

## Modelling the SOS Response by Semi-Stochastic Simulation \*

NI Ming(倪鸣)<sup>†</sup>, WANG Si-Yuan(王思远)<sup>†\*\*</sup>, OUYANG Qi(欧阳颀)<sup>\*\*\*</sup>

School of Physics and Center for Theoretical Biology, Peking University, Beijing 100871

(Received 31 March 2008)

The SOS (save our soul) response induced by DNA damage in bacteria *E coli* has raised a great interests in biophysics and has been extensively studied. Previously we have developed a stochastic simulation model to explain the oscillatory-like modulation of SOS gene expression observed in experiment. Here we present an improved semi-stochastic model which has higher simulation efficiency, taking into account the updated knowledge about SOS response. The improved model suggests that frequency of the modulation is controlled by the negative feedback in the system. DNA polymerase V, the key enzyme for error-prone translesion synthesis during SOS response, plays a major role in closing up the negative feedback. It is also indicated that the correlation between the modulation and cellular growth observed in experiment is due to DNA damage induced slowing down of transcription and translation.

PACS: 87.18.-h, 87.18.Tt, 87.18.Vf

The SOS response in bacteria *E coli* has been extensively studied as a typical DNA damage response model system.<sup>[1]</sup> However, a quantitative understanding of this control mechanism in molecular level is unclear at present. In 2005, Friedman *et al.*<sup>[2]</sup> investigated the promoter activity of SOS genes after ultraviolet (UV) radiation (a widely used way to produce DNA damage) at single cell level. It was shown that the promoter activity has oscillatory-like modulation profile. In order to explain this oscillatory-like behaviour, previously we developed a stochastic reaction model<sup>[3]</sup> based on Gillespie algorithm.<sup>[4]</sup> We suggested that the modulation is basically due to discontinuous collision between the replication fork and DNA lesion. In this Letter, we modify the model to take newly discovered features into account. The modified model improves simulation efficiency and fit the experimental facts better. The results also give us new insights to understanding of dynamics of SOS response.

Table 1. SOS genes involved in the model and their functions.

Name	Function
LexA	Transcription repressor of SOS genes
RecA	Catalysing LexA and UmuD auto-cleavage recombinational repair (RR)
UvrB	Nucleotide excision repair (NER)
UmuD	Turning into UmuD', which is subunit of PolV
UmuC	Another subunit of PolV, Translesion DNA synthesis (TLS)

As illustrated in Fig. 1(a), the SOS response is triggered when DNA replication is blocked at the DNA damage point.<sup>[1]</sup> RecA protein is then activated by binding to the single strand DNA (ssDNA) exposed

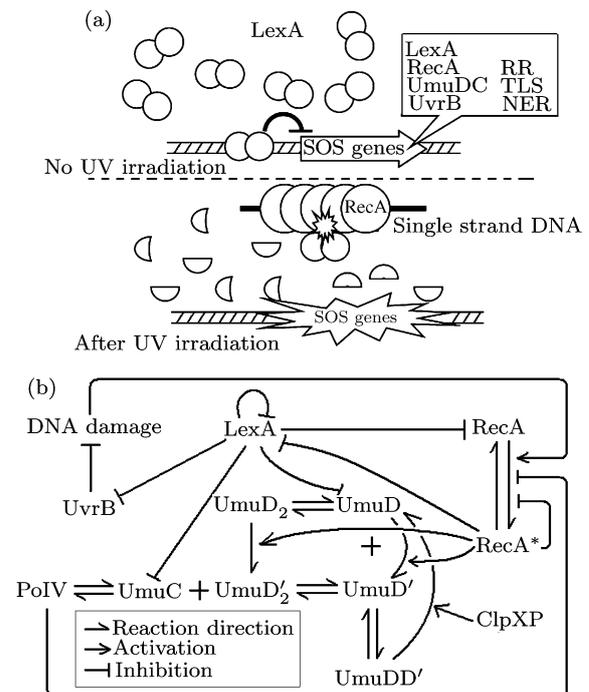


Fig. 1. (a) A sketch for SOS response induced by UV irradiation; (b) the control circuit of SOS pathway.

at the stalled replication fork. The activated RecA, named RecA\*, catalyses LexA cleavage. The transcription repressor LexA controls a series of SOS genes expression. These SOS genes have various functions as listed in Table 1. The stalled replication fork can restart via NER, TLS or RR pathway.<sup>[5]</sup> The essential control circuit is illustrated as in Fig. 1(b). Once replication resumes, RecA\* turns back into inactive form, LexA re-accumulates, and SOS response ends. In

\*Supported by the National Natural Science Foundation of China under Grant No 10574004, and the National Key Basic Research Programme of China under Grant No 2003CB715900.

\*\* Email: siyuanw@princeton.edu

\*\*\* Email: qi@pku.edu.cn

<sup>†</sup> The authors contributed equally to this work.

© 2008 Chinese Physical Society and IOP Publishing Ltd

the previous model, the Gillespie algorithm, a widely used computational algorithm for simulating stochastic processes of chemical reactions, is used because the limited number of replication forks in one cell could not be treated as continuous variable.<sup>[4]</sup> For the same reason, in the new model, the replication fork is treated as discrete variable governed by Markov process. However, for other elements in the model, the molecule numbers per cell are large enough to be considered as continuous variables, so they are described by ordinary differential equations (ODE). Further, we take into account the fact that UV irradiation slows down the cellular growth. In the ODEs,  $S$  is a coefficient reflecting the slowing down of transcription and translation coupled with the cell size doubling time  $T_D$ .  $S = 40 \text{ min}/T_D$ .  $D$  is the dilution rate. Therefore, we have  $D = T_D^{-1} \ln 2$ . According to Fig. 1(b), the dynamics of the SOS control system can be written as follows:

$$\frac{d[mLexA]}{dt} = \frac{v_{LexA} \times S}{1 + [LexA]/KM_{RecA}} - (kd_{mLexA} + D)[mLexA], \quad (1)$$

$$\frac{d[mRecA]}{dt} = \frac{v_{RecA} \times S}{1 + [LexA]/KM_{RecA}} - (kd_{mRecA} + D)[mRecA], \quad (2)$$

$$\frac{d[mUvrB]}{dt} = \frac{v_{UvrB} \times S}{1 + [LexA]/KM_{UvrB}} - (kd_{mUvrB} + D)[mUvrB], \quad (3)$$

$$\frac{d[mUmuD]}{dt} = \frac{v_{UmuD} \times S}{1 + [LexA]/KM_{UmuD}} - (kd_{mUmuD} + D)[mUmuD], \quad (4)$$

$$\frac{d[mUmuC]}{dt} = \frac{v_{UmuC} \times S}{1 + [LexA]/KM_{UmuC}} - (kd_{mUmuC} + D)[mUmuC], \quad (5)$$

$$\begin{aligned} \frac{d[LexA]}{dt} &= kp_{LexA} \times S[mLexA] \\ &\quad - kclv_{LexA}[LexA][RecA^*] \\ &\quad - (kd_{LexA} + D)[LexA], \end{aligned} \quad (6)$$

$$\begin{aligned} \frac{d[RecA]}{dt} &= kp_{RecA} \times S[mRecA] \\ &\quad - (kd_{RecA} + D)[RecA], \end{aligned} \quad (7)$$

$$\begin{aligned} \frac{d[UmuD]}{dt} &= -2 \times (k_{DD1}[UmuD]^2 \\ &\quad - k_{DD2}[UmuD_2]) \\ &\quad - (k_{DD'1}[UmuD][UmuD'] \\ &\quad - k_{DD'2}[UmuDD']) \\ &\quad - kclv_{UmuD}[UmuD][RecA^*] \\ &\quad + kd_{clp}[UmuDD'] \\ &\quad + kp_{UmuD} \times S[mUmuD] \end{aligned}$$

$$- (kd_{UmuD} + D)[UmuD], \quad (8)$$

$$\begin{aligned} \frac{d[UmuD_2]}{dt} &= k_{DD1}[UmuD]^2 \\ &\quad - k_{DD2}[UmuD_2] \\ &\quad - kclv_{UmuD}[UmuD_2][RecA^*] \\ &\quad - (kd_{UmuD_2} + D)[UmuD], \end{aligned} \quad (9)$$

$$\begin{aligned} \frac{d[UmuD']}{dt} &= -2 \times (k_{D'D'1}[UmuD']^2 \\ &\quad - k_{D'D'2}[UmuD'_2]) \\ &\quad - (k_{DD'1}[UmuD][UmuD'] \\ &\quad - k_{DD'2}[UmuDD']) + kclv_{UmuD} \\ &\quad \cdot [UmuD][RecA^*] - D[UmuD'], \end{aligned} \quad (10)$$

$$\begin{aligned} \frac{d[UmuD'_2]}{dt} &= k_{D'D'1}[UmuD']^2 \\ &\quad - k_{D'D'2}[UmuD'_2] \\ &\quad + kclv_{UmuD}[UmuD_2][RecA^*] \\ &\quad - (k_{DC1}[UmuD'2][UmuC] \\ &\quad - k_{DC2}[PolV]) - D[UmuD'_2], \end{aligned} \quad (11)$$

$$\begin{aligned} \frac{d[UmuDD']}{dt} &= k_{DD'1}[UmuD][UmuD'] \\ &\quad - k_{DD'2}[UmuDD'] \\ &\quad - (kd_{UmuDD'} + D)[UmuDD'], \end{aligned} \quad (12)$$

$$\begin{aligned} \frac{d[UmuC]}{dt} &= kp_{UmuC} \times S[mUmuC] \\ &\quad - (k_{DC1}[UmuD'2][UmuC] \\ &\quad - k_{DC2}[PolV]) \\ &\quad - (kd_{UmuC} + D)[UmuC], \end{aligned} \quad (13)$$

$$\begin{aligned} \frac{d[PolV]}{dt} &= k_{DC1}[UmuD'2][UmuC] \\ &\quad - k_{DC2}[PolV] - D[PolV], \end{aligned} \quad (14)$$

$$\begin{aligned} \frac{d[UvrB]}{dt} &= kp_{UvrB} \times S[mUvrB] \\ &\quad - (kd_{UvrB} + D)[UvrB], \end{aligned} \quad (15)$$

$$\frac{d[Lesion]}{dt} = -k_{NER}[Lesion][UvrB], \quad (16)$$

$$\begin{aligned} \frac{d[RecA^*]}{dt} &= k_{on} \times ssDNA[RecA] \\ &\quad - k_{off}[RecA^*]. \end{aligned} \quad (17)$$

The parameters used here accord to the previous model<sup>[3]</sup> with slight modification. The stochastic part contains the generation and elimination of the inducing signal, RecA\*. We simulate two replication forks travelling along the chromosome on which DNA lesions are uniformly distributed. When a replication fork hits a lesion, it stops and the parameter ssDNA in Eq. (17) increases by 1. Then, according to Eq. (17), RecA\* starts to accumulate. The probability for the

replication fork to restart between  $[t, t + dt]$  is written as  $(k_{NER}[UvrB] + k_{TLS}[PolV] + k_{RR})dt$ , where  $k_{NER}$ ,  $k_{TLS}$ ,  $k_{RR}$  represent the efficiencies of NER, TLS and RR pathway respectively. When replication restarts, ssDNA decreases by 1.

A new assumption is added to the new model. According to the latest experimental results,<sup>[6]</sup> we assume that once TLS happens, PolV hangs on the repli-

cation fork for 30 min so that in the next 30 min, if the replication fork is blocked by lesion again, TLS takes place immediately without RecA\* accumulation.

To take into account the environmental fluctuations involving temperature, nutrition, irradiation etc. that may majorly affect cellular growth rate,  $T_D$  is set as a uniform random variable between  $\bar{T}_D - 10$  min and  $\bar{T}_D + 10$  min.  $\bar{T}_D = 40 \text{ min} + [\text{UV}]^{[2]}$ , where  $[\text{UV}]$

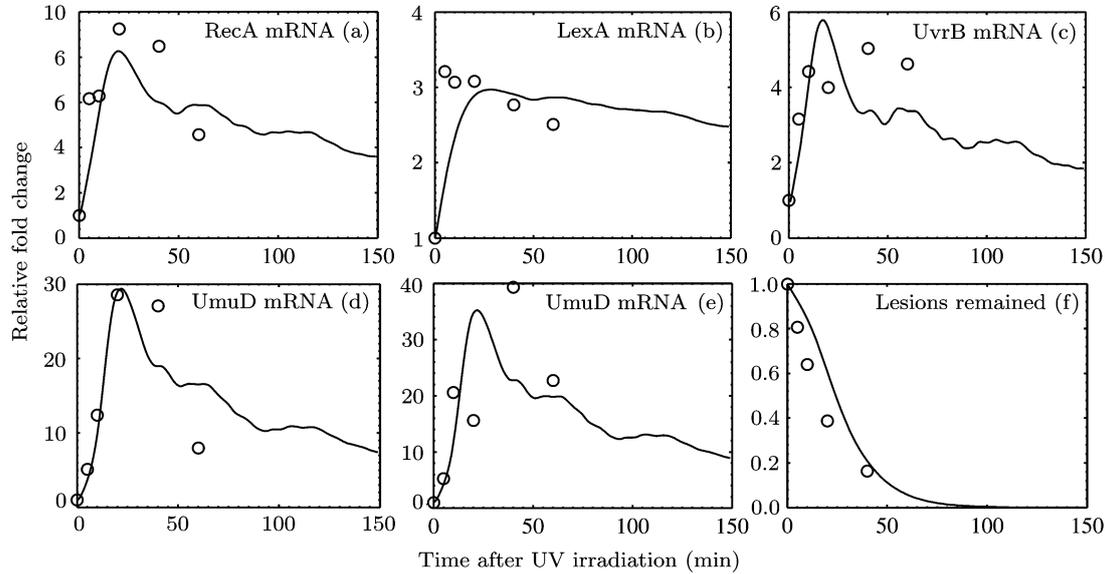


Fig. 2. Averaging simulation results for 50 runs. UV dose =  $40 \text{ J/m}^2$ . Lines denote simulation results. Circles denote experimental data: (a)–(e) mRNA level,<sup>[7]</sup> (f) lesions level.<sup>[8]</sup>

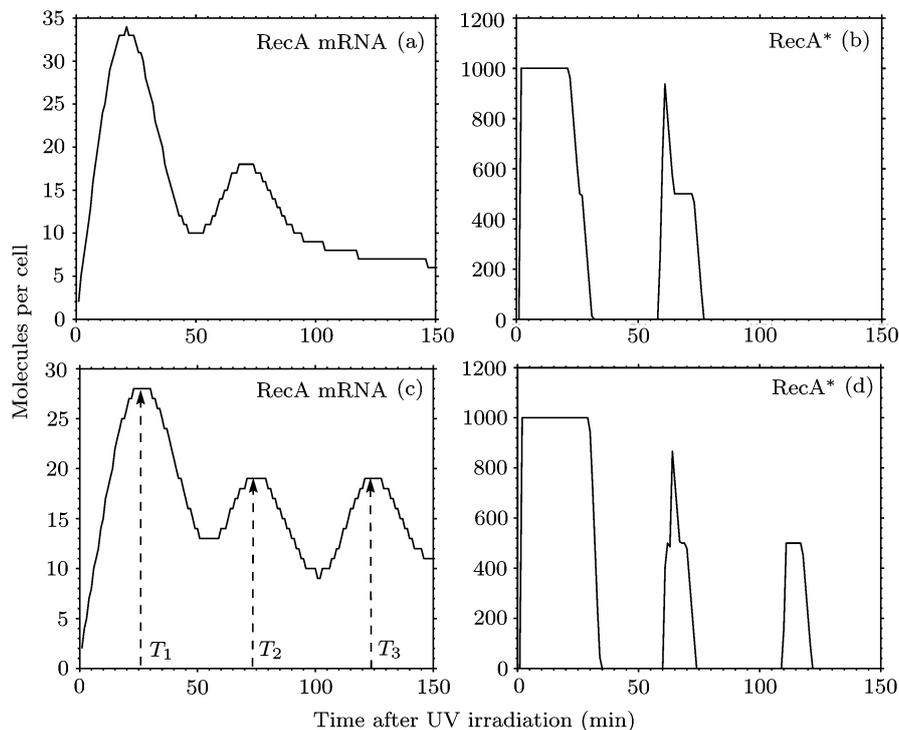
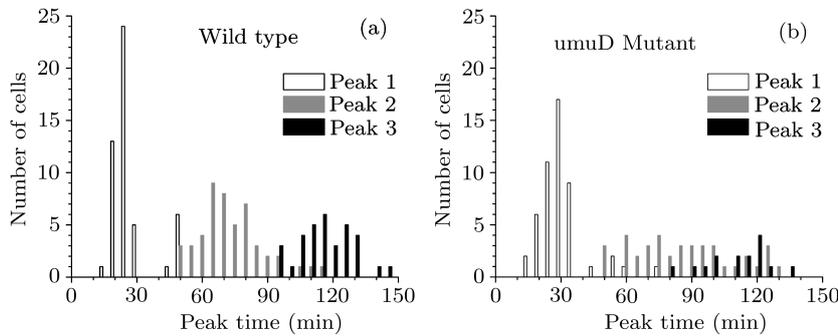


Fig. 3. Illustration of the RecA mRNA modulation, a case study. (a) RecA mRNA level. (b) RecA\* level. UV dose =  $20 \text{ J/m}^2$ . (c) RecA mRNA level. (d) RecA\* level. UV dose =  $40 \text{ J/m}^2$ .  $T_1$ ,  $T_2$  and  $T_3$  denote the peak time for the first, second and third peaks.

is the UV irradiation dose in units of  $\text{J}/\text{m}^2$ . Note that the simulation results are not sensitive to the 10 min-deviation assumption. Figure 2 shows the time course of the average results for 50 runs. The results capture the significant profile of the expression dynamics of SOS response.

In Fig. 3, two typical runs for  $20 \text{ J}/\text{m}^2$  and  $40 \text{ J}/\text{m}^2$  irradiations respectively are picked up as examples. In lower dose of UV irradiation ( $20 \text{ J}/\text{m}^2$ ) RecA\* shows two separate RecA\* peaks as a function of time (Fig. 3(b)). In higher dose of UV irradiation ( $40 \text{ J}/\text{m}^2$ ), it takes more time to remove the larger amount of lesions so that the SOS response persists longer. In this case, RecA\* shows the additional third peak (Fig. 3(d)). Since RecA\* controls the level of LexA, and LexA in turn controls SOS gene expression, the SOS gene mRNA level is controlled by RecA\*. As shown in Figs. 3(a) and 3(c), RecA mRNA level correlates with RecA\*. The results shown in Fig. 3 is consistent with the experimental observations.<sup>[2]</sup> As we sum up the peak position for 50 runs of  $40 \text{ J}/\text{m}^2$  UV

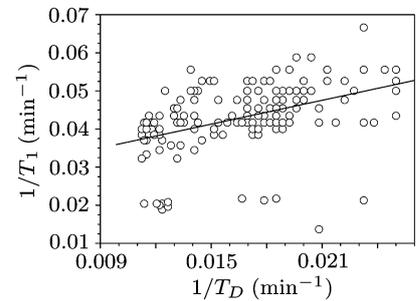
irradiation, the distribution of the first, second and third peak lies separately with the internal around 30 minutes (Fig. 4(a)). This result is also agree with the experimental result in Ref. [2] and fits better than the previous model.<sup>[3]</sup> If UmuD is deleted in the model, we obtain a different distribution in which the second and third peak positions vary so much that they actually overlap (Fig. 4(b)). This is again in agreement with the measurement for the UmuD mutant strain<sup>[2]</sup>. It indicates that PolV plays a significant role in modulating the SOS gene expression profile. The induction of RecA\* leads to SOS genes expression, including PolV. When PolV level is raised and TLS takes place, it will help the replication fork to bypass the DNA lesions. Therefore, RecA\* level drops down and so do PolV. In the absence of TLS function, the replication fork would be stopped by the lesions and induce RecA\* again. The correlation between RecA\* and PolV constructs a negative feedback and leads to oscillatory-like modulation.



**Fig. 4.** RecA mRNA dynamics statistic in 50 runs, UV dose =  $40 \text{ J}/\text{m}^2$ . (a) Wild type strain. (b) UmuD mutant strain.

The correlation between growth rate ( $1/T_D$ ) and the reciprocal of the time of the first peak ( $1/T_1$ ) is shown in Fig. 5. From this simulation result, we infer that the growth rate positively correlates with  $1/T_1$ <sup>[2]</sup> because the DNA damage simultaneously affects the time scale of both cellular growth and SOS response. There are various kinds of biochemical reactions involved in SOS response. According to the kinetic parameters,<sup>[3]</sup> the rate-limit steps lie in protein production, i.e. transcription and translation, and protein degradation. As shown in the ODEs,  $S$  and  $D$  control the protein production and degradation rate respectively and both of them are functions of  $T_D$ . Therefore, the first peak time  $T_1$  is determined by  $S$  and  $D$  and they are in turn determined by  $T_D$ .

In conclusion, our new model suggests that first, the frequency of the modulation is controlled by the negative feedback in the system and the interaction between DNA damage and PolV is responsible for



**Fig. 5.** Correlation between  $1/T_D$  and  $1/T_1$ . Correlation coefficient = 0.42 ( $p < 10^{-4}$ ). Slope of the regression line =  $1.0 \pm 0.2$ . UV dose = 40, 20,  $10 \text{ J}/\text{m}^2$ . Here 150 runs in total are included.

the feedback. Second, the time scale of the SOS response dynamic is govern by the protein production and degradation rate, which are affected by the extent of DNA damage.

## References

- [1] Errol C, Friedberg G C W, Siede W, Wood R D, Schultz R A and Ellenberger T 2005 *DNA Repair and Mutagenesis* 2nd edn (ASM Press)
- [2] Friedman N, Vardi S, Ronen M, Alon U and Stavans J 2005 *PLoS Biol.* **3** e238
- [3] Ni M, Wang S Y, Li J K and Ouyang Q 2007 *Biophys. J.* **93** 62
- [4] Gillespie D T 1977 *J. Phys. Chem.* **81** 2340
- [5] Kreuzer K N 2005 *Annu. Rev. Microbiol.* **59** 43
- [6] Schlacher K, Pham P, Cox M M and Goodman M F 2006 *Chem. Rev.* **106** 406
- [7] Courcelle J, Khodursky A, Peter B, Brown P O and Hanawalt P C 2001 *Genetics* **158** 41
- [8] Crowley D J and Hanawalt P C 1998 *J. Bacteriol.* **180** 3345