The role of the bacterial mismatch repair system in SOS-induced mutagenesis: A theoretical background

Oleg V. Belov, Ochbadrakh Chuluunbaatar, Mikhail I. Kapralov, Nasser H. Sweilam

Laboratory of Radiation Biology, Joint Institute for Nuclear Research, 6 Joliot–Curie Street, 141980 Dubna, Moscow Region, Russia
Laboratory of Information Technologies, Joint Institute for Nuclear Research, 6 Joliot–Curie Street, 141980 Dubna, Moscow Region, Russia
School of Mathematics and Computer Science, National University of Mongolia, 1 Central Ulaanbaatar, PO Box 464/523, 210646, Mongolia
Mathematics Department, Faculty of Science, Cairo University, Giza 12613, Egypt

HIGHLIGHTS

• A mathematical model of the bacterial DNA mismatch repair system is developed.
• Five key pathways of the Escherichia coli mismatch repair are simulated adequately.
• The relationships between SOS and MMR systems are described quantitatively.
• A possible mechanistic explanation of MMR role in UV-mutagenesis is shown.

ARTICLE INFO

Article history:
Received 29 August 2012
Received in revised form
20 April 2013
Accepted 22 April 2013
Available online 30 April 2013

Keywords:
SOS response
Mismatch repair
UV mutagenesis
Mathematical modeling

ABSTRACT

A theoretical study is performed of the possible role of the methyl-directed mismatch repair system in the ultraviolet-induced mutagenesis of Escherichia coli bacterial cells. For this purpose, mathematical models of the SOS network, translesion synthesis and mismatch repair are developed. Within the proposed models, the key pathways of these repair systems were simulated on the basis of modern experimental data related to their mechanisms. Our model approach shows a possible mechanistic explanation of the hypothesis that the bacterial mismatch repair system is responsible for attenuation of mutation frequency during ultraviolet-induced SOS response via removal of the nucleotides misincorporated by DNA polymerase V (the UmuD'2C complex).

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

One of the biological systems capable of correcting the non-complementary nucleotide pairs that appear as a consequence of certain factors is the methyl-directed mismatch repair system (MMR) (Lahue et al., 1989; Modrich and Lahue, 1996). The evidences of the functioning of this system were found in many organisms including bacteria, yeasts, and mammals. Despite high MMR conservability and the similarity of the repair mechanisms between bacteria and mammals, the interrelations of its pathways and other repair systems are well understood only for relatively simple biological objects like prokaryotic cells.

The factors which can start the MMR system may include the errors that occur during normal DNA replication and cell metabolism as well as a spectrum of DNA lesions induced by exposure to different agents of physical and chemical nature and the following DNA repair processes (Li, 2008). Among the physical factors capable of inducing this system, the action of radiations of different types is very interesting in terms of its use as an instrument for studying the MMR connections with other repair systems responsible for the mutagenic effects in the living organisms. A number of experimental facts indicate that MMR plays a certain role in the mutagenic effects of ionizing and ultraviolet (UV) radiation (Hongbo et al., 2000; Martin et al., 2010). Some of these facts suggest the involvement of MMR in mutagenic pathways of other repair systems.

Among the pathways leading to an increase in the mutation frequency and other negative effects under the influence of physical and chemical factors, an important role belongs to the SOS repair system (Radman, 1974; Witkin, 1976; Krasavin and Kozubek, 1991). Intense studies of the SOS response of prokaryotic cells have identified the key role of the specific PolV Mut complex comprising DNA polymerase V (or UmuD'2C) in the process of DNA
synthesis through the lesion which was called translation synthesis (TLS) (Wang, 2001). This mechanism is also realized not only in prokaryotic cells but in mammalian and human cells, too (Yang et al., 2003; Chiapperino et al., 2005).

Experimental studies have shown that PolV Mut demonstrates a relatively high error frequency during the incorporation of bases in nascent strands opposite the lesions which were not removed during the earlier stages of repair (Tang et al., 2000). However, the finally measured mutation frequency in individual genes is not so high as it might have been if all mismatches produced by the PolV Mut complex had been fixed as mutations. Our previous research related to the mathematical modeling of the mechanism of SOS-induced mutagenesis under 254 nm ultraviolet (UV) radiation demonstrated this fact by an interval of the free parameter value responsible for fixing the PolV-induced mismatches as mutations (Belov et al., 2009). These conclusions made us introduce in our model additional repair mechanisms at the final stages of SOS response. Taking into account the specific character of DNA synthesis by the PolV Mut complex and relying on the corresponding experimental facts, we have chosen the MMR system of *Escherichia coli* bacterial cells for the theoretical analysis of its influence on the UV-induced mutagenic effect. So, the main goal of this study is to identify the role of MMR in SOS-induced mutagenesis on the basis of the precise modeling of the enzymatic mechanisms of these two repair systems under exposure to radiation.

2. Mathematical model

2.1. A quantitative model of the SOS network

UV radiation induces two major types of photoadducts in DNA, namely, thymine–thymine cyc–syn cyclobutane photodimers and thymine–thymine pyrimidine (6–4) pyrimidone photoproducts. A significant portion of the primary induced thymine dimers are efficiently removed by pre-replication mechanisms such as photoreactivation (Rupert, 1975) and nucleotide excision repair (NER) (Sancar and Sancar, 1988). The lesions which were not removed by these repair mechanisms lead to the production of single-stranded DNA gaps (ssDNA) in bacterial chromosomes that prevent successfull process the remaining single-stranded gaps, cells actuate a number of specific repair mechanisms called SOS response (Radman, 1974). Therefore, ssDNA is regarded as an inducing signal for a cell to launch the SOS system.

To describe the dynamics of SOS signal induction, we used the mathematical approach developed earlier (Belov et al., 2009; Aksenov, 1999). The final dimensionless equations of this model are as follows:

\[
\begin{align*}
\text{for } r < r_2 & \quad x_0(r, \Psi) = \Psi \exp(-q_1 r) \int_0^q \exp(q_4 z) \frac{\exp(q_3 z)}{q_3 \exp(q_3 z)} dz, \\
\text{for } r \geq r_2 & \quad x_0(r, \Psi) = \Psi \exp(-q_1 r) \int_0^r \exp(q_4 z) \frac{\exp(q_3 z)}{q_3 \exp(q_3 z)} dz,
\end{align*}
\]

where

\[r_2 = \frac{1}{q_3} \ln(\exp(q_4)(1 + q_2 \Psi) - q_1 \Psi).\]

Here, \(x_0\) is the normalized intracellular concentration of ssDNA, \(\Psi\) is the UV energy fluence, \(r\) is the dimensionless time, \(r_2\) is the dimensionless time of replication termination, \(\xi\) is the integration variable, and \(q_i\) \((i = 1, \ldots, 4)\) are the dimensionless kinetic parameters of the model (see Appendix A).

Since most of the experiments on UV-induced mutagenesis are carried out in dark conditions, the influence of photoreactivation is excluded in this model. The source of the SOS-inducing signal is represented here as a superposition of two processes, namely, DNA replication and NER. The dynamics of ssDNA induction is determined on the basis of the balance between the five fractions of DNA disorders appearing in the bacterial chromosome after UV irradiation. They include the dimers initially produced by UV, the gaps opposite dimers, the dimers located in front of the replisome, the dimers removed by NER in front of the replisome, and the repaired gaps. This approach takes into account the facts that NER is pre-replication repair and its ferments operate in front of the replisome in the chromosome. The inducing signal is determined as the ss-gaps located opposite thymine dimers not removed by NER and formed as a consequence of replication restart beyond a dimer. DNA replication and dimer removal are regarded as parallel processes where NER leads. Therefore, the inducing signal terminates together with replication. The ssDNA sequences generated at any later stage of repair machinery are not considered as a contribution to the inducing signal for SOS response. Since NER cuts UV-generated dimers, it significantly reduces the number of primary DNA lesions. Thus, NER directly contributes to the total pool of SOS signal and indirectly impacts the number of mismatches and finally produced mutations.

The model assumptions mentioned above imply introducing the replication termination time, which depends on the number of produced dimers and, consequently, on UV energy fluence. Therefore, if the computation interval is less than the time of replication termination, the model of inducing signal should be calculated until the end of the stated interval. Otherwise, a calculation should be done until the replication time termination because after that, the inducing signal cannot be generated.

Further steps of SOS system functioning are connected with the activity of more than 40 genes controlled by LexA repressor. In our model, we consider four of these genes which make a major contribution to the regulation of the bacterial SOS function. They are the recA, lexA, umuD, and umuC genes. After ssDNA has been produced, the product of the recA gene binds to it and transforms into the active RecA protein. The activated RecA protein has a specific protease conformation, which makes it able to cleave the LexA repressor as well as a number of other proteins. A decrease in the LexA protein level leads to the enhanced expression of the repressed genes including recA, lexA, umuD, and umuC. Despite an increase in LexA gene expression, it does not lead to raising the level of LexA protein due to its immediate cleavage by RecA (Krasavin and Kozulek, 1991). The RecA protease also cleaves the UmuD protein, transforming it into the UmuD active form. The normal and active forms of the UmuD protein can form dimers of three types, namely, UmuD2, UmuDD, and UmuD2 (Burckhardt et al., 1988; Woodgate et al., 1989). These dimers are able to interact with UmuC, forming UmuD2C, UmuDDC, and UmuD2C complexes, respectively.

The main role in an induced mutation process in *E. coli* belongs to UmuD2C (or DNA polymerase V, PolV). This protein complex is able to form a multienzyme complex called PolV Mut, which includes molecules of the RecA-protease, SSB-proteins, and subunits of DNA polymerase III. The PolV Mut complex is able to fill the remaining single-stranded gaps by inserting nucleotides in a random manner. This allows a chromosome to be fully replicated. However, thedaughter DNA sequence contains defects due to the specificity of PolV Mut-mediated gap filling.

The UmuD2C and UmuDDC complexes play a supplementary role in the SOS function. UmuD2C is involved in cell cycle regulation. It stops replicative DNA synthesis and allows the TLS process to be realized in the presence of SOS-inducing damage. The UmuDDC has an inhibiting function in SOS mutagenesis that consists in suppressing the UmuD activity (Smith and Walker, 1998).

To simulate the above-mentioned stages of SOS repair, we used a quantitative model proposed earlier (Belov et al., 2009). The
dynamic change of the concentration of the main SOS proteins is expressed in general by the following differential equations:

\[
d\frac{X_i}{dt} = V_+ (X_i, X_0) - V_- (X_i, X_0),
\]

where \(X_i (i = 1, ..., n)\) is the \(i\)-th regulatory protein intracellular concentration, \(X_0\) is the level of an inducing signal (ssDNA), \(t\) is time, \(V_+\) is the regulatory protein synthesis rate, and \(V_-\) is the regulatory protein degradation rate. The functions \(V_+\) and \(V_-\) describe the protein accumulation and degradation, respectively. The equations reflecting the dynamic change of each SOS protein's concentration as well as the estimation of their parameters and initial conditions are presented in the respective subsections of Appendix A.

### 2.2. A model of translesion synthesis

To describe the induction of mismatches during SOS repair, we have used a translesion synthesis model published earlier (Belov et al., 2009). When we analyze the probability of errors appearing during TLS we need to model a random quantity which is the number of mismatches in DNA chain. These mismatches appear during fixed time under the assumption that all these events are independent and occur with some average fixed intensity. Since the number of nucleotides \(n\) supplied by PolV Mut is large (about several thousand) and a mutation probability in each individual nucleotide pasting is low (about \(10^{-4}\)–\(10^{-3}\) (Tang et al., 2000)), we can conclude that the number of mismatches \(m\) appearing in DNA chain is distributed by Poisson's law:

\[
P_n (m) = \frac{\alpha^m e^{-\alpha}}{m!}.
\]

The calculation of the Poisson distribution parameter \(\alpha\) is performed using a special code developed earlier (Belov et al., 2009). The code models PolV Mut moving along DNA and finds a fluence-time dependence of the parameter \(\alpha\) using the fluence-time dependence of the PoV complex, which is characterized by the variable \(x1\) from Eq. (A.1) (see Appendix A). Here we put the number of PolV Mut molecules equal to the number of PolV molecules because it is known that only one PolV molecules participate in the resynthesis of each ssDNA site (Pham et al., 2001). In this research, we also consider the possibility of PolV Mut involvement in the replication of ssDNA sites which do not contain thymine dimers. This assumption is based on the fact that PolV Mut can realize TLS at undamaged DNA (Tang et al., 2000). In our model, we separate the replication processes realized directly at the thymine dimers, at the rest of the ssDNA gaps without dimers, and at the other undamaged ssDNA sites contained in the cell at this moment. The developed code models the occurrence of nucleotide mismatches during TLS taking into account the individual peculiarities and characteristics of all three processes.

The replication at the thymine dimers described as a step-by-step calculation of number \(L_{\text{td}} (\tau, \Psi)\) of nucleotides inserted by PolV Mut. The formula for calculation is

\[
L_{\text{td}} (\tau, \Psi) = x_{11} (\tau, \Psi) v_{\text{td}} (\tau - \tau_m + \frac{I_1}{V_{\text{td}}})
\]

In Eq. (3), the number of the PolV Mut molecules is equal to the number of the DNA PolV molecules \(x_{11} (\tau, \Psi)\). \(v_{\text{td}}\) is the velocity of PolV Mut moving at the moment when it passes thymine dimers, \(\tau_m\) is the normalized time when the last of the \(x_{11} (\tau, \Psi)\) PolV molecules present at this moment was synthesized. The calculation of \(v_{\text{td}}\) is based on the data on the average length of the ssDNA gap \(I_1\) and the termination time of the whole TLS process for a specific UV energy fluence. \(v_{\text{td}}\) is the velocity of PolV Mut movement during the filling of the ssDNA sites.

The replication at the single-stranded DNA gaps excluding thymine dimers is described by the formula

\[
\begin{aligned}
L_{s} (\tau, \Psi) &= x_{11} (\tau, \Psi) v_{s} (\tau - \tau_m), \\
x_{11} (\tau, \Psi) &\leq N_1 (\Psi), \\
L_{s} (\tau, \Psi) &= N_1 (\Psi) v_{s} (\tau - \tau_m), \\
x_{11} (\tau, \Psi) &> N_1 (\Psi),
\end{aligned}
\]

where \(L_{s} (\tau, \Psi)\) is number of nucleotides supplied by PolV Mut at current moment, \(N_1 (\Psi)\) is the number of ssDNA gaps formed till the replication termination. \(N_1 (\Psi)\) was calculated as in detail described by Aksenov (1999) and Aksenov et al. (1997).

PolV Mut-mediated replication at undamaged ssDNA is described using the following approach. Since the PolV Mut complex has a great affinity to single DNA associated with the RecA protein, it can be concluded that the TLS begins and ends at ssDNA gaps. Consequently, the inclusion of PolV Mut in the process at the undamaged DNA occurs at a time when its number becomes to exceed the number of single-stranded gaps produced by nucleotide excision repair. Accordingly, the termination of synthesis on undamaged DNA occurs at a time when the number of molecules of PolV Mut becomes equal to the number of single spaces. Sequential calculation of the number \(L_{\text{und}} (\tau, \Psi)\) of supplied nucleotides in this case is described by the formula

\[
\begin{aligned}
L_{\text{und}} (\tau, \Psi) &= 0, \\
x_{11} (\tau, \Psi) &\leq N_1 (\Psi), \\
L_{\text{und}} (\tau, \Psi) &= v_{\text{sd}} (x_{11} (\tau, \Psi) - N_1 (\Psi)), \\
x_{11} (\tau, \Psi) &> N_1 (\Psi).
\end{aligned}
\]

The fluence-time dependence for the parameter \(\alpha\) is calculated in our code by the following formula:

\[
\alpha (\tau, \Psi) = P_a L_{\text{td}} (\tau, \Psi) + L_0 (\tau, \Psi) + P_{\text{td}} L_{\text{td}} (\tau, \Psi),
\]

where \(P_a\) is a probability of mismatch induction during nucleotide pasting by PolV Mut on DNA sites which do not contain thymine dimers; \(P_{\text{td}}\) is the mismatch induction probability during thymine dimer processing. The probability \(P_{\text{td}}\) was calculated as follows:

\[
P_{\text{td}} = P_1 P_A + P_2 P_B + P_3 P_C + P_4 P_D,
\]

where \(P_A\), \(P_B\), \(P_C\), and \(P_D\) are the probabilities of a single mismatch occurrence during processing TT (6–4) photoproduct with a 3-end, TT (6–4) photoproduct with a 5-end, cys–syn cyclobutane photodimer with a 3-end, and cys–syn cyclobutane photodimer with a 5-end, respectively. \(P_1\), \(P_2\), \(P_3\), and \(P_4\) are the probabilities of generating each of the four types of lesions. Estimation of the TLS model parameters is presented in Appendix A.

In our model, the variable \(\alpha (\tau, \Psi)\) calculated by formula (6) is the parameter through which the PolV concentration contributes first to the mismatch induction, and then to the mutation rate.

### 2.3. A model of DNA mismatch repair

Following the induction of SOS response, the frequency of misincorporated bases that are the substrate for MMR increases as compared to normal conditions. Recently a number of experimental observations led to the hypothesis that the MMR system significantly reduces the error rates during DNA replication by recognizing and correcting mismatches which prevent normal replication (Kornberg and Baker, 1992). It was also found that MMR can process the incorrect bases opposite UV-induced photoproducts which were not removed by early repair processes like photoreactivation or nucleotide excision repair and during SOS response (Hongbo et al., 2000). Summarizing all these findings, we consider the methyl-directed excision of incorrect bases inserted
MMR exonucleases could proceed from 5' to 3' or from 3' to 5' end to the mispair (Li, 2008). Exol and ExoVII digest the DNA strand in the 3' to 5' direction, RecJ degrades it from 5' to 3', and ExoVII can excise DNA in both directions (Dutra et al., 2007). The resulting single-stranded gap is filled by DNA polymerase III holoenzyme (PolIII) with SSB. The remaining DNA strand is joined to existing one by the DNA ligase (Modrich and Lahue, 1996). The reaction network, which highlights mass transfer and regulatory reactions, is presented in Fig. 2.

To simulate the dynamic changes of the MMR protein levels, we used reversible mass-action kinetics based on Eq. (2). We singled out two MMR pathways with different exonucleases possessing polarity 3' or 5'. The dimensionless equations for each protein and intermediate complexes of the MMR system are given in Appendix B (Eq. B.1). In this study, we assume that 3' and 5' MutH-mediated nicks as well as the involvement of exonucleases possessing the same end specificity are equally probable. We also take into account the fact that none of the MMR enzymes except UvrD are controlled by the SOS system, i.e. their synthesis is not controlled by the LexA protein. However the expression of the uvrD gene producing helicase II depends on the intracellular concentration of the LexA repressor (Easton and Kushner, 1983; Courcelle et al., 2001). To describe the regulation of the uvrD transcription by the LexA protein, we used the model of gene regulation used in many papers (Belov et al., 2009; Aksenov et al., 1997; Aksenov, 1999). The first term in the equation for the UvrD helicase ($y_{uvrD}$ in Eq. B.1) describes LexA-regulated synthesis.

As the function $a(\tau, \Phi)$ depends on UV energy fluence, the variable $y_0$ gives the fluence-time dependence of remained mismatches. In our model these mismatches are considered as a contribution of SOS network to the final mutation frequency. Here to calculate the mutagenesis in individual genes we applied a special model approach.

### 2.4. A model of UV-induced mutagenesis

The developed model allows quantitative estimation of the mutagenesis in individual genes of E. coli bacteria. The yields of UV-induced mutations measured in most cases can be described by the following expression based on the formula introduced by Krasavin and Kozubek (1991):

$$Z_m(\Psi)/Z(\Psi) = \theta_0 + \theta_1 \Psi + \theta_2 \Psi(1 - \exp(-\theta_2 \Psi)),$$

where $Z_m(\Psi)$ and $Z(\Psi)$ are the numbers of mutants and survived cells, respectively; $\Psi$ is the energy fluence of UV radiation. The
following interpretation of function components is suggested by Krasavin and Kozubek (1991). The linear component \( \theta_1Y \) characterizes the mutagenic lesions converted to stable mutations during constitutive repair or DNA replication. This process is seemingly defined by DNA PolIII processing effectiveness (Tang et al., 2000; Borden et al., 2002). \( \theta_2Y \) is proportional to the yield of premutational (or initial) DNA lesions in an individual gene. \( (1-\exp(-\theta_2Y)) \) is the fraction of mutations induced by mutagenic repair. In our study we have introduced an additional term \( \theta_0 \) into original formula proposed by Krasavin and Kozubek (1991). This parameter is the constant characterizing spontaneous level of mutagenesis in bacteria defective in some repair functions.

Taking into account the interpretation of terms in Eq. (8) we conclude that a contribution of SOS and MMR systems to the final mutation frequency can be realized only through the \( \theta_0 \) parameter. Using the models of SOS network, TLS, and MMR, it is possible to determine \( \theta_0 \) through the variable \( y_0 \) (Eq. B.1) which characterizes the amount of mismatches remained after MMR. Therefore, the full chain of events quantitatively described in our model is UV action \( \rightarrow \) thymine dimers \( \rightarrow \) NER and DNA replication \( \rightarrow \) ssDNA \( \rightarrow \) SOS network induction \( \rightarrow \) PolIV induction \( \rightarrow \) PolIV Mut action \( \rightarrow \) mismatch induction \( \rightarrow \) MMR action \( \rightarrow \) final mutation frequency \( Z_m(\Psi)/Z(\Psi) \).

3. Evaluation of the model parameters

The kinetic parameters of the models of the SOS network and TLS were estimated previously (Belov et al., 2009) and presented in Appendix A. Most of the MMR rate constants were determined by fitting the developed model (Eq. B.1) to the in vitro experimental data on the MMR kinetics for the 3' and 5' pathways (Pluciennik et al., 2009). Other parameters were obtained directly from experimental data or calculated in our previous papers (see Appendix B). To calculate the parameters of UV mutagenesis model we have proposed a special approach described below.

3.1. Parameters of the MMR model

3.1.1. Estimation of parameters for 3' and 5' incision

The first group of parameters was evaluated by fitting the model curve for \( y_8 \) to experimental data on the MutH-mediated incision stage of MMR (Pluciennik et al., 2009). For the incision in polarity 3', the parameters \( k_1, k_2, k_4, k_5, k_6, k_7, \) and \( k_8 \) were estimated. The relation of these dimensional parameters to dimensionless ones \( (p_l) \) is presented in the corresponding section of Appendix B. To estimate the first group of parameters, we have set the initial conditions for Eq. (B.1) according to the reactant concentrations for the in vitro reaction: \( Y_{00} = 2.4 \times 10^{-9} \text{ M (mismatches)} \), \( Y_{01} = 3.7 \times 10^{-8} \text{ M (MutS2)} \), \( Y_{03} = 2.5 \times 10^{-8} \text{ M (MutL2)} \), and \( Y_{05} = 1.0 \times 10^{-9} \text{ M (MutH)} \). Since the number of GATCm sequences equals the total number of mismatches, we set \( Y_{06} = Y_{00} \). Other MMR species were assumed to be zero at \( t = 0 \). The function \( a(r, \Psi) \) was also set to zero for all \( t \). The dimensionless initial conditions \( (y_{0l}) \) for Eq. (B.1) were set respectively. The parameter \( k_3 \) was also set to zero for all fitting procedures performed with in vitro data. In our model, \( k_3 \) is the rate constant of the non-specific losses of MutS2, MutL2, MutH, GATCm, UvrD, exonucleases, PolIII, and DNA ligase in cells growing exponentially. As the contribution of spontaneous protein degradation is negligible for all these species, this parameter has meaning for in vivo calculations exclusively. The results of the parameter evaluation for this MMR stage are presented in Fig. 3. The dashed curve corresponds to the \( y_8 \) variable recalculated in femtomoles.

3.1.2. Estimation of the parameters for 3' and 5' excision

The second group of parameters was estimated by fitting the model curve \( y_{14} \) to experimental data on the excision stage of MMR. To estimate \( k_{11}, k_{12}, k_{13}, k_{14}, k_{15}, \) and \( k_{16} \), we have used the data on excision mediated by ExoI (polarity 3') and RecJ (polarity 5') (Pluciennik et al., 2009). The initial conditions for Eq. (B.1) were set as follows according to the experimental procedure: \( Y_{00} = 2.4 \times 10^{-9} \text{ M (mismatches)} \), \( Y_{01} = 3.7 \times 10^{-8} \text{ M (MutS2)} \), \( Y_{03} = 2.5 \times 10^{-8} \text{ M (MutL2)} \), \( Y_{05} = 1.0 \times 10^{-9} \text{ M (MutH)} \), \( Y_{06} = Y_{00} \) (GATCm), \( Y_{09} = 1.2 \times 10^{-8} \text{ M (UvrD)} \), and \( Y_{012} = 1.8 \times 10^{-8} \text{ M (ExoI)} \). The initial conditions for other species as well as the first term in the equation for \( y_9 \) describing the synthesis of UvrD were assumed to be zero. The reaction rates \( k_1, k_2, k_4, k_5, k_7, \) and \( k_8 \) were taken from evaluating the 3' incision stage. The obtained curve corresponding to \( y_{14} \) is presented in Fig. 4.

Using experimental data for excision by RecJ, we have obtained the rate constants \( k_{17}, k_{18}, \) and \( k_{19} \). The initial conditions for the level of mismatches, MutS2, MutL2, MutH, GATCm, and UvrD were set as for the 3' excision. For RecJ, we have set \( Y_{015} = 7.8 \times 10^{-9} \text{ M} \). The resulting curve compared to experimental data is presented in Fig. 4. The values of the estimated parameters for excision stages '3' and '5' of MMR are presented in Table B.1.

3.1.3. Estimation of the parameters for the MMR polymerization stage

The last group of parameters was evaluated by fitting the curve for \( y_{19} \) to experimental data on single-stranded gap filling by DNA PolIII (Pluciennik et al., 2009). For the case when excision is performed by ExoI, we have set the initial conditions as follows: \( Y_{00} = 2.4 \times 10^{-9} \text{ M (mismatches)} \), \( Y_{01} = 3.7 \times 10^{-8} \text{ M (MutS2)} \), \( Y_{03} = 2.5 \times 10^{-8} \text{ M (MutL2)} \), \( Y_{05} = 1.0 \times 10^{-9} \text{ M (MutH)} \), \( Y_{06} = Y_{00} \) (GATCm), \( Y_{09} = 1.2 \times 10^{-8} \text{ M (UvrD)} \), \( Y_{012} = 1.8 \times 10^{-8} \text{ M (ExoI)} \),

![Fig. 3. Incision of a 3' (●) and 5' (▲) hemimethylated heteroduplexes by activated MutH in the presence of MutS and MutL. N is the concentration of incised DNA. The curves are the calculated results; the dots are the experimental data (Pluciennik et al., 2009).](image-url)
bacterial chromosome. Then the average number of premutational DNA lesions in a gene is $\theta_2 = L_1 m_0 / L_0$. As we are estimating the mutation frequency in the *lacZ* gene, we have to assign the corresponding length of this gene and the length of *E. coli*’s whole K-12 MG1655 genome (Table C.1).

3.2.2. $\theta_3$ parameter for *mut+$^* $ bacterial strains

To estimate the parameter $\theta_3$ through which the SOS and MMR systems contribute to the final mutation frequency, we used the following assumptions. The dependence of $y_0$ on UV energy fluence can be approximated by a linear function with a certain slope coefficient $k_s$. Such approximation is suitable for any time point within the calculation interval. The obtained slope coefficient is proportional to the parameter $\theta_3$. However, the meaning of $\theta_3$ implies finding the exact form of this proportionality. As we are calculating the yield of the mutations produced exclusively in an individual *E. coli*’s gene, we should take into account only the mismatches remaining within the length of this gene. Therefore, the final expression for this parameter will be $\theta_3 = L_1 k_s / L_0$.

To find the numerical value for $\theta_3$, we performed the following procedure consisting in running simultaneously models of the SOS-network, TLS, and MMR with corresponding sets of parameters and initial conditions. First, using the SOS network model, we have calculated the fluence-time dependence for PolV represented as the variable $x_{11}$ (Fig. 6). The SOS network model was computed using Eqs. (1) and (A1). The initial conditions and kinetic parameters were set as described in Appendix A. Then,

![Image](327x70 to 547x233)

![Image](334x282 to 540x437)

![Image](595.3x793.7)

**Fig. 5.** Gap filling of an excised 3' (●) and 5' (●) heteroduplexes by PolII. $N$ is the concentration of rebuilt DNA. The curves are the calculated results; the dots are the experimental data (Easton and Kushner, 1983).

and $Y_{017} = 7.9 \times 10^{-8}$ M (PolIII). The parameters $k_1, k_2, k_4, k_5, k_6, k_7, k_8, k_{11}, k_{12}, k_{13}, k_{14}, k_{15}$, and $k_{16}$ were set based on our previous findings. Finally, we have obtained the numerical values for $k_{20}$, $k_{31}$, and $k_{32}$. The dashed curve in Fig. 5 shows the agreement between the calculated curve ($Y_{10}$) and experimental data. The same values of the parameters $k_{20}$, $k_{21}$, and $k_{22}$ satisfy the experimental data for single-stranded gap filling, when excision was done by RecJ (see Fig. 5). In this case, we have set the same initial conditions except the levels of ExoI and RecJ: $Y_{012} = 0$ (ExoI), $Y_{013} = 7.8 \times 10^{-9}$ M (RecJ). The set of the newly obtained parameters is presented in Table B.1.

### 3.2. Estimation of the UV mutagenesis model parameters

#### 3.2.1. $\theta_0$, $\theta_1$, and $\theta_2$ parameters

The parameter $\theta_0$ depends on the bacteria strain and can be estimated directly from experimental data. For the strains with normal DNA repair functions, we assume this parameter to be zero. For the strains containing a mutation in either *mutS, mutL,* or *mutH* gene, we have set the corresponding values of $\theta_0$, $\theta_0$, $\theta_0$, and $\theta_0$, according to the experimental data (Hongbo et al., 2000). For the strain defective both in *umu* and *mut* genes, we have set the corresponding value for $\theta_0, \theta_{umu, mut}$ from the same paper. The coefficient $\theta_1$ of the linear component in Eq. (8) can be defined as the mutagenic effectiveness of DNA PolIII processing according to Drake (1969) (see Appendix C).

The coefficient $\theta_2$ characterizing the number of premutational lesions in a gene is defined as follows in our model. Let us assume that $L_1$ base pairs is the length of a gene, $L_0$ is the length of *E. coli*’s whole genome, and $m_0$ is the yield of the initial lesions per full

![Image](595.3x793.7)

**Fig. 6.** Three-dimensional plot showing the PolV level change with time and depending on the UV energy fluence. $N$ is the number of PolV molecules per one cell.

![Image](595.3x793.7)

**Fig. 7.** Dependence of the mean number of the occurring mismatches $a$ on time and UV energy fluence.

![Image](595.3x793.7)
using the TLS model, we have estimated the function $a(t, \Psi)$ (Fig. 7). Here we used Eqs. (3)-(7) of this model implemented in the developed program code. The parameters for this calculation are also presented in Appendix A. This code gives discrete values for $a(t, \Psi)$ to reflect the specifics of nucleotide pasting by PolIV Mut. The first large peak of the function $a(t, \Psi)$ is caused by processing the over-produced PolIV Mut molecules at undamaged DNA sites; other disturbances are caused by the process of thymine dimer bypass (Belov et al., 2009). The energy fluence range for these calculations was set as in Belov et al. (2009): $0.5 \leq \Psi \leq 100 \text{ Jm}^{-2}$. To compute all ODE systems presented in this paper, we have used the fourth-order Runge–Kutta method.

After that, we have calculated the dependence of the remaining mismatches $y_0$ on UV energy fluence and on time using the MMR model (Fig. 8). Since we use the function $a(t, \Psi)$ as the input data, we obtain the discrete values for $y_0$ as well. The MMR model was initialized with the kinetic parameters estimated earlier from the fitting procedure (Table B.1). The in vivo data on the normal intracellular levels of all MMR proteins were used as the initial conditions for calculating $y_0(t, \Psi)$. The numerical values for $Y_{01}$ (MutS2), $Y_{03}$ (MutL2), $Y_{05}$ (MutH), $Y_{09}$ (UvrD), $Y_{012}$ (3’ exonucleases), $Y_{015}$ (5’ exonucleases), $Y_{017}$ (PolIII), and $Y_{021}$ (DNA ligase) are estimated directly from experimental data and summarized in Table B.1. As we assume an equal probability of each pathway corresponding to different exonuclease activity, we set the normal intracellular levels of 3’ and 5’ exonucleases as follows:

\[ y_{012} = [\text{Exol} + 0.5[\text{ExoVII}] + [\text{ExoX}], \quad y_{015} = 0.5[\text{ExoVII}] + [\text{Rec}]. \]

The initial level of mismatches was set as $Y_{00} = a(0, \Psi) = 0$.

At the next step, we have approximated the obtained $y_0(t, \Psi)$ dependence by a linear function with the slope coefficient $k_s$. The approximation was carried out at the time point corresponding to the maximal concentrations of PolIV (Fig. 9). Using the parameter $k_s$, we have obtained the numerical value of $\theta_3$ (see Table C.1).

3.2.2. Parameters $\theta_{1, \text{mut}}$, $\theta_{3, \text{umu}}$ for mut$^-$ and umu$^-$ bacterial strains

To calculate the parameter $\theta_{1, \text{mut}}$ for bacteria defective in the MMR function, we have used the same procedure as for $\theta_3$ except for the different set of initial conditions for the MMR model. For the strains defective in either mutS, mutL, or mutH gene, we have set to zero the initial levels of the corresponding proteins: mutS$^-$, $Y_{01} = 0$; mutL$^-$, $Y_{03} = 0$; mutH$^-$, $Y_{09} = 0$. For all the mutants, we obtained the identical value of the slope coefficient $k_s$ mut$^-$ and then $\theta_{1, \text{mut}}$ (Table C.1).

We also estimated the parameter $\theta_{3, \text{umu}}$ for the umu$^-$ strains containing a mutation in the umu-genes of E. coli. This mutation prevents the formation of PolIV complex and, therefore, leads to the absence of the mutagenic effect of SOS repair. Here we set to zero the initial levels of either UmuD or UmuC protein in the SOS network model: $X_{04} = 0$ or $X_{05} = 0$, respectively. In both cases, we have obtained $\theta_{3, \text{umu}} = 0$. For the umu$^-$ mut$^-$ bacterial strains defective in both SOS and MMR systems, we also obtained $\theta_{3, \text{umu}, \text{mut}} = 0$.

4. Calculation of UV-induced mutagenesis

Using our model we have performed calculations of the mutation frequency in E. coli strains with different genotypes. The mutagenic effect of UV radiation was modeled for the cells with normal SOS and MMR functions and for mutants defective in different MMR genes. In this study, we have estimated the mutation frequency in the E. coli’s lacZ gene encoding $\beta$-galactosidase. The results of modeling were compared with experimental data on the revertant frequency in two alleles at lacZ codon 461, which reverts via CCC→CTC and CTT→CTC transitions (Hongbo et al., 2000).

4.1. Mutagenesis in bacteria with the normal SOS and MMR functions

First, we have simulated the UV-induced mutation frequency for the strains with the normal functioning of the SOS and MMR systems. The final mutation frequency was calculated by Eq. (8) using the parameter $\theta_3$. As we obtained a linear dependence of $\theta_3$ on UV energy fluence (Fig. 9), we can extend the range of $\Psi$ for calculations with formula (8) and start calculations of the mutation frequency at $\Psi = 0 \text{ Jm}^{-2}$. The parameter $\theta_3$ can be neglected in this case. Therefore, we set it to zero. The coefficients $\theta_1$ and $\theta_2$ are the same for the normal and mutagenic strains (Table C.1). The results of calculation for bacteria with the normal functioning of the SOS system and MMR are presented in Fig. 10 and indicated as mut$^+$. The corresponding experimental data are also shown in this plot (Hongbo et al., 2000).

4.2. Mutagenesis in bacteria defective in MMR functions

4.2.1. Calculations for mutS$^-$ strains

To estimate the UV-induced mutation frequency for the mutS$^-$ mutant strains, we have used the same $\theta_1$ and $\theta_2$ parameter values as for cells with normal MMR. The parameter $\theta_{0, \text{mutS}}$ was set according to the spontaneous mutation rate demonstrated by the mutS$^-$ strains (Hongbo et al., 2000). Therefore, at $\Psi = 0 \text{ Jm}^{-2}$, the curve computed
mutH energy by 2.6 times. The results of calculations are in accordance with the experimental data with their standard errors of the means (SEM), we can conclude that the model adequately reconstructs the observed mutagenic effect. Here we used the same \( \theta_1, \theta_2, \) and \( \theta_3, \) parameter values as for \( \text{mut}^S \) strains. The difference is just like in the experiment (Hongbo et al., 2000). The model meets the hypothesis on the possible role of the MMR in radiation-induced SOS mutagenesis. Choosing UV radiation as a mutagenic factor for this study is explained by the necessity to indicate the links between MMR and SOS response without any significant influence of other repair systems such as single- and double-strand break repair and base excision repair. Since most of the UV-induced thymine dimers are not removed by NER represent a substrate for SOS repair, it gives an opportunity to identify the direct connections between the biochemical mechanisms of these two systems.

The developed models provide a topological view of the MMR and SOS networks attempted to clarify their biological relations. Using our mathematical approach, we have analyzed all the chain of events from the primary DNA lesion appearance to fixing this lesion as a mutation. The model adequately describes the basic processes of the SOS response and MMR. Mathematical description of these two systems is carried out in compliance with concepts of modern system biology and with simulation methods of studying complex genetic networks.

A special issue should be addressed: the possible feedback of MMR to the SOS network. There are two major stages which could be considered as its possible pathways. The first one is the involvement of UvrD in MMR. As it is known, the synthesis of this protein is regulated by the LexA repressor and, therefore, the level of UvrD increases after UV irradiation. However, the amount of UvrD helicase molecules necessary for successful MMR functioning is not large (one molecule per one mismatch). Therefore, the interactions of UvrD during MMR should not affect the SOS functioning. As it is known, the synthesis of this protein is regulated by the LexA repressor and, therefore, the level of UvrD increases after UV irradiation. However, the amount of UvrD helicase molecules necessary for successful MMR functioning is not large (one molecule per one mismatch). Therefore, the interactions of UvrD during MMR should not affect the SOS functioning. As it is known, the synthesis of this protein is regulated by the LexA repressor and, therefore, the level of UvrD increases after UV irradiation. However, the amount of UvrD helicase molecules necessary for successful MMR functioning is not large (one molecule per one mismatch). Therefore, the interactions of UvrD during MMR should not affect the SOS functioning. As it is known, the synthesis of this protein is regulated by the LexA repressor and, therefore, the level of UvrD increases after UV irradiation. However, the amount of UvrD helicase molecules necessary for successful MMR functioning is not large (one molecule per one mismatch). Therefore, the interactions of UvrD during MMR should not affect the SOS functioning. As it is known, the synthesis of this protein is regulated by the LexA repressor and, therefore, the level of UvrD increases after UV irradiation. However, the amount of UvrD helicase molecules necessary for successful MMR functioning is not large (one molecule per one mismatch). Therefore, the interactions of UvrD during MMR should not affect the SOS functioning. As it is known, the synthesis of this protein is regulated by the LexA repressor and, therefore, the level of UvrD increases after UV irradiation. However, the amount of UvrD helicase molecules necessary for successful MMR functioning is not large (one molecule per one mismatch). Therefore, the interactions of UvrD during MMR should not affect the SOS functioning. As it is known, the synthesis of this protein is regulated by the LexA repressor and, therefore, the level of UvrD increases after UV irradiation. However, the amount of UvrD helicase molecules necessary for successful MMR functioning is not large (one molecule per one mismatch). Therefore, the interactions of UvrD during MMR should not affect the SOS functioning.
In this paper, we have shown how more or fewer functions connected with the activity of the \textit{mut} and \textit{umu} genes affect the mutation frequency, i.e. what influence the system’s different topologies have on the final cell response to irradiation. Here we have provided a possible mechanistic explanation of how a violation of the expression of these genes leads to an increase in mutagenesis in bacterial cells. It is clear that this method can be extrapolated to other SOS genes responsible for assembling the PolV Mut complex. According to our model, violations in the \textit{umuD} or \textit{recA} gene result in the same mutation frequency as in \textit{umuC}-defective strains.

Besides our previous studies, only a few papers are concerned with simulating some quantitative characteristics of TLS (Vaidyanathan and Cho, 2012; Malina et al., 2012). However, these approaches do not provide a system view of the process as well as do not focus on its probabilistic aspects and connections with other repair systems. One of the main features of our models is a clear representation of cause-and-effect relations between two replicate repair networks and the TLS effectiveness. In addition to the quantitative analysis of mutagenic effects, the developed models provide a tool for the detailed analysis of the protein–protein interaction dynamics of the SOS network and MMR system.

In this paper, we focused mainly on the problem of induced mutagenesis. Here we did not include a special mathematical description of the effects that contribute to spontaneous mutagenesis. This is the main reason for using the classical formula for the estimation of the final mutation frequency. The comprehensive reconstruction of the whole mutation process based only on mechanistic models requires the development of additional model approaches for other repair systems. Therefore, we used a more suitable approach based on modeling the biophysical processes behind the parameters of the classical equation.

Taking into account the knowledge of the molecular mechanisms of other \textit{E. coli}’s repair systems, we suggest that our model could be applied for the estimation of mutagenesis induced not only by UV radiation but also by ionizing radiations of different quality. The latter relates mostly to the repair of charged particle-induced clustered DNA lesions because it is supposed that these lesions make up the main substrate for mutagenic SOS repair.

**Acknowledgments**

This study was performed within the research project # 301 “Mathematical Modelling of Genetic Regulatory Networks in Bacterial and Higher Eukaryotic Cells” between Laboratory of Radiation Biology of Joint Institute for Nuclear Research and Cairo University. O. Chuluunbaatar acknowledges a financial support from the RFBR Grant no. 11-01-00523, and the JINR theme 09-6-1060–2005/2013 “Mathematical support of experimental and theoretical studies conducted by JINR”.

**Appendix A. Details of SOS network and TLS models**

Equations of the SOS network model

The equations reflecting the dynamical change of SOS protein’s concentration are designed in our previous study (Belov et al., 2009):

\[
\begin{align*}
\frac{dx_1}{dt} &= \frac{x_{01}(1 + q_{15})}{1 + (x_1/r_1N_A)^{a_1}} - q_9x_1x_2 - x_1, \\
\frac{dx_2}{dt} &= \frac{x_{02}(1 + q_{15})}{1 + (x_1/r_2N_A)^{a_1}} + q_1x_3 - q_8x_2x_3 - x_2.
\end{align*}
\]

\[
\begin{align*}
\frac{dx_3}{dt} &= q_9x_2x_3 - q_1x_3, \\
\frac{dx_4}{dt} &= \frac{x_{04}q_3(1 + q_{15})}{1 + (x_1/r_4N_A)^{a_1}} - q_{11}x_6x_{10} + q_{12}x_7x_7 + q_{13}x_6, \\
\frac{dx_5}{dt} &= \frac{x_{05}q_5(1 + q_{15})}{1 + (x_1/r_5N_A)^{a_1}} - q_{14}x_7x_7 - q_{15}x_8x_8 - q_{17}x_4x_4 - q_{18}x_4x_{11} - q_{19}x_4, \\
\frac{dx_6}{dt} &= q_{20}x_{20}(1 + q_{15})^{b_1} - q_{22}x_6x_7 - q_{23}x_8x_8 - q_{24}x_3x_3 - q_{25}x_5, \\
\frac{dx_7}{dt} &= q_{14}x_4x_4 + q_{16}x_4x_8 + q_{18}x_4x_{11} - q_{20}x_6^2 + q_{17}x_4x_6 - q_{11}x_6x_{10} - q_{12}x_6x_7 - q_{13}x_6, \\
\frac{dx_8}{dt} &= q_{15}x_8^2 - q_{22}x_5x_2 - q_{12}x_5x_7 - q_{27}x_7, \\
\frac{dx_9}{dt} &= q_{20}x_6^2 - q_{10}x_6x_8 - q_{22}x_5x_8 - q_{28}x_8, \\
\frac{dx_{10}}{dt} &= q_{17}x_4x_6 + q_{16}x_4x_8 + q_{18}x_4x_{11} - q_{20}x_6^2 - q_{17}x_4x_6, \\
\frac{dx_{11}}{dt} &= q_{22}x_6x_7 - q_{11}x_6x_{10} - q_{13}x_6, \\
\frac{dx_{12}}{dt} &= q_{24}x_3x_3 + q_{18}x_3x_{11} + q_{11}x_6x_{10} - q_{13}x_{12}. \\
\end{align*}
\]

(A.1)

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>α</td>
<td>0.0116 min⁻¹</td>
<td>Aksenov et al. (1997)</td>
</tr>
<tr>
<td>τ₁</td>
<td>0.17 min</td>
<td>Aksenov et al. (1997)</td>
</tr>
<tr>
<td>τ₀</td>
<td>40 min</td>
<td>Aksenov et al. (1997)</td>
</tr>
<tr>
<td>ν₁</td>
<td>0.7 min⁻¹</td>
<td>Aksenov et al. (1997)</td>
</tr>
<tr>
<td>ℓ₁</td>
<td>900 nucleotides</td>
<td>Aksenov et al. (1997)</td>
</tr>
<tr>
<td>δ₄</td>
<td>0.0231 min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>δ₅</td>
<td>0.0347 min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>δ₆</td>
<td>0.0173 min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>δ₇</td>
<td>0.0154 min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>δ₈</td>
<td>0.0116 min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>δ₉</td>
<td>0.0107 min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>δ₁₀</td>
<td>0.0224 min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>δ₁₁</td>
<td>0.0116 min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>δ₁₂</td>
<td>0.0154 min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>σ₁</td>
<td>1.1 × 10⁵ M⁻¹ min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>σ₂</td>
<td>7.7 × 10⁴ M⁻¹ min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>β₂</td>
<td>3.7 × 10⁴ M⁻¹ min⁻¹</td>
<td>Aksenov et al. (1997)</td>
</tr>
<tr>
<td>β₃</td>
<td>1.9 × 10⁴ M⁻¹ min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>η</td>
<td>4.0 × 10⁴ M⁻¹ min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>μ</td>
<td>5.8 × 10³ M⁻¹ min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>ψ</td>
<td>3.9 × 10³ M⁻¹ min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>a₁</td>
<td>6.2 × 10⁵ M⁻¹ min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>a₂</td>
<td>4.6 × 10⁴ M⁻¹ min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>a₃</td>
<td>4.3 × 10³ M⁻¹ min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>b₂</td>
<td>3.9 × 10³ M⁻¹ min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>γ₁</td>
<td>2.0 × 10⁴ M⁻¹ min⁻¹</td>
<td>Aksenov et al. (1997)</td>
</tr>
<tr>
<td>γ₂</td>
<td>2.0 × 10⁴ M⁻¹ min⁻¹</td>
<td>Aksenov et al. (1997)</td>
</tr>
<tr>
<td>γ₃</td>
<td>5.5 × 10³ M⁻¹ min⁻¹</td>
<td>Aksenov et al. (1997)</td>
</tr>
<tr>
<td>γ₄</td>
<td>2.0 × 10³ M⁻¹ min⁻¹</td>
<td>Aksenov et al. (1997)</td>
</tr>
<tr>
<td>h₁</td>
<td>2.6</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>h₂</td>
<td>2</td>
<td>Aksenov et al. (1997)</td>
</tr>
<tr>
<td>h₄</td>
<td>2</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>h₅</td>
<td>2</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>X₀₅</td>
<td>2.2 × 10⁶ M⁻¹ min⁻¹</td>
<td>Hegde et al. (1995)</td>
</tr>
<tr>
<td>X₀₆</td>
<td>1.19 × 10⁵ M⁻¹ min⁻¹</td>
<td>Banco and Kowalczykowski (1999)</td>
</tr>
<tr>
<td>X₀₄</td>
<td>2.99 × 10⁷ M⁻¹ min⁻¹</td>
<td>Sasanof and Roberts (1990)</td>
</tr>
<tr>
<td>X₀₅</td>
<td>2.49 × 10⁷ M⁻¹ min⁻¹</td>
<td>Goodman and Woodgate (2000)</td>
</tr>
<tr>
<td>X₀₇</td>
<td>2.24 × 10⁷ M⁻¹ min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>X₀₈</td>
<td>1.54 × 10⁷ M⁻¹ min⁻¹</td>
<td>Belowov et al. (2009)</td>
</tr>
<tr>
<td>ν₀</td>
<td>16.8 nucleotides per min</td>
<td>Fujii and Fuchs (2004)</td>
</tr>
</tbody>
</table>
The initial conditions for this model are the following:

\[ x_1(0) = x_{01}, \quad x_2(0) = x_{02}, \quad x_3(0) = 0, \quad x_4(0) = x_{04}, \quad x_5(0) = x_{05}, \quad x_6(0) = 0, \quad x_7(0) = x_{07}, \quad x_8(0) = 0, \quad x_9(0) = x_{09}, \quad x_{10}(0) = 0, \quad x_{11}(0) = 0, \quad x_{12}(0) = 0. \]

In Eq. (A.1), \( x_1, x_2, x_3, x_4, x_5, x_6, x_7, x_8, x_9, x_{10}, x_{11}, \) and \( x_{12} \) are the normalized intracellular concentrations of the LexA, RecA, RecA*, UmuD, UmuC, UmuD*, UmuD2, UmuD2*, UmuD2C, and UmuD2C proteins, respectively; \( q_i \) (\( i = 1, \ldots, m \)) is the normalized constant of the \( j \)-th protein–protein interaction.

Kinetic parameters of the SOS network model

(1) The parameters of Eq. (1): \( \tau = a t, \quad t_a = a t_2, \quad q_3 = \delta_3(\Psi)/\alpha, \quad q_2 = 25t_0/T_0, \quad q_1 = n_1/\alpha, \quad q_4 = n_1T_0, \) \( T_0 \) is replication duration at the normal growth conditions, \( \alpha \) is the rate constant of the processes of the nonspecific loss of the lexA gene product; \( N_\text{A} \) is the Avogadro constant; \( l_1 \) is an average length of a single-stranded DNA gap formed during the replication of sites containing thymine photodimers, and \( t_2 \) is the replication termination time. For \( \delta_3(\Psi) \) the following expression takes place (Belov et al., 2009):

\[ \delta_3(\Psi) = 0.147 \exp \left( \frac{1}{1 + 0.359\Psi} \right) \text{ min}^{-1}. \]

(2) The parameters of Eq. (A.1):

\[ x_i = X_N A_i \quad (i = 1, \ldots, 12), \quad q_8 = \frac{X_01}{T_1}, \quad q_9 = \delta_5(\Psi)/(\theta_0 x N A), \quad q_7 = \frac{X_01}{T_2}, \quad q_6 = \delta_6(\Psi)/(\theta_0 x N A), \]

\[ q_0 = (e x_0 + \delta_4)/x_0, \quad q_1 = \delta_5(\Psi)/x_0, \quad q_2 = \delta_6(\Psi)/x_0, \quad q_3 = \delta_7(\Psi)/x_0, \quad q_4 = \delta_8(\Psi)/x_0, \quad q_5 = \delta_9(\Psi)/x_0, \quad q_6 = \delta_{10}(\Psi)/x_0, \quad q_7 = \delta_{11}(\Psi)/x_0, \quad q_8 = \delta_{12}(\Psi)/x_0. \]

Here \( x_01 = X_01/X_01 = 1, \quad x_02 = X_02/X_02, \quad x_03 = X_03/X_03, \quad x_04 = X_04/X_04, \quad x_05 = X_05/X_05, \quad x_06 = X_06/X_06, \) and \( x_07 = X_07/X_07, \) \( x_08 = X_08/x_{08}, \) \( x_09 = X_09/X_09, \) \( x_{10} = X_{10}/X_{10}, \) and \( x_{11} = X_{11}/X_{11}, \) \( x_{12} = X_{12}/X_{12} \) are constitutive concentrations of the LexA, RecA, UmuD, UmuC, UmuD2, and UmuD2C proteins, respectively (Table A.1).

Parameters of the TLS model

The translesion synthesis model has the following parameter values: \( P_0 = 2.1 \times 10^{-4} \) (Belov et al., 2009), \( P_0 = 0.875, \) \( P_0 = 0.078, \) \( P_0 = 0.02, \) and \( P_0 = 0.048 \) (Tang et al., 2000; Livneh, 2000). We conclude that if we consider any of the two basic UV-induced DNA lesions (either TT (6-4) photoproduct—20% or cys–syn cyclobutane photodimer—80% (Wang, 1976; Cadet and Vigny, 1990)), then the 3′-end or 5′-end lesions are equally probable. Therefore, according to Eq. (7), the probabilities of generating each of the four types of lesions were found to be \( P_1 = P_2 = 0.1, \quad P_3 = P_4 = 0.4, \) and \( P_{id} = 0.12. \)

Appendix B. MMR model details

Equations of the MMR model

The dynamical change of MMR protein’s concentrations is described by the following system of ordinary differential equations:

\[ \frac{dy_0}{dt} = \alpha(\Psi) - p_1 y_1 y_2 - p_2 y_2, \]

\[ \frac{dy_1}{dt} = y_0 y_1 y_2 (p_3 + p_1 y_3) y_1 + p_1 y_2 y_3, \]

\[ \frac{dy_2}{dt} = p_1 y_0 y_1 y_2 (p_3 + p_4 y_3 y_4) y_3 y_4, \]

\[ \frac{dy_3}{dt} = y_0 y_1 y_2 y_3 (p_3 + p_4 y_3 y_4) y_3 y_4, \]

\[ \frac{dy_4}{dt} = p_1 y_0 y_1 y_2 y_3 y_4 (p_3 + p_4 y_3 y_4). \]

The initial conditions for this system are the following:

\[ y_0(0) = 0, \quad y_1(0) = y_{01}, \quad y_2(0) = 0, \quad y_3(0) = 0, \quad y_4(0) = y_{04}, \quad y_5(0) = 0, \quad y_6(0) = 0, \quad y_7(0) = 0, \quad y_8(0) = 0, \quad y_9(0) = 0, \quad y_{10}(0) = 0, \quad y_{11}(0) = 0, \quad y_{12}(0) = 0, \quad y_{13}(0) = 0, \quad y_{14}(0) = 0, \quad y_{15}(0) = 0, \quad y_{16}(0) = 0, \quad y_{17}(0) = 0, \quad y_{18}(0) = 0, \quad y_{19}(0) = 0, \quad y_{20}(0) = 0, \quad y_{21}(0) = 0, \quad y_{22}(0) = 0, \quad y_{23}(0) = 0. \]

In Eq. (B.1) \( y_0 \) is the normalized intracellular level of the mismatches (Mism), \( \alpha(\Psi) \) is the function describing the increase of mismatches produced by the PoIV Mut complex. \( y_1 \) is the concentration of the MutS dimer, which recognizes a mismatch and binds to it reversibly forming an intermediate MismMuts2, complex \( y_2 \). \( y_3 \) represents the normalized concentration of the MutL dimer, which joins the MismMuts2 complex and forms the next intermediate MismMuts2L2 (\( y_4 \)). \( y_5 \) is the concentration of the MutH protein interacting with the methylated GATCm sequence (\( y_6 \)) with the production of the GATCMutH complex (\( y_j \)). \( y_{i+4} \) represents the level of nicked DNA after the interaction of MismMuts2L2 complexes with GATCMutH. The molecules of the MutS2, MutL2, and MutH proteins remain joined to the nicked DNA strand. The following strand unwinding by the UvrD-helicase (\( y_9 \)) can be represented as a typical enzymatic reaction with the intermediate complex \( y_10 \) and resulting detachment of MutS2, MutL2, MutH, and UvrD. Since the synthesis of the UvrD helicase is SOS-dependent, we introduced the normalized concentration of the LexA protein (\( x_1 \)) into the equation for \( y_9 \). The kinetics of LexA is calculated using the model of SOS network (Eq. A.1). The action of UvrD leads to the formation of an unwound DNA sequence \( y_{11} \), which will be processed by two pathways with different exonucleases demonstrating 3′ or 5′ polarity. Here we assign the variables \( y_{12} \) and \( y_{15} \) to the levels of 3′ (ExoI, ExoVII, ExoX) and
Table B.1  
Parameters of MMR model.

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Value</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>$k_1$</td>
<td>$5.2 \times 10^7$ M$^{-1}$ min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_2$</td>
<td>0.195 min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_3$</td>
<td>0.0116 min$^{-1}$</td>
<td>Aksenov et al. (1997)</td>
</tr>
<tr>
<td>$k_4$</td>
<td>$1.3 \times 10^3$ M$^{-1}$ min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_5$</td>
<td>0.118 min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_6$</td>
<td>$1.4 \times 10^4$ M$^{-1}$ min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_7$</td>
<td>$0.088$ min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_8$</td>
<td>$7.0 \times 10^4$ M$^{-1}$ min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{11}$</td>
<td>$3.3 \times 10^{-4}$ min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{12}$</td>
<td>0.045 min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{13}$</td>
<td>0.221 min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{14}$</td>
<td>$3.3 \times 10^{-4}$ min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{15}$</td>
<td>$1.4 \times 10^{-4}$ min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{16}$</td>
<td>0.124 min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{17}$</td>
<td>$6.7 \times 10^3$ M$^{-1}$ min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{18}$</td>
<td>0.131 min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{19}$</td>
<td>0.255 min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{20}$</td>
<td>$3.9 \times 10^7$ M$^{-1}$ min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{21}$</td>
<td>0.052 min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{22}$</td>
<td>0.092 min$^{-1}$</td>
<td>This paper</td>
</tr>
<tr>
<td>$k_{23}$</td>
<td>$1.8 \times 10^5$ M$^{-1}$ min$^{-1}$</td>
<td>Below (2011)</td>
</tr>
<tr>
<td>$k_{24}$</td>
<td>0.667 min$^{-1}$</td>
<td>Below (2011)</td>
</tr>
<tr>
<td>$k_{25}$</td>
<td>0.021 min$^{-1}$</td>
<td>Below (2011)</td>
</tr>
<tr>
<td>$Y_{0}$</td>
<td>$1.4 \times 10^{-7}$ M</td>
<td>Drake (1969)</td>
</tr>
<tr>
<td>$h_0$</td>
<td>2</td>
<td>Feng et al. (1996)</td>
</tr>
<tr>
<td>$h_0$</td>
<td>$3.1 \times 10^{-7}$ M</td>
<td>Feng et al. (1996)</td>
</tr>
<tr>
<td>$h_0$</td>
<td>$1.9 \times 10^{-7}$ M</td>
<td>Feng et al. (1996)</td>
</tr>
<tr>
<td>$h_0$</td>
<td>$2.2 \times 10^{-7}$ M</td>
<td>Feng et al. (1996)</td>
</tr>
<tr>
<td>$h_0$</td>
<td>$5.0 \times 10^{-6}$ M</td>
<td>Feng et al. (1996)</td>
</tr>
<tr>
<td>$h_0$</td>
<td>$8.9 \times 10^{-5}$ M</td>
<td>Feng et al. (1996)</td>
</tr>
<tr>
<td>$h_0$</td>
<td>$6.3 \times 10^{-6}$ M</td>
<td>Feng et al. (1996)</td>
</tr>
<tr>
<td>$h_0$</td>
<td>$5.0 \times 10^{-6}$ M</td>
<td>Feng et al. (1996)</td>
</tr>
<tr>
<td>$h_0$</td>
<td>$5.0 \times 10^{-6}$ M</td>
<td>Feng et al. (1996)</td>
</tr>
</tbody>
</table>

5’ (ExoVII, Red) exonucleases. $y_{13}$ and $y_{16}$ are the intermediate complexes formed by 3’ and 5’ exonucleases respectively. $y_{14}$ represents the amount of single-stranded DNA remained after excision. $y_{17}$ is the normalized concentration of PolIII. $y_{18}$ describes the level of the intermediate complex representing PolIII molecules bound to a single-strand gap. $y_{19}$ is the level of the newly synthesized DNA sequence before ligation. The last MMR stage is characterized in the model by a reaction describing the ligation of a new sequence by a DNA ligase ($y_{20}$), where $y_{21}$ is the intermediate complex and $y_{22}$ is repaired DNA.

In our model $y_{01}$, $y_{02}$, $y_{05}$, $y_{06}$, $y_{08}$, $y_{09}$, $y_{012}$, $y_{015}$, $y_{017}$, and $y_{020}$ are the time-independent parameters representing the normalized initial levels of MutS2, MutL2, MutH, GATCm, UvrD, 3’ and 5’ exonucleases, PolIII, and DNA ligase, respectively. The initial concentrations of all intermediate complexes are assumed to be zero at the beginning of repair. The variables of the model are normalized per initial level of the MutS protein: $y_i = y_i / y_{0i}$. The values of the parameters $Y_{0i}$ are presented in Table B.1.

Kinetic parameters of the MMR model

The dimensionless parameters of Eq. (B.1) are $t = k_{21}$, $p_1 = k_1$ $Y_{01}/k_2$, $p_2 = k_3/k_2$, $p_4 = k_4 Y_{01}/k_2$, $p_5 = k_5/k_2$, $p_6 = k_6 Y_{01}/k_2$, $p_7 = k_7/k_2$, $p_8 = k_8 Y_{01}/k_2$, $p_9 = X_{01}/k_2$, $p_{10} = 1/(k_9 N_0)$, $p_{11} = k_2 Y_{01}/k_2$, $p_{12} = k_3/k_2$, $p_{13} = k_4 Y_{01}/k_2$, $p_{14} = k_5/k_2$, $p_{15} = k_6 Y_{01}/k_2$, $p_{16} = k_7/k_2$, $p_{17} = k_8 Y_{01}/k_2$, $p_{18} = k_9/k_2$, $p_{19} = k_{10}/k_2$, $p_{20} = k_{11} Y_{01}/k_2$, $p_{21} = k_{12}/k_2$, $p_{22} = k_{13}/k_2$, $p_{23} = k_{14} Y_{01}/k_2$, $p_{24} = k_{15}/k_2$, $p_{25} = k_{16}/k_2$. Here $t$ is the dimensional time; $k_0$ is the rate constant of the reverse reaction between MutS2 and a mismatch; $Y_{01}$ is the basal level of the MutS2 protein in the cell in the absence of MMR-inducing lesions; and $y_9$ is the dissociation rate constant of the LexA monomer from the uvrD gene operator.

The kinetic rates estimated by the fitting procedure are presented in Table B.1. Dissociation rate constant $y_9$ is assumed to be equal to the average value of the LexA dissociation rate from the SOS-box (Belov et al., 2009; Mohana-Borges et al., 2000). The value of the Hill coefficient $h_0$ is defined from the data on the binding cooperativity of the LexA repressor and uvrD regulatory region. As there is the only region of LexA binding to the uvrD operator (Smith and Walker, 1998), $h_0$ equals to 2 according to Aksenov et al. (1997). The value of the parameter $k_1$, was set equal to $a$ in the model of SOS network.

Appendix C. Parameters of UV mutagenesis model

See Table C.1.