JOURNAL OF THE ROYAL SOCIETY Interface

Elucidating the digital control mechanism for DNA damage repair with the p53 –Mdm2 system: single cell data analysis and ensemble modelling

Babatunde A Ogunnaike

J. R. Soc. Interface 2006 **3**, 175-184
doi: 10.1098/rsif.2005.0077

JOURNAL OF THE ROYAL SOCIETY Interface

J. R. Soc. Interface (2006) **3**, 175–184
doi:10.1098/rsif.2005.0077
Published online 19 September 2005

Elucidating the digital control mechanism for DNA damage repair with the p53–Mdm2 system: single cell data analysis and ensemble modelling

Babatunde A. Ogunnaike[†]

Department of Chemical Engineering, University of Delaware, Newark, DE 19716, USA

Recent experimental evidence about DNA damage response using the p53–Mdm2 system has raised some fundamental questions about the control mechanism employed. In response to DNA damage, an ensemble of cells shows a damped oscillation in p53 expression whose amplitude increases with increased DNA damage—consistent with 'analogue' control. Recent experimental results, however, show that the single cell response is a series of discrete pulses in p53; and with increase in DNA damage, neither the height nor the duration of the pulses change, but the mean number of pulses increase—consistent with 'digital' control. Here we present a system engineering model that uses published data to elucidate this mechanism and resolve the dilemma of how digital behaviour at the single cell level can manifest as analogue ensemble behaviour. First, we develop a dynamic model of the p53–Mdm2 system that produces non-oscillatory responses to a stress signal. Second, we develop a probability model of the distribution of pulses in a cell population, and combine the two with the simplest digital control algorithm to show how oscillatory responses whose amplitudes grow with DNA damage can arise from single cell behaviour in which each single pulse response is independent of the extent of DNA damage. A stochastic simulation of the hypothesized control mechanism reproduces experimental observations remarkably well.

Keywords: DNA damage response; p53–Mdm2 system; control systems engineering; systems biology; statistical data analysis; dynamic modelling

1. INTRODUCTION

The p53–Mdm2 system, which plays a crucial role in DNA damage repair, is one of the best-studied of the 'negative feedback motifs' known to be present in human cells (Piette *et al.* 1997; Vogelstein *et al.* 2000; Michael & Oren 2003). Studies of these systems typically involve perturbing cell populations with appropriate stimuli and monitoring total population response with immunoblots. Often such measurements of ensemble (i.e. population) behaviour are sufficient for understanding the molecular mechanisms underlying the phenomenon in question. In the case of the p53–Mdm2 system response to DNA damage however, Lahav *et al.* (2004), recently published experimental evidence that the dynamic behaviour of the ensemble is fundamentally different from that of individual cells, creating a dilemma about the underlying control system mechanism.

Specifically, in response to DNA damage, the observed ensemble response is a damped oscillation in p53 levels whose amplitude increases with increased DNA damage. This behaviour is consistent with analogue control and a few theoretical models are available that predict it reasonably well (e.g. Bar-Or *et al.* 2000; Mihalas *et al.* 2000). However, the data in Lahav *et al.* (2004), show that at the single cell level, the response to DNA damage is rather a series of discrete pulses in p53; furthermore, with increased DNA damage, neither the mean height nor the duration of the pulses changed, but the mean number of pulses increased. In addition, genetically identical cells in a population exposed to the same stimulus each showed a different number of pulses of p53. Taken together, the observed single cell behaviour is consistent with digital control, raising the obvious question: how can digital behaviour at the single cell level appear analogue at the ensemble level (Lahav *et al.* 2004)? The more fundamental issue concerning the underlying DNA damage response mechanism is captured by the following challenge (Lahav *et al.* 2004):

> What is the mechanism for digital oscillations in this system? Digital undamped oscillatory behaviour is a challenge to modelers because the simplest theoretical models of this negative feedback loop show damped analogue oscillations.

[†]Author for correspondence (ogunnaike@che.udel.edu).

One contribution of 8 to a themed supplement 'Statistical mechanics of molecular and cellular biological systems'.

Received 11 April 2005
Accepted 3 August 2005

We present here a comprehensive, systems engineering model that uses the Lahav data to elucidate this mechanism and resolve the dilemma.

The systems engineering paradigm we employ considers the entire DNA damage response apparatus as consisting of the following components:

(i) the DNA damage 'process' itself, by which, in response to a damage inducing 'input' signal, DNA is damaged to an extent dependent on the strength (magnitude) of the signal;
(ii) a sensor for detecting the presence of DNA damage;
(iii) a controller that is activated by the presence of DNA damage, responding by generating a digital command signal $c(t)$ that is implemented by
(iv) an effector system (the p53–Mdm2 system) which converts the controller signal $c(t)$ to the amount of p53 expressed in the cell.

A block diagram for the overall system (with the details to be discussed fully later) is shown in figure 1.

The results presented have been obtained by first developing appropriate models for the two central components of this system: (i) a dynamic model for the p53–Mdm2 effector system that reproduces the experimentally observed single non-oscillatory response to a stress signal; and (ii) a probabilistic model of the DNA damage response derived from a careful analysis of the Lahav data on the distribution of pulses observed in a cell population. Each model is analysed for insight before combining them with the simplest possible digital control algorithm into a complete control system. The overall control system is then shown, via simulation, to reproduce experimental observed behaviour quite well.

2. RESULTS AND ANALYSES

2.1. The p53–Mdm2 effector system model

The model derived here is an adaptation of the one presented in Bar-Or *et al.* (2000), modified as follows for simplicity and transparency.

(i) *p53 dynamics*: we explicitly separate out the 'stress response' signal to allow it to be determined by an explicit controller, not implicitly by the generic first order differential equation used to represent p53 activation upon exposure to DNA damage (eqn 5 of Bar-Or *et al.* 2000). This aspect is crucial for explaining the Lahav *et al.* observation (Lahav *et al.* 2004).
(ii) *Mdm2 dynamics*: the sequence of three equations employed in Bar-Or *et al.* (2000), necessitating the need for an 'intermediate' variable, are eliminated in favour of the simpler single differential equation incorporating a straightforward (gene-specific) time delay between a change in transcription rate and the corresponding change in the protein level (a sum total of the time required for completing the processes of transcription, mRNA processing, translation, and protein processing). This has become standard in many models of feedback inhibition of gene expression that have appeared since Bar-Or *et al.* (Bar-Or *et al.* 2000; Tiana *et al.* (2002); Monk 2003). The (gene-specific) delay may be estimated directly using genomic sequences, or else from data. In this case, as stated explicitly in Lahav *et al.* (2004), the composite time delay is ~100 min.
(iii) *Damage-induced activation*: eqs 6 and 7 in Bar-Or *et al.* (2000) introduced as a means of 'hard-coding' the stress response are eliminated. In our scheme, such activation is produced by the controller mechanism.

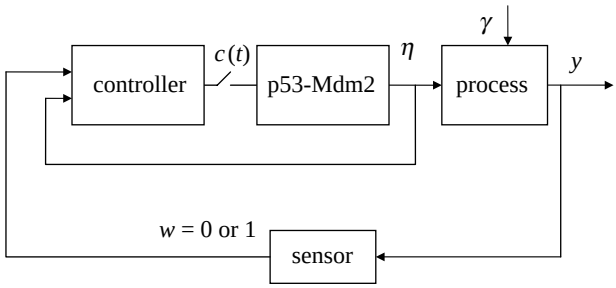

Figure 1. Postulated control system block diagram for the overall single cell DNA damage repair system.

On the basis of the foregoing, in the initial basal, unstressed state prior to sensing damage, the p53–Mdm2 system dynamics are represented by

$$\frac{\mathrm{d}C_{\mathrm{p}}^{0}}{\mathrm{d}t} = S_{\mathrm{p}} - k_{d_{\mathrm{p}}} C_{\mathrm{p}}^{0} - k_{\mathrm{c}}^{0} C_{\mathrm{p}}^{0} C_{\mathrm{m}}^{0}, \tag{2.1}$$

$$\frac{\mathrm{d}C_{\mathrm{m}}^{0}}{\mathrm{d}t} = S_{\mathrm{m}} - k_{d_{\mathrm{m}}} C_{\mathrm{m}}^{0} + k_{\mathrm{tt}} C_{p}^{0}(t-\tau). \tag{2.2}$$

Here $C_{\mathrm{p}}^{0}, C_{\mathrm{m}}^{0}$ represent respective basal concentrations of the proteins p53 and Mdm2, with respective steady state values $C_{\mathrm{p}}^{0*}, C_{\mathrm{m}}^{0*}$; S_{p}, S_{m} represent respective synthesis rates; $k_{d_{\mathrm{p}}}, k_{d_{\mathrm{m}}}$ are the respective natural degradation rate constants for each species. The final term in equation (2.1) represents Mdm2 interaction with p53, targeting it for ubiquitin-mediated degradation at a rate determined by k_{c}^{0}. The final term in equation (2.2) represents the p53-dependent transcription and translation of Mdm2; τ is the delay in the p53-dependent induction of Mdm2. (We opt for this simpler expression on the basis of the fact that by the very design of the p53-Mdmd2 system, p53 is not likely ever to be induced to the point where it will saturate the Mdm2 promoter.)

After stress, there is a dramatic disruption of the ability of Mdm2 to target p53 for degradation (effectively a reduction in k_{c}^{0}) resulting in a net decrease in the activity of Mdm2 in terms of its interaction with p53 (Lakin & Jackson 1999). Under these conditions, the model becomes

$$\frac{\mathrm{d}C_{\mathrm{p}}}{\mathrm{d}t} = S_{\mathrm{p}} - k_{d_{\mathrm{p}}} C_{\mathrm{p}} - k_{\mathrm{c}} C_{\mathrm{p}} C_{\mathrm{m}}, \tag{2.3}$$

$$\frac{\mathrm{d}C_{\mathrm{m}}}{\mathrm{d}t} = S_{\mathrm{m}} - k_{d_{\mathrm{m}}} C_{\mathrm{m}} + k_{\mathrm{tt}} C_{\mathrm{p}}(t-\tau). \tag{2.4}$$

By defining deviation variables

$$x_1 = C_{\mathrm{p}} - C_{\mathrm{p}}^{0*}, \tag{2.5}$$

$$x_2 = C_{\mathrm{m}} - C_{\mathrm{m}}^{0*}, \tag{2.6}$$

$$\Delta k_{\mathrm{c}} = k_{\mathrm{c}} - k_{\mathrm{c}}^{0}, \tag{2.7}$$

in terms of deviations from steady state basal, unstressed conditions, upon introducing a Taylor approximation for the single bilinear term in equation (2.3), and then subtracting appropriate corresponding equations given above, we obtain

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -(k_{d_{\mathrm{p}}} + k_{\mathrm{c}} C_{\mathrm{m}}^{0})x_1 - k_{\mathrm{c}} C_{\mathrm{p}}^{0} x_2 + f(t), \tag{2.8}$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = k_{\mathrm{tt}} x_1(t-\tau) - k_{d_{\mathrm{m}}} x_2. \tag{2.9}$$

Here the forcing function $f(t)$, defined by

$$f(t) = \Delta k_{\mathrm{c}} C_{\mathrm{m}}^{0} C_{\mathrm{p}}^{0}, \tag{2.10}$$

represents the *net effect* of the control mechanism's command signal activated by DNA damage.

We may now use this model to investigate the system response to short pulses of stress signals that perturb the system away from its initial operating point. We stress that the Taylor linearizing approximation used here is not necessary if one simply wants to obtain *specific* simulations of system responses. But the approximation provides invaluable insight into the *general* dynamic characteristics of the system of equations, albeit at the modest cost of some approximation error. This error is quite minimal, however, because first, we are investigating perturbations around the basal steady state caused by stress pulses. Furthermore, all higher order univariate partial derivatives in the truncation error associated with linear approximations to bilinear terms are identically zero, leaving only the sole surviving bivariate second order derivative, which in this case is a constant.

2.1.1. Theoretical analysis and general response characteristics. By defining (for notational simplicity)

$$\left.\begin{aligned} k_1 &= k_{d_{\mathrm{p}}} + k_{\mathrm{c}} C_{\mathrm{m}}^{0}, \\ k_2 &= k_{\mathrm{c}} C_{\mathrm{p}}^{0}, \\ k_3 &= k_{\mathrm{tt}}, \\ k_4 &= k_{d_{\mathrm{m}}} \end{aligned}\right\} \tag{2.11}$$

upon taking Laplace transforms, we obtain, from equations (2.8) and (2.9), the following transfer function models (Ogunnaike & Ray 1994; ch. 4)

$$x_1(s) = \frac{(s+k_4)}{(s+k_1)(s+k_4) + k_2 k_3 \mathrm{e}^{-\tau s}} f(s), \tag{2.12}$$

for how p53 (x_1) ultimately responds to the control stimulus f; and

$$x_2(s) = \frac{k_3 \mathrm{e}^{-\tau s}}{(s+k_4)} x_1(s), \tag{2.13}$$

for how Mdm2 (x_2) responds to changes in p53 (x_1).

(i) *Steady state step response*: in the (unlikely) event that the forcing function $f(t)$ is a step function of magnitude A (corresponding to a sustained activation of p53) the model indicates that the final value attained by the p53 level (as a deviation from the basal level) is given by

$$x_1^* = \left\{\frac{k_4}{k_1 k_4 + k_2 k_3}\right\} A. \tag{2.14}$$

Observe that if k_4 is small while k_1 is large, (possibly with k_2 and/or k_3 also relatively large), x_1^* will always be small even in the face of sustained activation, and the effect of such a drastic disturbance will be mitigated. This, of course, underscores the essence of the role of Mdmd2 in tightly regulating p53 levels (Momand *et al.* 2000). Thus the 'design' criteria for effective regulation of p53 (protecting the cell from unintended step changes in p53 level) are as follows:

(a) keep k_4 low (stable, longer half-life for Mdm2);
(b) keep k_1 large (short half-life for p53, in conjunction with a high basal level of Mdm2, and high initial rate of Mdm2 binding to p53);
(c) keep k_3 and/or k_2 relatively large (high rate of p53-dependent transcription and translation of Mdm2; and/or high rate of ubiquitin-mediated degradation of p53).

These conditions are all perfectly in keeping with what is known of this system (Lakin & Jackson 1999).

(ii) *Conditions for overdamped (non-oscillatory) response*: in the undelayed case, it is easy to show (Ogunnaike & Ray 1994) that the second order transfer function model will exhibit a non-oscillatory response when

$$k_2 k_3 < \tfrac{1}{4}(k_1 - k_4)^2. \tag{2.15}$$

With the delay, the transcendental function $\mathrm{e}^{-\tau s}$ contributes an infinite number of poles (Ogunnaike & Ray 1994, pp. 495–496) that can only be computed explicitly for specific transfer function parameters (without the benefit of a closed-form expression).

The condition for obtaining (underdamped) oscillatory response follows so that, for the undelayed case,

$$k_2 k_3 > \tfrac{1}{4}(k_1 - k_4)^2. \tag{2.16}$$

2.1.2. Single cell response. Figure 2 shows the predicted single cell response to a forcing function $f(t)$ realized as a rectangular pulse of unit height, duration 250 min (Lahav *et al.* 2004), passed through a first order filter (time constant 40 min) to smooth out the sharp transitions of the rectangular pulse and yield a more physiologically realistic stress signal (Lakin & Jackson 1999). The parameters are $k_1 = 0.1$, $k_2 = 0.004$, $k_3 = 0.065$, $k_4 = 0.05$ (satisfying the conditions required for non-oscillatory response; the same combination of parameters but with higher values for k_1 give similar responses). The delay is chosen as $\tau = 100$ min, in accordance with the experimental results in Lahav *et al.* (2004).

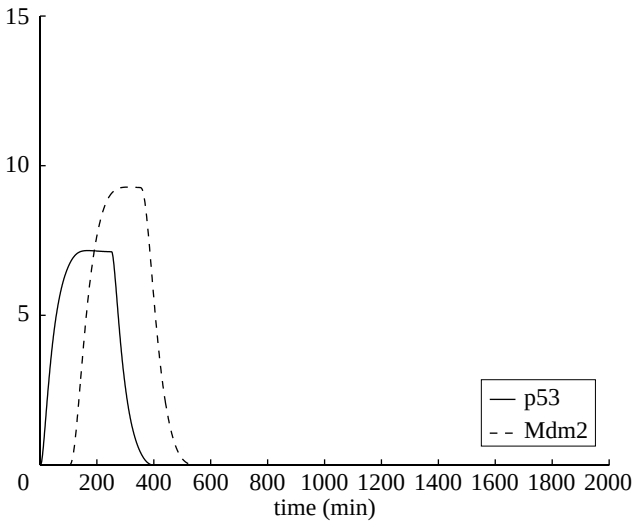

Figure 2. Single cell response of p53 and Mdm2 to a single stress signal pulse; as deviations from basal levels; in arbitrary units.

(If the value chosen for k_2 were increased by an order of magnitude to 0.04—satisfying the conditions required for oscillations—oscillatory responses (not shown) are indeed obtained. The model is therefore sufficiently general to obtain various types of responses and has not been artificially constrained to produce only non-oscillatory responses.)

As discussed more fully later, consider a control strategy in which each time it is activated by DNA damage, the controller in figure 1 responds according to the single p53 pulse response obtained above, which we will represent as $\eta(t)$ (with a pulse width, w). If, after a period T, the single p53 pulse is sufficient to accomplish complete damage repair, the controller remains inactive thereafter; if DNA damage persists, the controller responds with yet another 'identical' pulse, $\eta(t)$, which now appears T time units after the previous one. Thus, with extensive damage, the controller response will be a sequence of several such pulses, each one appearing T time units after the previous one, with the total number dependent on the extent of damage.

Observe therefore that, mathematically, we may use $\eta(t)$ as the 'basis function' to derive a compact expression for the controller response in terms of a sequence of j such pulses (represented as $\eta_j(t)$) as follows:

$$\eta_1(t) = \eta(t) \tag{2.17}$$

$$\eta_2(t) = \eta(t) + \eta(t-T) \tag{2.18}$$

$$\vdots \qquad \vdots$$

$$\eta_j(t) = \eta(t) + \eta(t-T) + \eta(t-2T) + \ldots + \eta(t-(j-1)T) \tag{2.19}$$

or

$$\eta_j(t) = \sum_{i=1}^{j} \eta(t-(i-1)T), \tag{2.20}$$

where the indicated delayed functions are defined in the usual manner for a delay α,

$$h(t-\alpha) = \begin{cases} 0, & t < \alpha, \\ h(t-\alpha), & t \geq \alpha. \end{cases} \tag{2.21}$$

Observe that for the pulses to be separate and distinct it is necessary that $T \geq w$. In fact, in Lahav *et al.* (2004), the average experimentally observed pulse width, w, is 350 min (± 160); the average period, T is 440 min (± 100). The pulse width of the response shown in figure 2 is ≈ 400 min.

2.2. DNA damage response model

Let Y represent the (unobserved) measure of the 'extent of DNA damage' resulting from exposing an ensemble of cells to gamma-irradiation of dose γ Gy; and let Z be the corresponding number of p53 pulses observed as a consequence. From the experimental results in Lahav *et al.* (2004), Z is clearly dependent on Y; also genetically identical cells exposed to the same stimulus show different numbers of pulses. Thus, both Z and Y are *random variables*, so that the specific values z and y pertaining to *each* cell in the ensemble will differ from cell to cell.

It is customary to describe the ensemble by probability distribution functions $f_z(z)$ and $f_y(y)$, which, respectively, quantify the probability that $Z=z$ and $Y=y$ (i.e. $f_z(z)=\text{Prob}(Z=z)$; $f_y(y)=\text{Prob}(Y=y)$).

Table 1 contains data (extracted from Lahav *et al.* 2004) on the fraction of cells showing $z=0$, 1, 2 or more pulses, for various irradiation doses. This is, in fact, an empirical probability distribution of the number of p53 pulses observed in the ensemble of cells used in their experimental system.

The question of interest to us here may now be stated as follows:

> Given experimental values for the distribution of Z, the number of p53 pulses resulting from exposing an ensemble of cells to various irradiation doses, γ, what is the distribution of Y, the unobserved extent of DNA damage, and how is it related to γ?

We are able to show that a Poisson distributed Y with the probability distribution function

$$f_y(y) = \frac{\lambda^y e^{-y}}{y!} \tag{2.22}$$

that is related to the random variable Z as follows,

$$\begin{aligned} z = 0 \quad &\text{if} \quad y = 0, \\ z = 1 \quad &\text{if} \quad y = 1 \text{ or } 2, \\ z = 2 \quad &\text{if} \quad y = 3 \text{ or } 4, \\ \ldots \quad &\qquad \ldots \\ z = n \quad &\text{if} \quad y = (2n-1) \text{ or } (2n), \end{aligned} \tag{2.23}$$

matches the observed data remarkably well, with the Poisson parameter λ dependent on the irradiation dose, γ, according to the power law relationship

$$\lambda = a\gamma^b. \tag{2.24}$$

2.2.1. Data analysis and DNA damage response model. For each radiation dose, we obtain a maximum likelihood estimate of the Poisson parameter λ such that, from the resulting $f_y(y|\lambda)$, the predicted distribution $\hat{f}_z(z)$, obtained from the definition given in

Table 1. Empirical distribution of the number of p53 pulses for various irradiation doses (Lahav *et al.* 2004).

z	$f_z(z)$ $\gamma=0$ Gy	$f_z(z)$ $\gamma=0.3$ Gy	$f_z(z)$ $\gamma=2.5$ Gy	$f_z(z)$ $\gamma=10$ Gy
0	0.95	0.25	0.12	0.05
1	0.05	0.60	0.53	0.35
2+	0.00	0.15	0.35	0.6

equation (2.23), i.e.

$$\begin{aligned} \hat{f}_z(z=0) &= f_y(y=0) \\ \hat{f}_z(z=1) &= f_y(y=1)+f_y(y=2) \\ \hat{f}_z(z=2) &= f_y(y=3)+f_y(y=4) \\ \cdots &= \cdots \\ \hat{f}_z(z=n) &= f_y(y=2n-1)+f_y(y=2n) \end{aligned} \tag{2.25}$$

optimally matches $f_z(z)$, the experimentally observed distribution of p53 pulses. Because of the 'lookup-table-like' form of equation (2.25), there is no closed form analytical expression relating z to y; the estimation is therefore carried out numerically.

The results of the analysis using the data in table 1 are summarized as follows: (i) table 2 shows estimates obtained for the Poisson variable λ as a function of the irradiation dose γ; (ii) appendix A shows a comparison of the predicted $\hat{f}_z(z)$ with the corresponding observed experimental values; and (iii) a regression analysis of the data in table 2 is shown in figure 3 along with the resulting regression equation (and some goodness-of-fit metrics), establishing the power law relationship in equation (2.24).

The implications of these results are as follows.

(i) Following irradiation of dose γ Gy, the extent of DNA damage experienced by a cell is characterized by a random variable Y possessing a Poisson distribution with parameter λ, the ensemble mean extent of DNA damage.

(ii) λ depends on radiation dose according to the power-law relationship in equation (2.24); in the specific case of the experimental system studied in Lahav *et al.* (2004), the estimated values are $a=10^{0.192}$, $b=0.365$.

(iii) $\hat{f}_z(z)$, the distribution of p53 pulses predicted from the Poisson model $f_y(y)$, in conjunction with the Y-to-Z relationship not only matches the Lahav data remarkably well; it also allows an extrapolation that indicates what fraction of cells would have shown 3 or more pulses had the experiment continued long enough. (It was explicitly noted in Lahav *et al.* (2004), that 'some cells began a third pulse towards the end of the movies.')

(iv) We do not currently know precisely what this (unobserved) variable Y is physiologically; but it is unlikely to be a straightforward linear function of the number of double-stranded DNA breaks (DSB). This is because of the well-established fact that the number of DSBs varies linearly with γ-irradiation doses (Lahav *et al.* 2004) whereas, λ, the mean value of the random y varies with dose in the power law form shown earlier.

(v) Whatever the variable Y turns out to be physiologically, statistically, it is Poisson distributed, and the relationship between it and the observed number of pulses, Z, is the 'look-up-table-like' expressions in equation (2.23): no p53 pulses are recorded for zero extent of damage; $z=1$ pulse is observed when this measure of the 'extent of damage' is 1 or 2; $z=2$ pulses are observed when the extent of damage is 3 or 4, etc.

Conversely, this suggests that if the extent of damage variable y is binned as follows: (0), (1,2), (3,4), …, etc., then each pulse of p53 reduces y to the next lower bin. Thus, for example, if y was initally in the (3,4) bin, the first p53 pulse will reduce it to (1,2), and the next one will reduce it to (0), eliminating the damage.

Table 2. Maximum likelihood estimates of the Poisson variable λ (Lahav *et al.* 2004).

Radiation Dose (Gy)	λ
0.00	0.05
0.30	1.30
2.50	2.10
10.00	3.10

We are now in a position to combine the single cell model with the probability model for DNA damage response to obtain the ensemble response.

2.3. Ensemble response model

We are concerned here with deriving an expression for $U(t)$, the ensemble p53 response of a population of N_T cells, each individually showing $i=0,1,2,3,\ldots$ pulses of p53 represented by $\eta_i(t)$ in each case.

Now, let us represent as ϕ_i, the fraction of the total number of cells in the population showing i pulses. (Note that from a classical frequency interpretation of probability, ϕ_i also represents the empirical probability that a particular cell in the ensemble shows exactly i pulses.) For a total number of N_T cells in the population, therefore, the ensemble response will be given by

$$U(t) = N_T \sum_{i=0}^{\infty} \phi_i \eta_i(t) \tag{2.26}$$

(essentially the mathematical expectation of the p53 response function), which expands out to give

$$U(t) = N_T\{\phi_1\eta_1(t)+\phi_2\eta_2(t)+\phi_3\eta_3(t)+\cdots+\}, \tag{2.27}$$

since $\eta_0=0$. Recalling equation (2.20) for $\eta_j(t)$ and further expanding out gives

$$\begin{aligned} U(t) = N_T\{&\phi_1\eta(t) \\ &+\phi_2\eta(t)+\phi_2\eta(t-T) \\ &+\phi_3\eta(t)+\phi_3\eta(t-T)+\phi_3\eta(t-2T) \\ &+\phi_4\eta(t)+\phi_4\eta(t-T)+\phi_4\eta(t-2T) \\ &+\phi_4\eta(t-3T)+\ldots\} \end{aligned} \tag{2.28}$$

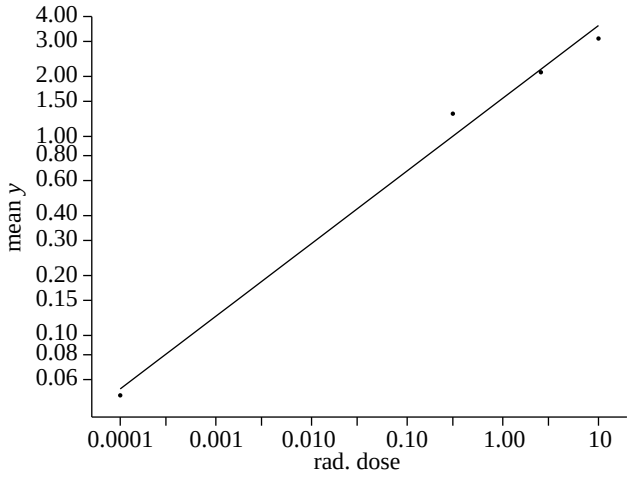

Figure 3. Mean damage response as a function of radiation dose. Log(mean y) = 0.192 + 0.365log(rad. dose). Residual error mean sum of squares, $S = 0.096$; goodness-of-fit measures, $R^2 = 99.1\%$ and $R^2(\text{adj}) = 98.6\%$. These values imply a good model fit to the data without overparametrization (Draper & Smith 1981).

or

$$U(t) = N_T\{\bar{\Phi}_1\eta(t) + \bar{\Phi}_2\eta(t-T) + \bar{\Phi}_3\eta(t-2T) + \ldots\}, \tag{2.29}$$

where, since $\sum_{i=0}^{\infty}\phi_i = 1$ (the probabilities of all possible outcomes sum to 1), the indicated $\bar{\Phi}_i$ coefficients are defined as follows,

$$\begin{aligned}\bar{\Phi}_1 &= \phi_1 + \phi_2 + \phi_3 + \ldots = 1 - \phi_0,\\ \bar{\Phi}_2 &= \phi_2 + \phi_3 + \ldots = 1 - \phi_0 - \phi_1,\\ \ldots &= \ldots\end{aligned}$$

so that, in general

$$\bar{\Phi}_j = \sum_{i=j}^{\infty}\phi_i = 1 - \sum_{i=0}^{j-1}\phi_i \tag{2.30}$$

is the fraction of cells with j or more pulses, satisfying the recursion

$$\bar{\Phi}_{j+1} = \bar{\Phi}_j - \phi_j \tag{2.31}$$

with the initial condition $\bar{\Phi}_0 = 1$.

It is now most important to observe that $\bar{\Phi}_i$ is a *monotonically decreasing* sequence, with the following implications.

(i) From (2.29) the resulting ensemble p53 response will be a sequence of pulses $\eta(t)$ repeated at a period T, *but with decreasing amplitude* $\bar{\Phi}_i$ dictated by the extent of damage—giving the appearance of a damped oscillation.

(ii) With low damage, the distribution of ϕ_i is skewed heavily to the low end (mostly ϕ_0), with most cells showing few, if any, pulses; as damage increases, the distribution shifts to the right with more cells experiencing more damage and hence showing more pulses (see Lahav *et al.* 2004, fig. 3*a*).

(iii) Furthermore, the coefficients, $\bar{\Phi}_i$, represent the fraction of cells experiencing i or more pulses: this quantity increases with extent of damage, hence with increasing damage, the amplitude of the response in (2.29) increases uniformly.

These results are best illustrated concretely with the Lahav *et al.* (2004) data.

2.3.1. Illustrative examples from Lahav data. The distribution of ϕ_i from the data in Lahav *et al.* (2004) (with values for $i = 3, 4, 5$, extrapolated from the model; see appendix A) are shown below in table 3. The resulting equations for the ensemble p53 response $U_\gamma(t)$ (or $\bar{U}_\gamma(t)$ when normalized by N_T) for each indicated irradiation dose, γ are

$$\bar{U}_{0.3}(t) = 0.75\eta(t) + 0.15\eta(t-T) + 0.02\eta(t-2T), \tag{2.32}$$

$$\bar{U}_{2.5}(t) = 0.88\eta(t) + 0.35\eta(t-T) + 0.06\eta(t-2T) + 0.004\eta(t-3T), \tag{2.33}$$

$$\bar{U}_{10}(t) = 0.95\eta(t) + 0.6\eta(t-T) + 0.2\eta(t-2T) + 0.04\eta(t-3T) + 0.005\eta(t-4T), \tag{2.34}$$

and using the single response in figure 2 as the basis function $\eta(t)$ ($w = 400$ in this case), the resulting responses are shown in figures 4, 5 and 6 for a choice of $T = 500$. Note how, with increasing irradiation dose, the amplitude of each response increases (from 0.75 to 0.88 to 0.99 for the first pulse, and from 0.15 to 0.35 to 0.6 for the second, etc.), as does the number of apparent 'oscillations'. As noted earlier, because $T > w$, the pulse train appears as a sequence of distinct pulses, even as the amplitudes diminish monotonically.

3. THE COMPLETE DIGITAL CONTROL SYSTEM

The foregoing results and analyses and other evidence in the literature lead us to postulate the following hypothesis for the single cell DNA damage repair control mechanism shown in figure 1.

(i) Upon exposure to irradiation (dose γ Gy), the cell experiences DNA damage, the extent of which is quantified by y.

(ii) A sensor system (specifically, the protein complexes MRE11–RAD50–NBS1 (D'Armours & Jackson (2000)) and RAD9–RAD1–HUS1 (Zhou & Bartek 2004)) detects the presence or absence of DNA damage (not the extent), akin to converting y to a binary number $m = 0, 1$; this is communicated to the controller.[1]

(iii) If there is no damage, $m = 0$, and the controller does not respond ($c(t) = 0$); if there is damage, the controller issues an output $c(t)$ that is a pulse of unit height and duration δ min. (In control system engineering terms, this is equivalent to the switch indicated in figure 1 closing for the duration of δ mins and remaining open thereafter.) This

[1]Portions of this postulate have recently been confirmed experimentally in Abraham *et al.* (2005) and Lee & Paull (2005).

Table 3. Fraction of cells with $i=0, 1, 2, 3, 4, 5$ pulses as a function of irradiation dose, γ: values for 0, 1, 2 are obtained from Lahav *et al.* (2004); values for $i=3, 4, 5$ are extrapolated from the model (see appendix A).

ϕ_I	$\gamma=0.3$	$\gamma=2.5$	$\gamma=10.0$
ϕ_0	0.25	0.12	0.05
ϕ_1	0.60	0.53	0.35
ϕ_2	0.13	0.29	0.40
ϕ_3	0.02	0.056	0.16
ϕ_4	0.0	0.004	0.035
ϕ_5	0.0	0.0	0.005

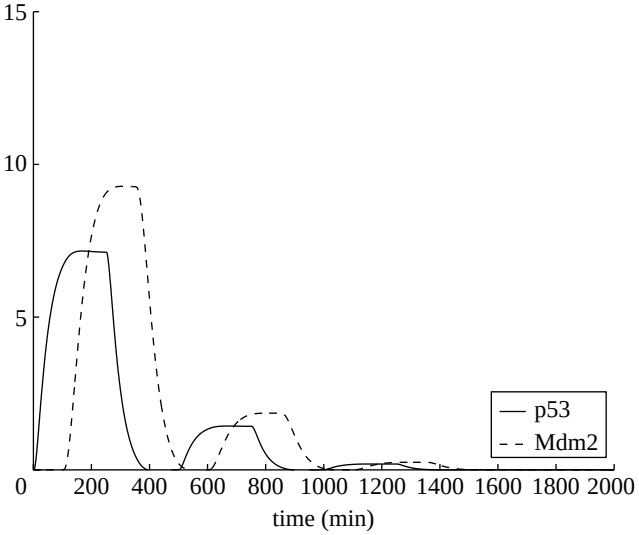

Figure 4. Ensemble response: $\gamma=0.3$; p53 (solid line) and Mdm2 (dashed line); as deviations from basal levels; in arbitrary units.

control signal is translated in the cell to the forcing function $f(t)$ shown in equation (2.10) as an input to the p53–Mdm2 effector system.

(iv) The net result of the controller action is manifested as $\eta(t)$, a single pulse of p53 such as shown in figure 2 (and also of its regulatory foil, Mdm2).

(v) An internal feedback mechanism monitors the p53 level: while higher than basal level, any further control action is suspended; only if the p53 level is low, *and* damage is still present ($m=1$), will the next pulse $c(t)$ be initiated. The period T (min) between the initiation of each pulse is thus determined by how long it takes for the amount of p53 expressed in the cell to return to basal level while damage persists.

(vi) This sequence is repeated until the damage is repaired. If damage persists after a (possibly physiologically predetermined?) number of pulses, the apoptotic program is initiated and the cell dies.

Note that for low-level damage, this mechanism will produce only a few identical pulses of p53; the number of pulses will grow with more extensive damage, precisely as observed in Lahav *et al.* (2004).

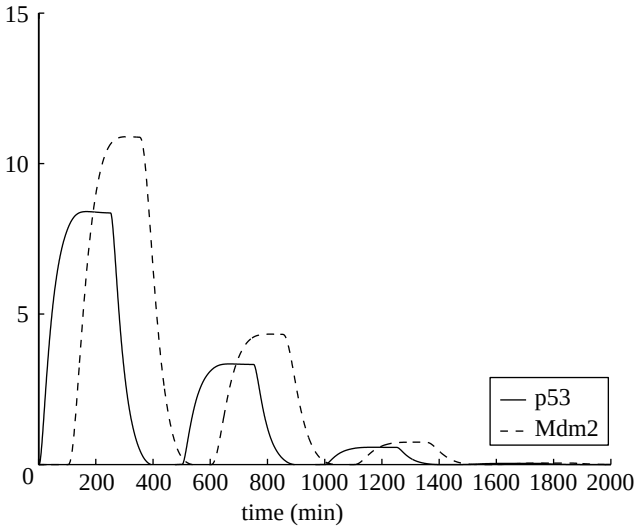

Figure 5. Ensemble response: $\gamma=2.5$; p53 (solid line) and Mdm2 (dashed line); as deviations from basal levels; in arbitrary units.

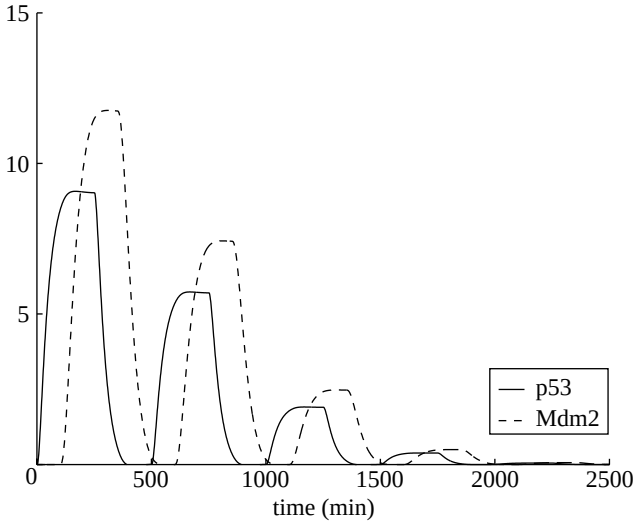

Figure 6. Ensemble response: $\gamma=10.0$; p53 (solid line) and Mdm2 (dashed line); as deviations from basal levels; in arbitrary units.

3.1. Stochastic ensemble simulation: process and control system

The entire preceding discussion and the generation of the ensemble responses shown thus far have been based on an ensemble of perfectly synchronized cells. All the p53 pulses are activated at precisely the same time, and last for precisely the same duration, the only variability considered being the Poisson-distributed extent of induced damage.

In reality, of course, there will be variability not only in the extent of damage experienced by each cell, but in the activation time of the p53 pulse and in its duration. To capture this behaviour requires a stochastic simulation, and one with the following characteristics was carried out.

(i) The population consists of 100 individual cells.

(ii) For each irradiation 'experiment' $\gamma=0.3$, 2.5 and 10, the distributions of y are taken as in appendix A as appropriate to each dose.

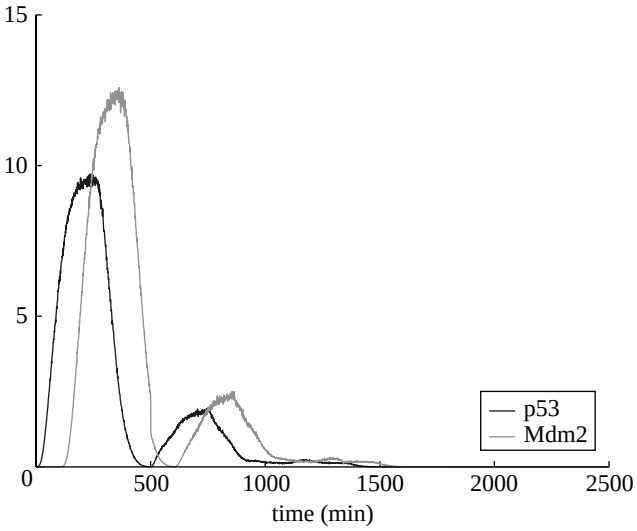

Figure 7. Stochastic simulation: population average p53 response to irradiation dose $\gamma = 0.3$ Gy.

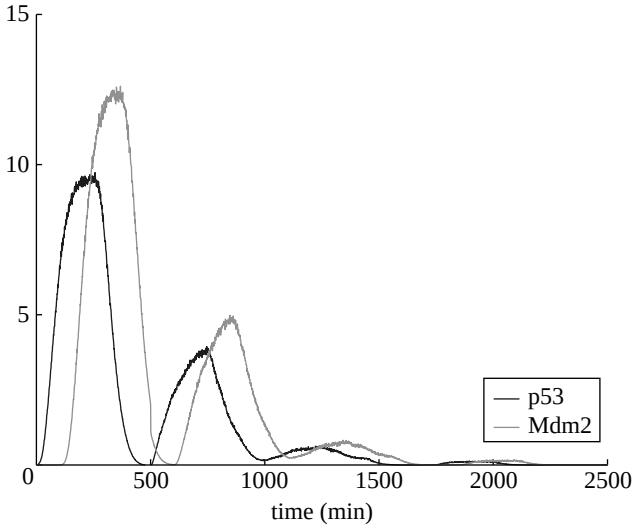

Figure 8. Stochastic simulation: population average p53 response to irradiation dose $\gamma = 2.5$ Gy.

(iii) The initiation of damage in each cell is taken to occur at time t_{0_i} obtained as

$$t_{0_i} = |\mathcal{N}(0, 50)|, \tag{3.1}$$

where $\mathcal{N}(0, 50)$ indicates a Gaussian random number with mean value 0, and standard deviation 50 (in min).

(iv) Each p53 pulse is taken to be as prescribed by the function $\eta(t)$ computed from the p53–Mdmd2 system model, but with a modest zero mean, standard deviation 0.1, Gaussian random noise component added.

(v) For each cell, the extent of damage variable y is binned as (0), (1,2), (3,4), (5,6), etc. Each pulse of p53 reduces the damage in 'bins' so that, for example, if y started out as 5 or 6, at the end of the first period after the application of the first pulse, y is reduced to 3 or 4, and after the next round to 1 or 2—thus requiring 3 pulses to eliminate the damage completely.

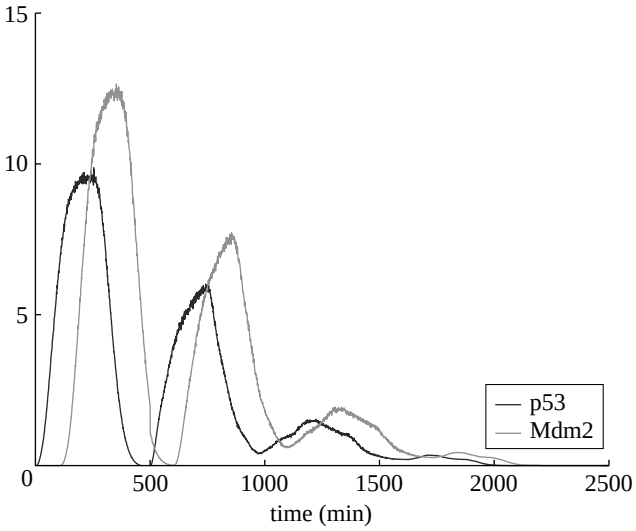

Figure 9. Stochastic simulation: population average p53 response to irradiation dose $\gamma = 10.0$ Gy.

(vi) In accordance with the control scheme shown above, no new pulse is initiated until the level of p53 is back to a tolerance level of 0.2 (two standard deviations) *and* there is still residual damage. This introduces further variability in the periodicity of each pulse.

The resulting simulation of the population average p53 responses are shown in figures 7–9; they are realistic and reminiscent of the experimental responses described in Bar-Or *et al.* (2000) and Lahav *et al.* (2004, see fig. 3*g*).

4. DISCUSSION

By employing dynamic modelling synergistically with engineering systems analysis and statistics, we have been able to resolve the multi-scale dilemma in which single cell non-oscillatory pulses in p53 appear as damped oscillations in the ensemble. We showed that explicit individual component models, a dynamic transfer function model-guided choice of model parameters, and explicitly accounting for cell-to-cell variability in the ensemble were all crucial to the final results.

Specifically, we found that a relatively simple two-state dynamic model (with four kinetic parameters and a time delay) was sufficient to capture the essential dynamic behaviour of the p53–Mdm2 effector system; a Poisson model for a yet unknown 'extent-of-damage' variable, y, whose mean value (the Poisson variable, λ) varied with irradiation dose via a power law relationship (two parameters), was consistent with experimental data (Lahav *et al.* 2004) in describing the DNA damage response process. We have also proposed a plausible digital control mechanism, a stochastic simulation of which gave realistic results.

Several issues remain unresolved, however. First, it will be important to investigate the physiological meaning of the variable y. This variable is the unobserved stimulus for the control action observed in the cell as pulses of p53 in response to DNA damage. If, as we have suggested here, y is a *primary* DNA

damage 'response' variable indicating the extent of damage (as opposed to the level of p53 which is a secondary response), what does its quantization as an integer variable mean; what is the biological implication of its Poisson-distributed characteristics; and why is it related to the number of p53 pulses in the specific look-up-table-like fashion as shown in equation (2.25)? Why is it that a single pulse of p53 apparently reduced the variable y in blocks of two units? Is y related to the number of DNA double-stranded breaks? Or does the fact that y is an integer suggest that it is more likely to be the output of a sensor (such as the MRE11–RAD50–NBS1 protein complex)?

It will also be interesting to analyse the postulated control system, investigate the implications of the structure for robustness, and stability, and obtain some insight into the design principles it represents. For example, a control system in which the primary sensor is only required to indicate the presence or absence of a particular signal ($m=0, 1$) will be more robust than one in which the sensor is required to provide a precise measure of the signal in question.

It is also possible that a better understanding of this control system will enable the elucidation of the cell-fate decision mechanism by which the choice is made between repair, cell-cycle arrest, or apoptosis. For example, is there an additional component that monitors the total number of p53 pulses? What determines the threshold value above which the cell is irreversibly committed to the apoptotic program? This will require complementary experimental and computational programs for developing of a fuller picture of the control system that incorporates additional components involved in the cell-fate decision process.

5. METHODS

5.1. Computational platforms

All dynamic simulations were carried out using the MATLAB/SIMULINK computational platform as described in the main body of the manuscript. The statistical analyses were all carried out using the software package MINITAB.

5.2. Data source

The illustrative data set (in table 1) was obtained from Lahav *et al.* (2004); figure 3*a*.

I thank Dan Zak, Haiping Hao, Jan Hoek and Boris Kholodenko for useful discussions on the subject, and Claudio Gelmi for assistance in generating the stochastic simulation plots. I am also grateful to the anonymous reviewers for their helpful comments. A significant portion of the work was completed while I was at Cambridge University's Isaac Newton Institute for Mathematical Sciences, for the workshop on *Statistical Mechanics of Cellular Systems* during the months of January and February of 2004; their hospitality and financial support is gratefully acknowledged. Additional funding was provided by the Delaware Biotechnology Institute.

APPENDIX A

Tables of probability model predictions and data.

Table A1. Model prediction and data—distribution of p53 pulses: $\gamma=0.0$.

y damage	z p53 pulses	$f_y(y\|\lambda=0.05)$	$\hat{f}_z(z)$ predicted	$f_z(z)$ observed
0	0	0.951	0.951	0.95
1	1	0.048	0.049	0.05
2	1	0.001	0.049	0.05
3	2	0.000	0.001	0.00
4	2	0.000	0.001	0.00
5	3	0.000	0.000	0.00
6	3	0.000	0.000	0.00
7	4	0.000	0.000	0.00
8	4	0.000	0.000	0.00
9	5	0.000	0.000	0.00
10	5	0.000	0.000	0.00

Table A2. Model prediction and data—distribution of p53 pulses: $\gamma=0.3$.

y damage	z p53 pulses	$f_y(y\|\lambda=1.30)$	$\hat{f}_z(z)$ predicted	$f_z(z)$ observed
0	0	0.273	0.273	0.25
1	1	0.354	0.584	0.60
2	1	0.230	0.584	0.60
3	2	0.100	0.132	0.15
4	2	0.032	0.132	0.15
5	3	0.008	0.010	0.00
6	3	0.002	0.010	0.00
7	4	0.000	0.000	0.00
8	4	0.000	0.000	0.00
9	5	0.000	0.000	0.00
10	5	0.000	0.000	0.00

Table A3. Model prediction and data—distribution of p53 pulses: $\gamma=2.5$.

y damage	z p53 pulses	$f_y(y\|\lambda=2.1)$	$\hat{f}_z(z)$ predicted	$f_z(z)$ observed
0	0	0.122	0.122	0.12
1	1	0.257	0.527	0.53
2	1	0.270	0.527	0.53
3	2	0.189	0.290	0.35[a]
4	2	0.099	0.290	0.35[a]
5	3	0.042	0.056	0.00
6	3	0.015	0.056	0.00
7	4	0.004	0.005	0.00
8	4	0.001	0.005	0.00
9	5	0.000	0.000	0.00
10	5	0.000	0.000	0.00

[a] Lahav data is for $z\geq 2$ whereas the model prediction provides individual probabilities for all higher number of pulses. The experimental value 0.35 is to be compared with $0.29+0.056+0.005=0.351$.

Table A4. Model prediction and data—distribution of p53 pulses: $\gamma = 10.0$.

Y damage	z p53 pulses	$f_y(y\|\lambda=3.1)$	$\hat{f}_z(z)$ predicted	$f_z(z)$ observed
0	0	0.045	0.045	0.05
1	1	0.140	0.356	0.35
2	1	0.216	0.356	0.35
3	2	0.224	0.396	0.60[a]
4	2	0.173	0.396	0.60[a]
5	3	0.107	0.163	0.00
6	3	0.056	0.163	0.00
7	4	0.025	0.035	0.00
8	4	0.010	0.035	0.00
9	5	0.003	0.004	0.00
10	5	0.001	0.004	0.00

[a] Lahav data is for $z \geq 2$ whereas the model prediction provides individual probabilities for all higher number of pulses; the experimental value 0.60 is to be compared with $0.396 + 0.163 + 0.035 + 0.004 = 0.598$.

REFERENCES

Abraham, R. T. & Tibbetts, R. S. 2005 Guiding ATM to broken DNA. *Science* **808**, 510–511. (doi:10.1126/science.1112069)

Bar-Or, R. L., Maya, R., Segel, L. A., Alon, U., Levine, A. J. & Oren, M. 2000 Generation of oscillations by the p53–Mdm2 feedback loop: a theoretical and experimental study. *Proc. Natl Acad. Sci.* **97**(21), 11 250–11 255. (doi:10.1073/pnas.210171597)

D'Armours, D. & Jackson, S. P. 2000 The MRE11 complex: at the crossroads of DNA repair and checkpoint signalling. *Nat. Rev. Mol. Cell Biol.* **3**, 317–327. (doi:10.1038/nrm805)

Draper, N. R. & Smith, H. 1981 *Applied regression analysis*. New York: Wiley.

Lahav, G., Rosenfeld, N., Sigal, A., Geva-Zatorsky, N., Levine, A. J., Elowitz, M. B. & Alon, U. 2004 Dynamics of the p53–Mdm2 feedback loop in individual cells. *Nat. Genet.* **36**, 147–150. (doi:10.1038/ng1293)

Lakin, N. D. & Jackson, S. P. 1999 Regulation of p53 in response to DNA damage. *Oncogene* **18**, 7644–7655. (doi:10.1038/sj.onc.1203015)

Lee, J.-H. & Paull, T. T. 2005 ATM Activation by DNA double strand breaks through the Mre11–Rad50–Nbs1 complex. *Science* **808**, 551–553. (doi:10.1126/science.1108297)

Michael, D. & Oren, M. 2003 The p53–Mdm2 module and the ubiquitin system. *Semin. Cancer Biol.* **13**, 49–58. (doi:10.1016/S1044-579X(02)00099-8)

Mihalas, G. I., Simon, Z., Belea, G. & Popa, E. 2000 Possible oscillatory behavior in p53–Mdm2 interaction computer simulation. *J. Biol. Syst.* **8**, 21–29.

Momand, J., Wu, H. H. & Dasgupta, G. 2000 Mdm2-master regulator of the p53 tumor suppressor protein. *Gene* **242**, 15–29. (doi:10.1016/S0378-1119(99)00487-4)

Monk, N. A. M. 2003 Oscillatory expression of Hes1, p53 and NF(B driven by transcriptional time delays. *Curr. Biol.* **13**, 1409–1413. (doi:10.1016/S0960-9822(03)00494-9)

Ogunnaike, B. A. & Ray, W. H. 1994 *Process dynamics modeling and control*. New York: Oxford University Press.

Piette, J., Neel, H. & Marechal, V. 1997 Mdm2: keeping p53 under control. *Oncogene* **15**, 1001–1010. (doi:10.1038/sj.onc.1201432)

Tiana, G., Jensen, M. H. & Sneppen, K. 2002 Time delay as a key to apoptosis induction in the p53 network. *Eur. Phys. J. B* **29**, 135–140. (doi:10.1140/epjb/e2002-00271-1)

Vogelstein, B., Lane, D. & Levine, A. J. 2000 Surfing the p53 network. *Nature* **408**, 307–310. (doi:10.1038/35042675)

Zhou, B. B. S. & Bartek, J. 2004 Targeting the checkpoint kinases: chemosensitization versus chemoprotection. *Nat. Rev. Cancer* **4**, 1–10. (doi:10.1038/nrc1296)