cite{2,10}, but rather to address general questions about the behaviour of this type of switch. Our model is designed to be as simple as possible while retaining the key features of DNA inversion and orientational control; its simplicity allows us to obtain analytical results and to explore a wide range of parameter space.

We simulated the reaction scheme (1) using a continuous time Monte Carlo scheme \cite{11}. A typical trajectory is shown in Fig. 1 where we plot the number \( n \) of \( R_1 \) molecules and the switch state as functions of time. When the switch is in the on state, \( n \) increases (on average) towards a plateau value of \( k_{2}/k_{1} \), while in the off state \( n \) decays towards zero. We now obtain an exact analytical solution for the case where \( k_{3}^{off} = 0 \). This case is relevant to the \textit{fim} switch, where FimE catalysates almost exclusively on to off switching. In the following, all our analytical results will correspond to \( k_{3}^{off} = 0 \), while simulation results will be presented also for \( k_{3}^{off} > 0 \). Analytical results for \( k_{3}^{off} > 0 \) will be published elsewhere.

\textit{Steady State.}—We first consider the statistics of \( n \) in the steady state and compute the long time, joint probability \( p_s(n) \) that the switch is in state \( s \) and there are \( n \) molecules of \( R_1 \). The system of birth–death equations for \( p_{on}(n) \) and \( p_{off}(n) \) becomes in the steady state

\[
(n + 1)k_{1}p_{on}(n + 1) + k_{2}p_{on}(n - 1) + k_{4}^{off}p_{off}(n) = (nk_{1} + k_{2}n^{on} + k_{3}^{on})p_{on}(n), \tag{2a}
\]

\[
(n + 1)k_{1}p_{off}(n + 1) + nk_{3}^{off}p_{on}(n) + k_{4}^{on}p_{on}(n) = (nk_{1} + k_{4}^{off})p_{off}(n). \tag{2b}
\]

In order to decouple the above set of equations, we solve (2a) for \( p_{off} \), then insert the result into (2b) to give a decoupled equation for \( p_{on} \). Introducing the generating function \( G_{on}(z) = \sum_{n} z^{n}p_{s}(n) \) (where \( s = \{\text{on, off}\} \)), the decoupled equation for \( p_{on} \) reduces to a second order differential equation for \( G_{on} \).

Defining then a new variable \( u \equiv u_{on} = k_{2}z/(k_{1} + k_{3}^{on}) - k_{1}k_{2}/(k_{1} + k_{3}^{off})^{2} \), the latter equation reads:

\[
u G_{on}'(u) + (a - u)G_{on}(u) - bG_{on}(u) = 0,
\]

where \( a = 1 + u_{1} + (k_{4}^{off} + k_{3}^{off})/(k_{1} + k_{3}^{on}) \) and \( b = 1 + k_{4}^{off}/k_{1} \). Expanding the solution as a regular power series (i.e. \( G_{on}(u) = \sum_{n}a_{n}u^{n} \)), one finds that \( G_{on}(z) = a_{0}F_{1}(b, a, u) \), where \( F_{1} \) is a confluent hypergeometric function and \( a_{0} \) is an integration constant which can be determined with the normalisation condition \( \sum_{n}p_{on}(n) + p_{off}(n) = 1 \) (or equivalently \( G_{on}(1) = G_{off}(1) = 1 \)). This result can be rewritten in terms of the original variable \( n \) as

\[
p_{on}(n) = a_{0}(u_{1} - u_{0})^{n}(b)_{n}/(a)_{n}F_{1}(b + n, a + n, u_{0}) \tag{4}
\]

where \((a)_{n} = (a + 1) \ldots (a + n - 1)\). An expression for \( p_{off} \) can be derived by inserting Eq. (4) into Eq. (2a). In Fig. 2 we compare the result for \( p(n) = p_{on} + p_{off} \) to simulations, obtaining perfect agreement. Figure 2 also illustrates the effects of the different timescales for switch flipping and production/decay of \( R_1 \). For small \( k_{4} \), switch flipping is slow compared to the rate of change of \( n \). In this case, when the switch is in the on state, the number of recombinase has time to reach a plateau before the switch flips off. The on and off switch states are then each associated with a different value of \( n \) and the distribution \( p(n) \) is bimodal. In contrast, when \( k_{4} \) is large, the switch flips back and forth much more rapidly than recombinase production or removal. Then, the fraction of time spent in the on state is \( k_{4}^{off}/(k_{4}^{off} + k_{3}^{off}) \), and \( p(n) \) tends to the Poisson distribution expected for a birth–death process with birth rate \( k_{4}^{off}/(k_{4}^{off} + k_{3}^{off}) \) and death rate \( k_{1} \).
Flipping time distributions.—To determine how orientational control affects switch function, we compute flipping time distributions. The flipping time $T$ can be defined in two different ways. In the first scenario, which we call the Switch Change Ensemble (SCE), we define $T$ as the time spent in a particular switch state—for example, $F_{\text{on}}^{\text{SCE}}(T)$ is the probability distribution for the time between the moment the switch enters the on state and the moment it flips from the on to the off state. In the second scenario, which we call the Steady State Ensemble (SSE), we start observing the cell at a random moment and measure the time interval between this moment and its next flip into the other state. $F_{\text{on}}^{\text{SSE}}(T)$ and $F_{\text{off}}^{\text{SSE}}(T)$ may be relevant to the response of a population of switching cells to a sudden environmental change. They also correspond to an experiment where one measures the time until the next flip, for cells sampled in the steady state [12].

To compute these distributions, we define $F_{\text{on}}(n|0)$ as the probability that the system begins at $t = 0$ in the $s$ state with $n$ recombinase and flips for the first time at $T$. Note that for $k_{\text{off}}^3 = 0$, the off to on flipping process does not depend on $R_1$ and is governed by $k_{\text{off}}^3$; thus $F_{\text{off}}(T) = k_{\text{off}}^3 \exp(-k_{\text{off}}^3 T)$ is independent of $n_0$ (for both the SCE and the SSE). However, the on to off flipping rate is $n_0$ dependent, so that we average over the ensemble of initial states (characterised by the probability $W_{\text{on}}(n)$ of having $n$ recombinase at the start of our measurement) to obtain the flip time distribution $F_{\text{on}}(T) = \sum_{n_0} W_{\text{on}}(n_0) F_{\text{on}}(T|n_0)$.

For the SCE, the initial condition is taken just after a flip, which implies that $W_{\text{on}}^{\text{SCE}}(n_0) = p_{\text{off}}(n_0)/G_{\text{on}}(1)$. For the SSE, the initial condition is sampled in the steady state, yielding $W_{\text{on}}^{\text{SSE}}(n_0) = p_{\text{on}}(n_0)/G_{\text{off}}(1)$. To compute $F_{\text{on}}(T)$, we first define the survival probability $h_{\text{on}}(n, t)$ that, at time $t$, the switch is in the on state with $n$ $R_1$ molecules, without having flipped, given the initial condition $h_{\text{on}}(n_0, 0) = W_{\text{on}}(n_0)$. The evolution equation for $h_{\text{on}}$ is:

$$\partial_t h_{\text{on}}(n, t) = (n + 1)k_1 h_{\text{on}}(n + 1, t) + k_2 h_{\text{on}}(n - 1, t) - (nk_1 + k_2 + n k_{\text{on}}^3 + k_{\text{off}}^3) h_{\text{on}}(n, t).$$

(5)

Defining a generating function $\tilde{h}_{\text{on}}(z, t) = \sum_n z^n h_{\text{on}}(n, t)$, it follows that $F_{\text{on}}(t) = -\partial_z \tilde{h}_{\text{on}}(1, t)$. Eq. (5) reduces to a partial differential equation for $\tilde{h}_{\text{on}}(z, t)$ which has the initial condition $\tilde{h}_{\text{on}}(z, 0) = \sum_{n_0} W_{\text{on}}(n_0) z^n$. Its solution is

$$\tilde{h}_{\text{on}}(z, t) = e^{-z(k_{\text{on}}^3 + k_2(1 - k_1 \tau_\text{on}))} k_2 \tau_\text{on} (z - k_1 \tau_\text{on}) (1 - e^{-z/k_{\text{on}}^3})$$

$$\times \tilde{h}_{\text{on}}(k_1 \tau_\text{on} + e^{-z/k_{\text{on}}^3} (z - k_1 \tau), 0),$$

(6)

where $\tau_\text{on} = (k_1 + k_{\text{on}}^3)^{-1}$. $F_{\text{on}}(T)$ can then be computed for the different measurement scenarios using $\tilde{h}_{\text{on}}(x, 0) = G_{\text{on}}(x)/G_{\text{on}}(1)$ with $s = \text{off}$ for SCE and $s = \text{on}$ for SSE.

Peak in the distribution.—Our results, illustrated in Fig. 3, show a striking effect of orientational control for this model switch: for the SCE, we can obtain a peak in the flipping time distribution. Such a peak has been postulated for the fim switch [9], where it might imply that the switch tends not to leave the on state before it has had time to synthesise fimbriae [3, 9, 10]. Time spent in the on state may also influence recognition by the host immune system. This peak in $F_{\text{on}}^{\text{SCE}}$ is a consequence of the feedback between the switch state and the level of $R_1$. Once in the on state, the rate of on to off flipping increases with time as $R_1$ is produced. In contrast, for a Poissonian switch, the rate of flipping is constant in time. For a peak to occur, the slope of $F_{\text{on}}^{\text{SCE}}$ at the origin must be positive, which implies

$$k_2 - \langle k_{\text{on}}^3 \rangle^2/k_{\text{on}}^3 - (k_1 + 2k_{\text{on}}^3) \langle n_0 \rangle - k_{\text{on}}^3 \langle n_0^2 \rangle > 0,$$

(7)

where $\langle \ldots \rangle$ denotes an average using the weight $W_{\text{on}}$. The l.h.s. of (7) can be evaluated numerically using the exact result [4] to compute $\langle n_0 \rangle$ and $\langle n_0^2 \rangle$. We can then determine the regions of parameter space where a peak exists, as shown in the shaded region of Fig. 3 for the SCE. Our results show that the presence of a peak isfavoured by large values of $k_2$ (strong production of $R_1$ in the on state), and suppressed by very large values of $k_{\text{on}}^3$ (strong $R_1$—mediated on to off switching) or by very small values of $k_{\text{on}}^3$ (switching dominated by $R_1$—independent mechanism). Likewise, when $k_3$ is increased, the range of values over which the peak exists is decreased, since $R_1$—independent Poissonian switching tends to dominate.

For the SSE, on the other hand, we did not find any parameter values where inequality (7) is verified. The peak in the SCE appears because the number of recombinase, and hence the flipping probability, is typically low immediately after entering the on state and increases significantly thereafter. In contrast, in the SSE one typically starts a measurement when $n$, and hence the flipping probability, is already high. This tends to suppress the peak in the SSE flipping time distribution.

Correlated flips.—Another potentially important effect of the feedback between the switch state and the production of recombinase $R_1$ may be to cause correlations in the waiting
times between successive flips (for example, a particularly short time before a flip might lead to subsequent flips occurring in quick succession). Such correlations might allow a population of switching microbes to “remember” the history of past environmental changes. We define the switch period \( \tau_i \) as the time from when the switch enters the on state from the off state for the \( i \)th time, until it enters the on state for the \((i+1)\)th time [cf. Fig. 1]. In Fig. 3 we plot simulation results for the correlation function for switch periods \( \tau_i \) and \( \tau_j \), as a function of \( j - i \). For \( k^{\text{off}}_3 = 0 \), weak correlation is observed between subsequent periods \( \tau_i \) and \( \tau_{i+1} \). Correlations are weak because when \( k^{\text{off}}_3 = 0 \), the off to on switching process does not depend on \( R_1 \) and is an uncorrelated Poisson process. When \( k^{\text{off}}_3 \neq 0 \), correlations are much stronger and extended correlated sequences of flips emerge.

Discussion.—We have presented a generic model for a DNA inversion switch with orientational control. By solving the model analytically in the case \( k^{\text{off}}_3 = 0 \) (relevant to the \( \text{fim} \) switch), and using stochastic simulations, we have shown that this type of switch can display markedly non–Poissonian behaviour, including a peaked flipping time distribution for intermediate values of \( k^{\text{off}}_3 \) and, for \( k^{\text{off}}_3 > 0 \), correlated sequences of flips. Non–Poissonian behaviour has been postulated to be a consequence of orientational control [3, 9]. The model presented here allows us to analyse the origins and effects of this behaviour in detail, and provides analytical results which can be used as a basis for more complex models [13]. Suggested evolutionary roles for orientational control include rapid response to environmental change [10], as well as peaked flipping time distributions [3, 9]. This study raises interesting questions about the consequences of non–Poissonian switching for population dynamics in changing environments. Several models have been proposed for the growth of populations of switching cells in stochastically and periodically changing environments (see, for example [14]). These models assume Poissonian switch flipping. The analytical solutions presented here should make it possible to extend such models to the case of non-Poissonian flips. Non-Poissonian switch flipping opens up the possibility that lineages of cells may ‘remember’ (in a statistical sense) the history of their recent phenotypic states. This is likely to have important consequences for models which include selection according to the fitness of different phenotypic states in changing environments. We speculate that bacteria which adopt a non-Poissonian flipping strategy may be able to maximise the evolutionary advantages of using some knowledge of the likely future behaviour of the environment, combined with the benefits of a stochastic strategy as an insurance against sudden and unpredictable environmental changes. These avenues will be the subject of future work.

The authors are grateful to A. Adiciptaningrum, G. Blakely, M. Cates, C. Dorman, A. Free, D. Gally, N. Holden, S. Tănase-Nicola and S. Tans for valuable discussions. R.J.A. was funded by the RSE. This work was funded by the EPSRC under grant EP/E030173.

[1] M. W. van der Woude and A. Bäumler, Clin. Microbiol. Reviews 17, 581 (2004).
[2] M. W. van der Woude, FEMS Microbiol. Lett. 254, 190 (2006).
[3] I. C. Blomfield, Adv. Microb. Physiol. 45, 1 (2001)
[4] H. D. Kulasekara et al., Mol. Microbiol. 31, 1171 (1999).
[5] S. A. Joyce and C. J. Dorman, Mol. Microbiol. 45, 1107 (2002).
[6] D. Dubnau and R. Losick, Mol. Microbiol. 61, 564 (2006).
[7] J. L. Cherry and F. R. Adler, J. Theor. Biol. 203, 117 (2000).
[8] P. B. Warren and P. R. ten Wolde, J. Phys. Chem. B 109, 6812 (2005); Phys. Rev. Lett. 92, 128101 (2004).
[9] D. W. Wolf and A. P. Arkin, OMICS 6, 91 (2002).
[10] D. Chu and I. C. Blomfield, J. Theor. Biol. 244, 541 (2007).
[11] D. T. Gillespie, J. Comput. Phys. 22, 403 (1976); A. B. Bortz et al., J. Comput. Phys. 17, 10 (1975).
[12] Experiments are likely to be complicated by division events between flips (not considered here).
[13] N. J. Holden and D. L. Gally, J. Med. Microbiol. 53, 585 (2004).
[14] M. Thattai and A. van Oudenaarden, Genetics 167, 523 (2004).