



OPEN

Simplified modeling of *E. coli* mortality after genome damage induced by UV-C light exposure

Jaime Sánchez-Navarrete¹, Nancy Jannet Ruiz-Pérez¹, Armando Guerra-Trejo² & Julia Dolores Toscano-Garibay¹✉

UV light is a group of high-energy waves from the electromagnetic spectrum. There are three types of UV radiations: UV-A, -B and -C. UV-C light are the highest in energy, but most are retained by the ozone layer. UV-A and -B reach the earth's surface and cause damage on living organisms, being considered as mutagenic physical agents. Numerous test models are used to study UV mutagenicity; some include special lamps, cell cultures and mathematical modeling. Mercury lamps are affordable and useful sources of UV-C light due to their emission at near the maximum absorption peak of nucleic acids. *E. coli* cultures are widely used because they have DNA-damage and -repairing mechanisms fairly similar to humans. In here we present two simple models that describe UV-C light incidence on a genome matrix, using fundamental quantum-mechanical concepts and considering light as a particle with a discontinuous distribution. To test the accuracy of our equations, stationary phase cultures of several *E. coli* strains were exposed to UV-C light in 30 s-intervals. Surviving CFUs were counted and survival/mortality curves were constructed. These graphs adjusted with high goodness of fit to the regression predictions. Results were also analyzed using three main parameters: quantum yield, specific speed and time of mortality.

The main source of energy of our planet is the Sun. The nuclear fusion occurring within its core follows a fixed proton-proton chain reaction known as the Critchfield's cycle that ultimately results in light and energy dispersed in a wide electromagnetic spectrum ($[p^+ + p^+ \rightarrow H_2 + e^+ + U_e, H_2 + p^+ \rightarrow He_3 + \textit{gamma, visible or UV}]; [He_3 + He_3 \rightarrow He_4 + p^+ + p^+]$)¹.

Ultraviolet light (UV light) is the portion of such spectrum constituted by high-energy wavelengths, classified just below ionizing radiations. There are three main UV light regions: type A (UV-A) (315–400 nm), type-B (UV-B) (280–315 nm) and type-C (UV-C) (200–280 nm)². UV-C light is the highest in energy and the more dangerous to living organisms, nonetheless most of it is retained by the ozone layer (OL) in the stratosphere along with other radiations such as cosmic-, gamma- and X-rays and normally do not reach the surface^{3–6}. Sun-emitted light that hits our planet is made up of visible (62.7%), infrared (31.9%), UV-A (5.1%) and UV-B (0.3%) lights⁷. Biological damages during sun exposure are caused by UV-B, hence is considered the main mutagenic physical agent on earth^{8,9}.

Since the 1970's, NASA has monitored the changes in the OL-thickness observing a gradual but controlled deterioration, particularly at the Antarctica, and an elevation in the frequency of ozone depletion events (miniholes) on both hemispheres^{10–13}. This phenomenon is hindering OL protective capacity against radiations and consequently increasing the probability of high-energy UV light exposure, which might be an important factor contributing to the occurrence of cancer-associated mutations in humans (photocarcinogenesis)^{8,9}.

Multiple experimental designs are available to describe and predict the biological effects of UV light. Some of the simpler test systems include the use of artificial light sources inciding on cell cultures and the mathematical modeling of the survival/mortality kinetics.

The most affordable and widely distributed artificial UV light sources are the lamps of low or medium pressure made up of quartz tubes with tungsten electrodes and an atmosphere of mercury/argon. The electric arc produced through such steam, generates light wavelength of 253.7 nm for low-pressure and 240–270 nm for

¹Laboratorio de Investigación en Toxicología, Hospital Juárez de México, Av. Instituto Politécnico Nacional #5160 Col. Magdalena de Las Salinas, Ciudad de México, Mexico C.P. 07760, Mexico. ²Departamento de Biofísica, Escuela Nacional de Ciencias Biológicas, Instituto Politécnico Nacional, Prolongación de Carpio Y Plan de Ayala S/N, Col. Santo Tomás, Ciudad de México, Mexico C.P. 11340, Mexico. ✉email: julia.toscano@gmail.com

medium-pressure lamps, which are close to the maximum absorption peak of nucleic acids (260 nm)¹⁴. This wavelength also coincides with the maximum growth inactivation peak in *E. coli* (268 nm) that makes mercury lamps excellent germicides and also a useful tool for the study of UV-provoked biological damage in this bacteria¹⁵.

Specifically, genomic effects of UV-C light exposure occur through different mechanisms, the most important are by transference of charge through endogenous chromophores to DNA modifying molecules¹⁶ or by direct absorption of photons by nitrogenous bases¹⁷. The main DNA lesions by direct absorption are those where photons are strongly acquired by adjacent pyrimidines, which eventually form dimers. There are two main types of dimerization products: 5–6-*cis*-syn-cyclobutane (CPD) and cross-adducts of pyrimidine-6–4 pyrimidones (6-4PP)^{18–20}. Pyrimidic dimerization halts DNA replication by interfering with polymerases normal gliding and consequently cell reproduction stops (a process called *inactivation*). In *E. coli*, the formation of thymine-thymine, thymine-cytosine, and/or cytosine-cytosine dimers have a frequency of 59%, 34% and 7%, respectively. Nonetheless, there are several natural mechanisms of dimer-repairment that lead to reactivation of *E. coli* cells, from those that restore DNA under visible light (photoreactivation) to the prone to error SOS system. The nucleotide excision repair mechanism (NER), also referred as dark repair of the SOS response, is a pathway predominantly repairing dsDNA damaged by UV-emissions and it is driven by the UvrABC endonuclease multienzyme complex. NER starts when a complex of two UvrA and one UvrB subunits (A2B) scans DNA in search of injured sites. Then, the UvrA dimer is released and UvrB recruits UvrC that in turn hydrolyses specific phosphodiester bonds around the lesion. UvrD helicase then sets loose a 12–13 nt fragment leaving a gap that is finally filled by the DNA polymerase Pol I.

Although this mechanism has been extensively described in bacteria, the process is universal. In humans, NER is conducted by a much larger number of proteins that essentially perform the same steps as UvrABC in *E. coli*, but the difference relies only in the length of the fragment released from the DNA filament (27–29 nt)^{21–23}. Hence, *E. coli* cultures under a luminous stimulus are a good test system for the determination of UV light-induced DNA damage, for the observation of cellular repairing capacities and consequently for the description of survival and/or mortality behaviors of cells.

Finally, several mathematical models have been developed that addressed different aspects to describe the dynamics of the aforementioned phenomenon. Some authors consider a wave-like conduct and others a particle nature of UV-C light, however, there are notable differences between each behaviour of light and the damage caused to the cell interior. For instance, propagation of UV light through a cell matrix could be considered as continuous when is thought as a wave, while it should be reckon that DNA bases would discontinuously absorb impacting photons when UV light is visualized as a corpuscular flow.

On the other hand, models such as the Chick-Watson model, consider a single impact stoichiometry for UV light incident on DNA, that is to say, one photon is sufficient to carry out a photochemical reaction²⁴. However, these models do not fully explain the entire cell inactivation processes²³. Instead, Adair-Aguiar et al. had consider a multiple impact kinetics, where is necessary that two or more photons hit on the same site of a thymine molecule in order to cause a dimerization reaction²⁵. *Jagger and Harm*, on the other hand ponder that a number of photons in the order of 10^4 is required to generate thymine dimers as result of a photoreaction²⁶.

Most of the mathematical models are complex descriptions that assume indistinctly a wave/corpuscular behavior of UV light and use multiple hard-to-measure parameters, making them impractical for further applications.

The aim of this work was to develop two simple models to describe and predict the kinetics of viability and mortality of genome damage induced by UV-C light on *E. coli*, considering both: a multiple impact and a corpuscular behavior of light. We used a set of five strains that were differentially mutated as the experimental system; taking advantage of its fast-growing capability, its well-known mutation and DNA repairment mechanisms and the fact that its genomic structure is described in detail²⁷. Our mathematical models are of deductive and mechanistic nature; both are based on fundamental quantum concepts and the results were analyzed through three main variables: quantum mortality yield, minimum time of mortality and specific speed of mortality.

Materials and methods

Physical and mathematical modeling. *General assumptions.* Models were constructed using differential and integral calculus, mean value theorem, regression method by least squares and determining fiducial limits. We made the following assumptions:

1. According to quantum theory, UV light has a corpuscular behavior.
2. Dissemination of the total energy of the incident light is discontinuous in a genomic matrix.
3. The energy of two or more photons is absorbed by one base. The energy of one photon could not be enough, but two or more would break/form a molecular bond (multiple impact).
4. The physical model was based on the Planck's equation, where the energy of each photon is conversely proportional to the wavelength of the incident light ($E = n_0hc/\lambda$). Energy is expressed in einsteins, and n_0 corresponds to the Avogadro's constant.
5. Pyrimidine dimers are the major and most important photochemical products on nucleic acids. Purine bases are less sensitive to UV-C light^{12,13,23,28}.
6. Both models were based on the relationship between a specific speed of mortality (SSM) and the energy absorbed in a particular exposure time.
7. We considered a hypothesis, named the *proportionality hypothesis*, where the number of base pairs of a genome is proportional to the lethal impact number (*LIN*).

<i>E. coli</i> Strain	Genotype	Phenotype	Reference
HfrH180	(K12 <i>hfr+</i>)	Complete repairing systems	Lavoie and Mathieu ²⁹ , Trudel et al. ³⁰
W3110	<i>F⁻ λ⁻ rph1, IN(rrnD, rrnE)</i>	Slow pyrimidine de novo synthetic pathway	Bachmann ³¹
ATCC25922	Clinical isolate	Complete repairing systems	Bergey's manual ³²
<i>polA</i> ⁺	<i>W3110 thy⁻ thi⁻ PolA1⁺</i>	Expression of <i>polA</i> polymerase	Espinosa-Aguirre et al. ³³
<i>polA</i> ⁻	<i>P3478 thy⁻ thi⁻ PolA⁻</i>	Lack of <i>polA</i> polymerase	Espinosa-Aguirre et al. ³³
DH5athy ⁻	<i>recA⁻ thy⁻</i>	Lack of <i>recA</i>	Manrique-Suárez et al. ³⁴

Table 1. *E. Coli* strains with genotypes.

We used several variables for the mathematical modeling including the effective impact section (EIS), molecular excitation time, frequency of impact, lethal impact number (LIN), number of bases by unit of genomic volume (N^*), lethal radiation dosage at 50% (LRD_{50}), absolute death state, minimum mortality time (MMT), specific speed of mortality (SSM), and cell quantum mortality yield (cQMY). The genomic volume for each *E. coli* strain was also calculated using LIN and cQMY. The fitness of the developed equations was challenged on the laboratory.

Experimental procedures. Table 1 enlists the strains of *E. coli* used for the biological experiments. Each strain was selected with specific mutations affecting repairing enzymes.

Strain DH5athy⁻ was derived from DH5α mutagenized with methyl-*N'*-nitro-*N*-nitrosoguanidine (MNNG, Sigma chemical CO, St Louis MO, USA) according to Manrique-Suarez et al.

Treatment of *E. coli* strains. Each strain was grown during 18 h in 50 ml of nutrient broth No. 2 (Cat. No. CM0067, OXOID) until density reached a turbidity equivalent to a 0.5 McFarland standard or $\approx 10^8$ cells. Then, 3 ml of each culture were directly poured on dry Petri dishes (area 75.42 cm²) and were exposed to UV-C light using a low-pressure mercury lamp (TUV, input voltage 15 W and output 3 W, Cat. No. G15T8, PHILLIPS, Holland) with 253.7 nm in wavelength that was located at a fixed distance of 33 cm, hence a fluence of 2.75 mJ/s cm². Lamp was pre-heated for 10 min before use. Ten dishes of each strain culture were exposed for different times, in 30 s-intervals, up to a maximum exposure lapse of 300 s (fluence: 83–830 mJ/cm², dosage range: 0.627–6.27 J/Petri dish). Subsequently, 500 μl of the content of each dish were collected, serially diluted on saline solution (five dilutions on a decimal scale). Final dilution was mixed with soft agar and spread over nutrient agar plates (Cat.No. OX-CM0003B, OXOID) before incubation at 37 °C during 24 h. Control were obtained as experimental mock-ups. Briefly, 3 ml of the initial stationary culture were poured on dry plates, recovered, serially diluted and cultured under the same conditions as the experimental samples. The number of colony forming units per milliliter (CFU/ml) was determined with a colony counter (Accu-Lite Colony Counter Model 133–8,002, FISHER) and accounting the dilution grade. The number of CFUs from unexposed cultures was considered as 100% of viability. Three experiments were executed by triplicate for each exposure time.

Statistical analysis. Data on viable counts were processed using a linear regression analysis and calculating the correlation and determination coefficients, fiducial limits, confidence intervals, and finally subjected to a t-Student test by one mean. We report mean values of data from each cultured strain.

Results

Construction of the models. *Physical model.* Our physical model assumes that the occurrence of photoreactions depend on the number of photons impacting each nitrogenous base, and hence, on the energetic content of the radiation^{25,28,35}. From the quantum perspective is expected, first, that an “activation” of the genome occurs when hit by a luminous radiation, meaning that part of the incident energy has been absorbed but that is still not enough to produce bond breakage. Second, that the energy of the activated molecule increases as the wavelength diminishes; and third, that an activated molecule absorbs a fixed number of quanta to become an excited molecule where energy is suffice to provoke the dislocation of electrons (Fig. 1). If each excited molecule then results in a product molecule, the yield of such reaction will have values under one because it is a multiquantum process, therefore, the absorbed light influences the rate of cell death in a wavelength-dependent manner^{36,37}. According to this, for any genomic matrix undergoing a photochemical reaction, each molecule accepts luminous energy independently instead of being evenly distributed among the whole bases.

We defined mortality as the condition were a cell is non-reproductive and has no metabolic activity and the cell quantum mortality yield, or cQMY (Φ_μ), as the capacity of quanta with a particular wavelength to cause chemical changes in sensitive bases. In a condition where the stimulus is continuous, each dead cell is the result of the sum of the alterations in the bases (or number of pyrimidine dimers), so cQMY can be expressed in terms of absorbed Einsteins as³⁷:

$$\Phi_\mu = \frac{\text{Number of dead cells (N) on a given time (t)}}{\text{Number of absorbed einsteins on the same time}}$$

According to this, the rate of cell death or specific speed of mortality (SSM or $k = dN/dt$) would be proportional to the activation energy or radiation dosage:

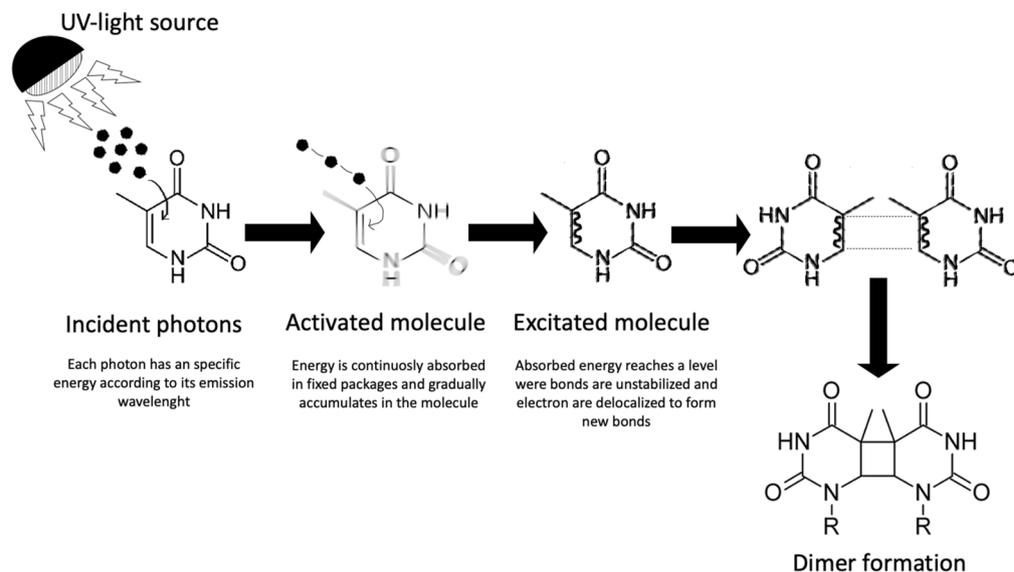


Figure 1. Schematics of basic concepts for the physical model.

$$k \propto \frac{\Delta E}{t} \text{ or } \frac{dN}{dt} \propto \frac{\Delta E}{t}$$

When these expressions are multiplied by a proportionality constant and rearranged, a first order and first grade differential equation is obtained:

$$\frac{dN}{dt} - \Phi_{\mu} (\Delta E/t) = 0 \text{ or } \Phi_{\mu} = \frac{dN}{dt} \frac{t}{\Delta E} \quad (1)$$

also:

$$\frac{dN}{dt} = \Phi_{\mu} (\Delta E/t) \quad (2)$$

cQMY is actually the proportionality factor defined as the efficiency by which a luminous stimulus changes the cellular mortality. $\frac{dN}{dt}$ corresponds to SSM, ΔE is the activation energy (also E_a), t is the exposure time and N is the number of dead cells. In other words, cQMY corresponds to the number of dead cells due to the formation of dimers in a time unit divided by the number of absorbed Einsteins.

Integrating Eq. (2) and assuming that Φ_{μ} is an unchanging and cell-specific value, we have the following mortality expression:

$$N = \Phi_{\mu} E_a \ln(t) \quad (3)$$

Considering that after UV light exposure, the number of initial cells (N_0) is given by the sum of dead plus surviving cells (N_1), then $N_0 = N + N_1$ and survival is defined by the equation:

$$N_1 = -\Phi_{\mu} E_a \ln(t) + N_0 \quad (4)$$

where $N_1 < N_0$ is always true. From here, E_a can be expressed in terms of wavelength (λ) accordingly with assumption 4, so:

$$N_1 = -\frac{\Phi_{\mu} n_0 h c}{\lambda} \ln(t) + N_0 \quad (4a)$$

The latter equation shows that specific speed of mortality by UV-C light is a converse function of the wavelength and is directly related to the total activation energy.

Besides, depending on cQMY, the time for cell inactivation depends on the nature of the sensitive chemical groups on the genome, its base distribution, sequence of sensitive sites and the functionality of cellular repairing mechanisms^{12,25,38}.

Noticeably, cQMY has a linear relationship according to Eq. (4a), through measurable variables (wavelength and mortality). Hence, it is possible to calculate cQMY by counting dead cells after in at least two exposure times. Another way to obtain cQMY is to construct a graphic of N versus $\ln(t)$ where the slope corresponds to:

$$k = \text{slope} = \Phi_{\mu} E_a = \Phi_{\mu} n_0 h c / \lambda \quad (3a)$$

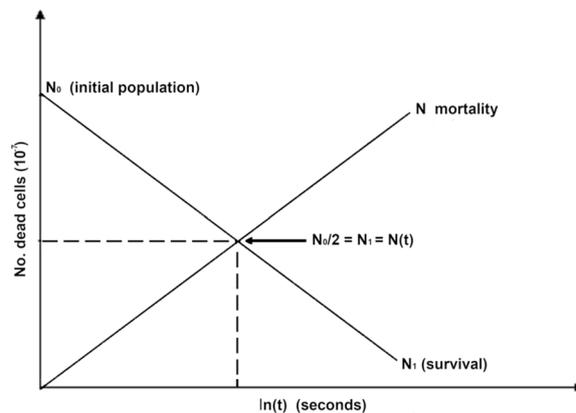


Figure 2. Viability and mortality curves.

On the other hand, when N_1 is charted versus $\ln(t)$, the expected curve is a straight line passing through the ordinate-axis and moving up or down by changes in the negative slope. Such intersect corresponds to N_0 . The slope of this survival curve is associated to the target size, this is to say, to the size of the genome at a constant repairing time. In the case where values of $cQMY$ are the same for two microorganisms, its ratio to the relative genome sizes could be calculated from each slope, thus allowing the differentiation between species (Fig. 2). However, Φ_μ changes with the repairing capacities of each cell and as a function of the degree of damage^{13,39}. It is also worth noting that microorganisms with the same quantity of DNA and similar pyrimidine/purine ratios might have different sensitivities to the lethal effects of radiation depending on its genomic sequence; therefore each species have unique mortality times.

From Fig. 2, it can be observed that the intersection point between the mortality and survival curves corresponds to an equilibrium state where the number of dead is the same as the surviving cells ($N = N_1$), this point is also known as the Lethal Dose of Radiation (LDR_{50}).

Lethal dose of radiation. LDR_{50} is defined as the exposure time needed to kill half a microbial population. Given Eq. (3) and considering that at LDR_{50} $N = N_0/2$, the intersection point using Eq. (4) is resolved as:

$$t = e^{N_0/(2\Phi_\mu E_a)} = e^{N_0/(2k)} \quad (5)$$

In the special case where the population is only two microbial cells irradiated ($N_0 = 2$):

$$t = e^{1/k} \quad (5a)$$

from here we know that $\ln(t) = 1/k$ and this value corresponds to the minimal exposure period required to produce one dead cell, also known as Minimum Mortality Time (MMT, t_u):

$$t_u = 1/k \quad (6)$$

Minimum time of mortality. The energy required to produce the death of one cell is provided by the incidence of UV-C light quanta on each nitrogenous base, instead of the uniform distribution of total energy on a wide genomic area that is contemplated by the wave-like perspective of the diffusion of light over matter (assumption 2).

Cell inactivation occurs in the same time-lapse at which photoreactions are happening. This brief period is such that the sum of the times of each individual photon-base collision is equal to t_u , accounting that the repairing mechanisms are not fast enough to contain the damage.

On the other hand, we can express Eq. (6) on terms of wavelength as:

$$t_u = \frac{\lambda}{\Phi_\mu n_0 hc}$$

the latter equation also shows that the variation on the time for cell death on each species is directly proportional to the wavelength and conversely proportional to $cQMY$. It also can be observed that MMT is defined by the combination of three fundamental constants of physics with $cQMY$, so both variables could be considered as intrinsic biological constants, associated to the nature and composition of a genome. MMT is a crucial concept for the development of our model and deserves special description.

Mortality is frequently a polemic issue because cell death does not happen instantly and it can be described by a series of biochemical events and photochemical reactions. In here we adopted a simplified approach where MMT coincides with the formation of pyrimidine dimers in a context of insufficient repairing systems, the latter is considered as cellular inactivation.

Parameter	Definition	Units
δ	Absorbed fraction	–
l	Mean free path	cm
N^*	Number of total bases per genomic volume	Bases/cm ³
N_j	Number of base pairs	bp
V_G	Genomic volume	cm ³
A_G	Genomic area	cm ²
Z	Effective macroscopic section (EMS)	cm ⁻¹
$f(t_u)$	Lethal impact number symbol (LIN)	–
Φ_u	Cell quantum mortality yield (cQMY)	mol/erg.s
E_a	Activation energy	Einstein, erg/mol
N	Number of dead cells	–
t	Exposure time	s
k	Specific speed of mortality	s ⁻¹
N_0	Number of initial cells	–
N_1	Number of surviving cells	–
RLD_{50}	Radiation Lethal dose at 50%	s, min
MMT	Minimum mortality time	S
t_u	Minimum mortality time symbol	S

Table 2. Definition of the parameters used at physical and mathematical modeling (–) adimensional

Parameter	
$n_0 = 6.0225 \times 10^{23} \text{ mol}^{-1}$	Avogadro's constant
$h = 6.6256 \times 10^{-27} \text{ erg.s}$	Planck's constant
$c = 2.9979 \times 10^{10} \text{ cm/s}$	Speed of light
$\lambda = 2.537 \times 10^{-5} \text{ cm}$	Wavelength of UV-C light lamp
$V_G = 4.3716 \times 10^{-13} \text{ cm}^3$	Mean genome volume
$\Delta X_i = 5.5111 \times 10^{-5} \text{ cm}$	Mean genome thickness
$\sigma = 4.9627 \times 10^{-16} \text{ cm}^2$	Effective impact section
$N_j = 5.2 \text{ Mpb}$	Mean number of base pairs

Table 3. Physical parameters of the models (constants). *See the Appendix 1 derivation of EIS²².

Absolute death state. The absolute death of a microorganism population is such that the number of surviving cells must be $N_1 = 0$, under such premise, according to Eq. (4) the time required to reach this state (T) could be expressed as:

$$T = e^{\left(\frac{N_0}{k}\right)} \quad (6a)$$

Mathematical model. Analytic approach. The physical model describes the survival and mortality kinetics for entire cellular populations. We then aim to mathematically derive equations from molecular events as the initial point and deducing the behavior of individual cells and populations.

In here, the total sum of accumulated photon-base collisions determining the death of a single cell is called *unitary mortality process*. The photodynamics of this process obliges to define the properties of cell populations in function of discrete variables with fixed values contained on discrete parameters as listed on Tables 2 and 3.

Mean free path. In here, we used an statistical mechanical concept known as free path⁴⁰ to define the trajectory that a photon follows from the moment of its impact on the genome until its absorption by a nitrogenous base. We assume that during its displacement, a photon moves at the speed of light (c) in a straight line and the mean distance of these motions is called mean free path (MFP). For photons passing without impacting any molecule, MFP would be infinite. For molecules with a big area without internal movement of photons, MFP would be zero. Actually, every nitrogenous base has a finite extension, then MFP is related to the size of the target, to the genome density and to the speed of impact. Since the photon is considered as a particle with rapid movement, it interacts with the target molecule during a short interval and the impact section depends on the speed (or energy).

To calculate the MFP in a simple manner, an area around the target could be constructed such that an absorption occurs if the trajectory of an impinging photon passes through it. This area is known as Effective Impact Section (EIS) and is represented by σ . A photon flux in a time t could traverse a volumen with a σ of ct and photons would collide with all the molecules with centers included in such volumen. For a continuous flow, the first photon covers a distance $c\Delta t_1$ finishing at the moment of its impact, a second photon travels $c\Delta t_2$, a third $c\Delta t_3$, and so on, in such way that the total distance is:

$$c\Delta t_1 + c\Delta t_2 + c\Delta t_3 + \dots + c\Delta t_i$$

$$\Delta x = c \sum_i \Delta t_i$$

where each time interval has a unique value.

If there are N^* bases in a genomic volumen (V_G), the number of collisions at the time t is the number of remaining available bases ($2N_j$) in the cross-section σ . This is the probability of occurrence of an impact, a.k.a. absorbed fraction (δ):

$$\delta = \sigma N^* c \sum_i \Delta t_i = \sigma N^* \Delta x_i$$

The mean free path (\bar{l}) is the average distance between impacts and then Δx_i is the total distance at the total time invested on photonic impacts. The latter expression is equivalent to the genome thickness divided by the number of impacts occurring in the lapse $\sum \Delta t_i$:

$$\bar{l} = \frac{c \sum_i \Delta t_i}{\sigma N^* c \sum_i \Delta t_i} = \frac{1}{\sigma N^*} \quad (7)$$

In order to find the average impact time, we considered that the time between absorption events is:

$$\Delta t_i = \frac{1}{c} = \frac{1}{N^* \sigma c} \quad (8)$$

The latter equation represents the excitation time^{37,41}, which is an even briefer lapse than MMT. Thus the impact frequency or number of impacts in a time unit results from the reciprocal of Eq. (8):

$$f = N^* \sigma c = \frac{1}{\Delta t_i} \quad (9)$$

To obtain these equations we supposed that photons reach immovable targets. On the other hand, the collision process for a photon flux could be considered as elastic collisions because several particles moving at the same speed with different paths reaching one base would made the times of consecutive hits of single photons on multiple adjacent bases.

The term ($N^* \sigma$) is known as effective macroscopic section (EMS) and is represented by symbol Z . Considering the relationship between the time a photon requires to pass through the entire genome and the mean absorption time we have:

$$\delta = \frac{\Delta t}{\Delta t_i} = \frac{\Delta x/c}{1/N^* \sigma c} = N^* \sigma \Delta x \quad (10)$$

Estimation of the molecular excitation time.

Considering that EIS is $4.9627 * 10^{-16} \text{ cm}^2$ and substituting on Eq. (8), the average time among impacts is then:

$$1/N^* \sigma c = 1 / \left(2.3791 * 10^{19} \frac{\text{bases}}{\text{cm}^3} \right) (4.9627 * 10^{-16} \text{ cm}^2) (2.9979 * 10^{10} \text{ cm/sec})$$

$$1/N^* \sigma c = 2.8252 * 10^{-15} \text{ sec}$$

Estimation of the absorbed fraction. The number of available molecules in a genome could be defined in terms of its area (A_G) and thickness (Δx) as:

$$N^* A_G \Delta x$$

If each base has a defined EIS, the total impact area is: $\sigma (N^* A_G \Delta x)$, henceforth the probability of photon impacting the genome is a ratio of the total area or: $\sigma (N^* A_G \Delta x) / A_G$. To determine such ratio, we matched the equation to the absorbed fraction:

$$\delta = \sigma (N^* \Delta x) = N^* \sigma c \Delta t \quad (10a)$$

Since $\Delta x = c \Delta t$ and considering that $N^* = 2N_j / V_G$, then substituting on Eq. (10) we have:

$$\delta = \frac{2N_j \sigma A_G \Delta x}{V_G A_G} = \frac{2N_j \sigma \Delta x}{V_G}$$

Also, if $A_G \Delta x$ is the volume occupied by the genome, then $A_G \Delta x / V_G = 1$ and the absorbed fraction is:

$$\delta = \frac{2N_j \sigma}{A_G} \quad (10b)$$

An equivalent expression is:

$$\delta = \frac{\overline{\Delta x}}{l} \quad (10c)$$

Taking a mean genome size of 5.2 millions of base pairs (Mbp)⁴² for *E. coli*, a genome volume of $4.3716 \times 10^{-13} \text{ cm}^3$ and a thickness of $\Delta x = 5.5111 \times 10^{-5} \text{ cm}$ ⁴³⁻⁴⁵, then substituting in Eq. (10):

$$\delta = \left(2.3790 \times 10^{19} \frac{\text{bases}}{\text{cm}^3} \right) (4.9627 \times 10^{-16} \text{ cm}^2) (5.5111 \times 10^{-5} \text{ cm})$$

$$\delta = 0.6506 \text{ bases; and } Z = N^* \sigma = 1.1806 \times 10^4 \text{ cm}^{-1}$$

Lethal impact number. From Eq. (9), the product of the impact frequency times MMT is equal to the number of impacts at the time $t = t_u$, also denominated lethal impact number (LIN) or the number of photonic impacts received by a particular molecule in the time needed to reach mortality, and it is proportional to the number of bases in the genome (see next section).

Then:

$$f(t_u) = \frac{N^* \sigma c}{k} \quad (12)$$

And accordingly,

$$k = \frac{N^* \sigma c}{f(t_u)} \quad (13)$$

where it can be seen that k is directly proportional to the number of bases in the genome volume, and also, to the EIS.

Generalization of the Unitarian equation system for mortality and impact.

The frequency of impacts occurring in the time interval necessary for the death of one cell, gives an adequate argument to know the mortality kinetics, since the change rate in a population represents a continuous function in the close interval $t_a \leq t \leq t_b$ and is derivable in the open interval $t_a < t < t_b$; hence, for a value of $t = (t_u)$, enclosed between $t_a < t < t_b$, is verifiable that:

$$\frac{N_b - N_a}{t_b - t_a} = \frac{\Delta N}{\Delta t} = \frac{dN(t)}{dt} \quad (14)$$

then the interval between t_a and t_b constitutes a differential element of the mortality curve. Knowing that $f(t_u)$ happens in $t = (t_u)$, then:

$$f(t_u) = \frac{N^* \sigma c t_u}{t} \quad (15)$$

rearranging the equation for $t_u = t_b - t_a$ and considering Eq. (14) we get the following expression:

$$N_b = N_a + (t_b - t_a) \frac{dN(t)}{dt} = N_a + (t_u) \frac{dN(t)}{dt} \quad (16)$$

Substituting in Eq. (16), from $N_a = 0$ until $N = N_b = 1$ and for the death of a single cell as a condition of Eq. (15) with $t_a = 1$, we have:

$$\frac{dN(t)}{dt} = \frac{N^* \sigma c}{f(t_u) t_a} = k$$

This last expression is the definition of the SSM. Solving the differential equation we get the following equality for the mortality:

$$N = \frac{N^* \sigma c}{f(t_u)} \ln(t) \quad (17)$$

On the other hand, the number of surviving microorganisms is:

$$N_1 = -\frac{N^* \sigma c}{f(t_u)} \ln(t) + N_0 \quad (18)$$

Equations (17) and (18) are analogue expressions to Eqs. (3) and (4), of our physical model. From Eq. (16) we can deduce that the MMT is given by:

$$\frac{(1 - 0)f(t_u)(1)}{N^* \sigma c} = (t_b - t_a) = t_u$$

and then:

$$t_u = \frac{f(t_u)}{N^* \sigma c} = f(t_u) \Delta \bar{t}_i \quad (19)$$

The latter equation indicates that the sum of the individual collisions corresponds to the MMT or to the number of impacted bases in the same time.

Under this argument and from Eq. (16), again is possible to deduce the unitary equation of mortality and impact:

$$N_b = N_a + (t_b - t_a) \frac{N^* \sigma c}{f(t_u) t_a}$$

when $N = N_b = 1$, because $N_a = 0$, then:

$$N = \frac{N^* \sigma c}{f(t_u) t_a} = 1$$

Replacing t_u for $1/k$ and $t_a = 1$

$$N = \frac{N^* \sigma c}{f(t_u) k} = 1 \quad (20)$$

This last equation is a particular case of Eq. (17) which shows, along with Eq. (18), that SSM is directly associated to the genome size, EIS and c but conversely to the number of impacts received by each base in the MMT. The total number of bases is constant and characteristic of each species of microorganism (see Appendix 2 for derivation).

Predictions and theoretical results. The results obtained by the developed equations in the mathematical model allow us to estimate, under the proportionality hypothesis of assumption 7, theoretical values for the studied parameters (SSM, MMT and cQMY). Data on the number of base pairs associated to each species and genus were obtained from the literature; these are related to the dimension of their nucleoid and to the bacterial size. Other values are related to the corpuscular nature of light.

Given the Eq. (17) for cellular mortality, where the value for the SSM is given by Eq. (13), knowing the implicit parameters and considering that $\bar{N}_j \propto f(t_u)$ or $\bar{N}_j \approx f(t_u)$, we have:

$$\bar{k} \approx \frac{2\sigma c}{\bar{V}_G} \quad (13a)$$

where \bar{V}_G is the average of genome volume⁴⁴⁻⁴⁷. Substituting the known values for *E. coli*, we get:

$$\bar{k} \approx \frac{2\sigma c}{\bar{V}_G} \approx \frac{2(4.9627 * 10^{-16} \text{ cm}^2)(2.9979 * 10^{10} \text{ cm/s})}{4.3716 * 10^{-13} \text{ cm}^3}$$

$$\bar{k} \approx 6.8065 * 10^7 \text{ s}^{-1}$$

If MMT is the reciprocal of k , then we have a mean value of:

$$\bar{t}_u = 1/\bar{k} \approx 1.4691 * 10^{-8} \text{ s}$$

In the molecular context, the energy distribution in the excited bases or “hot bases” is widely enriched as compared to the molecules in basal states; this is, electronically, vibrationally and rotationally. According to our estimations, the excited molecules will have a half-life of 10^{-8} s^{37,41}, and its energy would be either transferred to the adjacent pyrimidine to form a dimer or released with the consequent recovery of the basal state. This process occurs in the whole genome matrix, and could inactivate the cell if the UV-C light stimuli are constant and the repairing mechanisms are not sufficient.

The mean cQMY was calculated from Eqs. (13a) and (3a):

$$\bar{\Phi}_\mu \approx \frac{2\bar{\sigma}\lambda}{V_G n_0 h} \quad (21)$$

Using the values of the proportionality hypothesis and the result for the SSM, we can calculate the cQMY as:

$$\bar{\Phi}_\mu \approx \frac{2(4.9627 * 10^{-16} \text{ cm}^2)(2.537 * 10^{-5} \text{ cm})}{(4.3716 * 10^{-13} \text{ cm}^3)(6.0225 * 10^{23} \text{ mol}^{-1})(6.6256 * 10^{-27} \text{ erg s})}$$

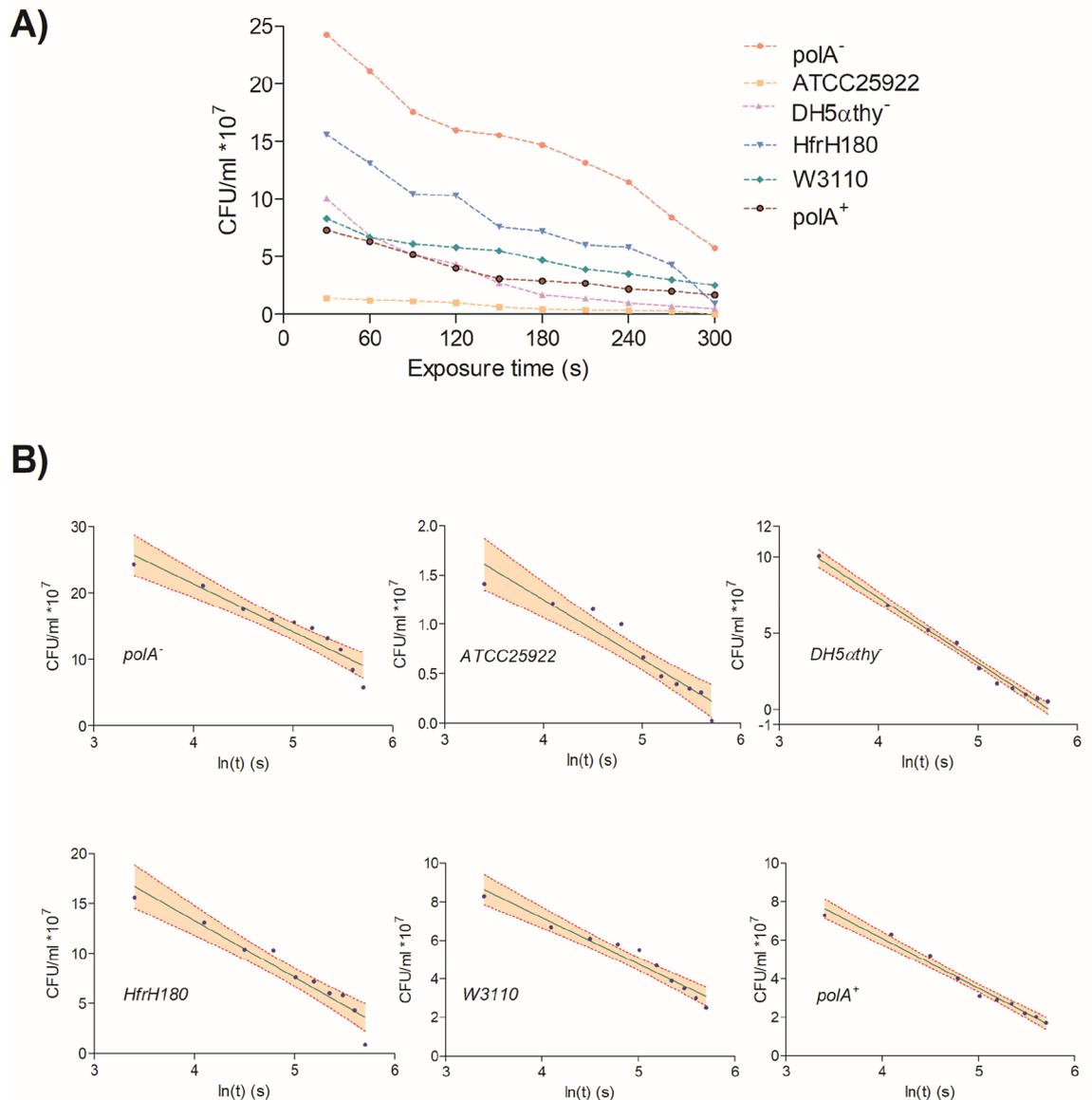


Figure 3. Survival curves of E coli strains. (a) Representation of survival of all used strains. (b) Dispersion plots with fiduciary limits (red lines) and confidence intervals (fawn colored area).

$$\bar{\Phi}_{\mu} \approx 1.4435 * 10^{-5} \text{ mol/erg.s}$$

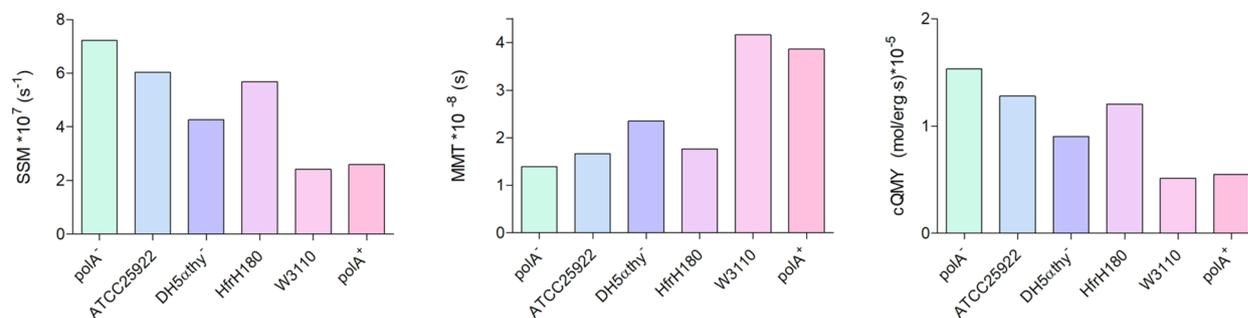
Noticeable, data is on the range of magnitude orders of those previously reported by others^{23,26,48}.

Experimental results. We statistically analyzed the data obtained in the laboratory to graph the natural logarithm of exposure time and the mortality/survival of the exposed population. To simultaneously proof the results obtained through the models, regression coefficients and the intercepts were determined.

The first step was to elaborate a plot using the CFU/ml counted for each strain to decide which was the best model describing the survival of *E. coli* under UV-C light exposure at several times (Fig. 2A). Data were distributed in a logarithmic curve (survival decreases with exposure time), with N_0 and k as constants for the initial number of microorganisms ($t = 0$). We use a regression model for the best fit, to establish the relationship between the variables and to check the theoretical prediction obtained from the mathematical models.

It can be seen from the Fig. 3A that data had a trend towards a linear behavior, with a slope or regression coefficient corresponding to the SSM. After constructing the confidence interval, it can be seen that 90% of the data were included inside the ranges (Fig. 3B).

Then, cQMYs were estimated using the regression coefficients (Fig. 4). It is also shown a complementarity between the variables, for instance, polA⁻ strain (light green bar) had the highest SSM meaning that it dies with a shorter exposure time (lower MMT) and it had a high susceptibility to UV-C light exposure (high cQMY).



Strain	[SSM]*10 ⁷ (k, s ⁻¹)	MMT*10 ⁻⁸ (t _u ,s)	cQMY*10 ⁻⁵ (Φ _μ , mol/erg.s)	R	R ²
polA ⁻	7.2239	1.3843	1.5320	0.9558	0.9135
ATCC25922	6.0322	1.6578	1.2793	0.9526	0.9074
DH5athy ⁻	4.2609	2.3469	0.9036	0.9949	0.9898
HfrH180	5.6813	1.7602	1.2049	0.9627	0.9267
W3110	2.4085	4.1519	0.5108	0.9722	0.9451
polA ⁺	2.5937	3.8555	0.5501	0.9901	0.9803

Figure 4. Graphical representation and calculated values of the measured parameters. SSM and R are presented in absolute values. R: correlation coefficient, R²: determination coefficient.

Strain	Wavelength (λ, nm)	Assumption 7; Eq. (22)	D _p	Reference
<i>Experimental results</i>				
polA ⁻	253.7	1.0613	65,274	-
ATCC25922	253.7	0.8862	78,168	-
DH5athy ⁻	253.7	0.6260	110,664	-
HfrH180	253.7	0.8346	82,994	-
W3110	253.7	0.3538	195,771	-
polA ⁺	253.7	0.3810	181,798	-
<i>Referenced values</i>				
K12 (1981)	260	1.0578	65,488	Jagger ²⁶
-	254	1.4076	49,213	Harm ⁴⁸
IFO 3,301	254	2.7709	25,000	Oguma ¹²
K12 (wild type)	254	4.4335	15,625	Howard-Flanders ³⁹

Table 4. Assumption (7) for Eq. (22) in experimental and other *E. coli* strains.

To compare the observed means of SSM against the estimated values, a t-Student test was performed considering 0.97 of confidence, 0.03 as significance level. The results showed $p \geq 0.045$ (data not shown). Hence, values had not a statistically significant difference.

Proportionality hypothesis. The quantity and distribution of the nitrogenous bases and the genomic volume are intrinsic features of each genus and species and these parameters are determinant when calculating the SSM, because it changes inversely with V_G . In here, we hypothesized that the number of photons impacting a genomic matrix is proportional to the number of base pairs and that the coefficient of these two values must be nearly equal to one^{23,28}.

Then, using Eqs. (2a), (13) and (13a), we derived an equation to test the validity of the proportionality hypothesis:

$$\frac{N_j}{f(t_u)} = \frac{\Phi_\mu n_0 h V_G}{2\sigma \lambda} \quad (22)$$

In some of our experimental observations this coefficient resulted in values close to one, except for W3110 and polA⁺ strains. To compare other proportionality coefficients, Eq. (22) was applied to data previously reported in the literature (Table 4).

Nonetheless, Eq. (22) does not describe the number of photons necessary for the formation of pyrimidine dimers; to that end, we obtained the following expressions:

$$\Phi_\mu = D^{\wedge} p / \gamma_{\text{abs}} \quad (23)$$

$$D_p = 1/\Phi_\mu = \gamma_{abs}/D^p \quad (24)$$

where D_p is the number of formed pyrimidine dimers, D_p is the number of photons necessary for the formation of a single dimer and γ_{abs} is the number of absorbed photons per cell. The reciprocal values of cQMY are shown in Table 4, confirming postulate 7 and the multi-impact nature of UV-C light propagation on DNA.

Discussion

According to our equation system, the determination of cQMY, SSM, MMT and D_p are crucial when describing UV-C light incidence on a genome. These four parameters are sensible indicators of the degree of damage of an exposed genomic matrix.

Beggs²³ reported that Φ_μ might be ranging from 10^{-6} to 10^{-5} and in Table 4 we enlisted other cQMYs^{12,26,39,48}. Noticeably, the yields calculated experimentally are on the same order of magnitude, showing a high level of fitness in our equations. The proportionality coefficient of Eq. (22), varies from 1.03 to 4.00 (Table 4), we attribute these discrepancies to the diversity of methodologies used by each author.

The mathematical model predicted a value of SSM of $6.8 \times 10^7 \text{ s}^{-1}$ and the experimental results show a range of $2-7 \times 10^7 \text{ s}^{-1}$, without a significant difference between both values ($p=0.045$). In regard to MMT, the experimental results are within the range of the predicted value ($1.4 \times 10^{-8} \text{ s}$) (Fig. 3) and coincide with other mortality times previously reported for the duration of the excitation state. We determined that photons could lead a base into its excitation state at a rather fast time ($2.8252 \times 10^{-15} \text{ s}$, $\lambda = 253.7 \text{ nm}$), which is also in accordance with other reports³⁵.

Another useful data in the simulation of the process is the absorbed fraction (δ). In our experiments, $\delta = 0.65$, implying that approximately 35% of an *E. coli* genome is not hit.

Variations on this value are due to the genomic area, base-distribution and sequence configuration of each strain.

We used 5.2 Mbp as the mean genomic size according to the GENOME database for *E. coli*⁴², although other authors use 4.6 Mbp as the mean size (4'639,221 bases)^{13,49}, and even in the literature, values ranged from 4.6 to 5.5 Mbp; nonetheless, by substituting these alternative values the results are approximate to those previously reported.

In regard to the physical model, although single impact developments are used in other biological models of UV-inactivation, they do not completely explain the dynamics of the process²³, nor does the models with a continuous light diffusion approach (undulatory behavior). We considered that the corpuscular approach is more adequate since a quantum model results useful to explain why several photons are required to cause a pyrimidic dimerization. Besides, the repairing mechanisms play an important role because a cell with a low repairing rate could not overcome dimer formation and it might be unable to maintain its integrity. cQMY is directly associated with the wavelength of the incident light and to the genome size, which make it an appropriate measure of the effectiveness of a photochemical reaction (Eq. 21) and could be considered as a biological constant, analogous to the quantum yield. Our results are mutually compatible with the unitary equations of mortality and impact and in agreement with the experimental values reported by Kumiko et al. of 4×10^{-5} for a genome size of 4.6 Mbp and the 6.4×10^{-5} reported by Howard-Flanders et al.^{26,39,48}.

Sensitivity. We showed that SSM changes directly with cQMY and conversely with MMT. Nonetheless, the variability of data between each *E. coli* strain is related to the characteristics of each genome and/or its repairing mechanisms. Pyrimidine dimers are eliminated by the nucleases coded in *uvrABC* and *PolA* genes. When the filling of gaps is produced after replication, DNA repairing could be done by recombination with an undamaged parallel strain by means of Rec proteins, particularly by RecA⁵⁰. This protein is also necessary for the expression of the SOS regulon, because it promotes the proteolysis of the repressor (LexA) hence allowing the expression of *umuCD* genes⁵¹. On the other hand, it is known that deletion-insertion mutants of the DNA helicase II (*uvrD*) are UV-C light sensitive due to the high variability of its allelic genotypes⁵².

In here, the elimination genotypes of each strain produce different values for MMT, cQMY and SSM. The strain HfrH180, which does not have deficiencies on any DNA repairment system^{53,54} was moderately sensitive to UV-C light effects.

PolA⁻ strain presented a higher sensitivity to the UV-C light exposure compared to PolA⁺, probably due to its lack of the polymerase. In fact, the pair of PolA⁻/PolA⁺ strains is widely used on the study of geno-lethal effects of diverse substances on bacterial test panels⁵⁵.

Youngs et al.⁵⁶ showed that a higher sensitivity to UV light was presented by strains with a *recA*⁻ genotype, being the strains *recA*⁻/*uvrB*⁻ the most susceptible. In here, the strain DH5 α thym⁻ was less sensitive to UV-C light than HfrH180 although it does not have mutations on the *uvrABC* system. The ATCC25922 strain had a relatively high sensitivity only below *polA*⁻, which was unexpected, although, it partially could be explained because this strain is intrinsically sensitive to lethal agents. In fact, ATCC25922 is recommended by the CLSI as a reference microorganism to evaluate antimicrobial sensitivity. Finally, *E. coli* W3110, previously reported as UV-sensitive and with a pyrimidine auxotrophy, had low SSM but a still measurable mortality.

Regression, correlation and determination. The regression analysis showed a strong trend towards a straight line on the dispersion diagram [$\ln(t)$ vs N , Fig. 2B] and the regression coefficients were in agreement with the theoretical calculations (SSM, k). The fact that the data behavior was not accurately predicted for a sample of the variables [$(\ln(t_i), N_j)_{ij}$] depends on the experimental variability and the complex biological nature of test system. Let us also recall that we assume that the number of surviving or dead microorganisms is a random variable, which depends on the exposure time, that it has homogeneous variances, independence of N and N_j

values and a normal distribution. The proposed models are adequate to describe the travel of UV-C light in a genomic environment in a more complete and real manner.

Regarding the correlation coefficients, the points on each strain showed a marked association between both variables $[\ln(t), N]$ within the interval $0.9526 \leq R \leq 0.9949$. Hence, there is good evidence that the variable $\ln(t)$ contributes to explaining N and N_1 .

The values for the determination coefficients had a high goodness of fit to the projected curve because the explained fractions are in the interval of $0.9074 \leq R^2 \leq 0.9898$, revealing a cause-effect relationship between the variables.

Conclusions

The quantum depiction proposed in these models was useful to prove that the distribution of the activation energy from UV-C light through the genome is discontinuous. On the other hand, these are multiple-impact mechanistic models that allowed the calculation of mean population values from individual effects at the photon-molecule interaction. Finally, other population kinetics of mortality and survival could be simulated by taking into account the wavelength of the lamp used, the volume and base distribution in the genome of the microorganism in study and by contrasting data with simple viable counts measures from the resulting cultures.

Perspectives

We believe that is necessary to consider the DNA repairing mechanisms to extend the parameters involved, then to improve the mathematical modeling to include eukaryotic cells or tissues. Additionally, antimutagenic capacity of candidate compounds projected to be used as protection from the effects of UV light, can be further studied since the equations describe the mechanisms associated with damage to nucleic acids and allows its quantization. cQMY, SSM and MMT could be used as classification parameters for the mutagenic/antimutagenic potency of virtually any molecule; this is to say, to establish the grade (low, mild, high or very high) on which a substance will damage/protect a cell from UV light effects.

The system presented in here, could also be extended to other wavelengths by performing the pertinent adjustments for other affected biomolecules (proteins, lipids, carbohydrate, etc.). On the other hand, in eukaryotic cells it could be applied in tumoral tissue cells, in radiotherapy assays and in carcinogenesis animal models. Alternatively, the fitness of models could be challenged by adding molecules with known photonic absorption (molecular actinometers) to improve their accuracy.

Received: 1 October 2019; Accepted: 9 June 2020

Published online: 09 July 2020

References

- Atkinson, R. D. E. & Houtermans, F. G. Zur Frage der Aufbaumöglichkeit der Elemente in Sternen. *Zeitschrift für Physik* **54**, 656–665 (1929).
- Meulemans, C. C. E. The basic principles of UV-disinfection of water. *Ozone Sci. Eng.* **9**, 299–313. <https://doi.org/10.1080/01919518708552146> (1987).
- Perdiz, D. *et al.* Distribution and repair of bipyrimidine photoproducts in solar UV-irradiated mammalian cells. Possible role of Dewar photoproducts in solar mutagenesis. *J Biol Chem* **275**, 26732–26742 (2000).
- Clydesdale, G. J., Dandie, G. W. & Muller, H. K. Ultraviolet light induced injury: immunological and inflammatory effects. *Immunol Cell Biol* **79**, 547–568 (2001).
- de Gruijl, F. R., van Kranen, H. J. & Mullenders, L. H. UV-induced DNA damage, repair, mutations and oncogenic pathways in skin cancer. *J. Photochem. Photobiol. B* **63**, 19–27 (2001).
- van der Leun, J. C. & de Gruijl, F. R. Climate change and skin cancer. *Photochem. Photobiol. Sci.* **1**, 324–326 (2002).
- Svobodova, A., Psotova, J. & Walterova, D. Natural phenolics in the prevention of UV-induced skin damage. A review. *Biomed. Pap. Med. Fac. Univ. Palacky Olomouc Czech Repub.* **147**, 137–145 (2003).
- Kozmin, S. G., Pavlov, Y. I., Kunkel, T. A. & Sage, E. Roles of *Saccharomyces cerevisiae* DNA polymerases Poleta and Polzeta in response to irradiation by simulated sunlight. *Nucleic Acids Res.* **31**, 4541–4552 (2003).
- Gonzales-Pumariega, M., Vernhes-Tamayo, M. & Sanchez-Lamar, A. Ultraviolet radiation: its incidence in the human health. *Theoria* **18**, 69–80 (2009).
- Newman, P. A. & Nash, E. R. NASA ozone watch. https://ozonewatch.gsfc.nasa.gov/facts/miniholes_NH.html (2018).
- Stenke, A. & Grewe, V. Impact of ozone mini-holes on the heterogeneous destruction of stratospheric ozone. *Chemosphere* **50**, 177–190 (2003).
- Oguma, K., Katayama, H. & Ohgaki, S. Photoreactivation of *Escherichia coli* after low-or medium-pressure UV disinfection determined by an endonuclease sensitive site assay. *Appl. Environ. Microbiol.* **68**, 6029–6035 (2002).
- Oguma, K. *et al.* Determination of pyrimidine dimers in *Escherichia coli* and *Cryptosporidium parvum* during UV light inactivation, photoreactivation, and dark repair. *Appl. Environ. Microbiol.* **67**, 4630–4637 (2001).
- O'Brien, W., Hunter, G., Rosson, J., Hulsey, R. & Carns, K. In Disinfecting wastewater for discharge & reuse speciality conference of the water environment federation. 2-11-12-22 (WEF).
- Von Sonntag, C. & Schuchmann, H.-P. UV disinfection of drinking water and by-product formation-some basic considerations. *Aqua J. Water Supply Res. Technol.* **41**, 67–74 (1992).
- Verschooten, L., Claerhout, S., Van Laethem, A., Agostinis, P. & Garmyn, M. New strategies of photoprotection. *Photochem. Photobiol.* **82**, 1016–1023 (2006).
- Rochette, P. J. *et al.* UVA-induced cyclobutane pyrimidine dimers form predominantly at thymine-thymine dipyrimidines and correlate with the mutation spectrum in rodent cells. *Nucleic Acids Res.* **31**, 2786–2794 (2003).
- Douki, T., Court, M., Sauvaigo, S., Odin, F. & Cadet, J. Formation of the main UV-induced thymine dimeric lesions within isolated and cellular DNA as measured by high performance liquid chromatography-tandem mass spectrometry. *J. Biol. Chem.* **275**, 11678–11685 (2000).
- Park, H. *et al.* Crystal structure of a DNA decamer containing a cis-syn thymine dimer. *Proc. Natl. Acad. Sci. USA* **99**, 15965–15970 (2002).
- Wang, X. & Yu, H. The effect of DNA backbone on the triplet mechanism of UV-induced thymine-thymine (6–4) dimer formation. *J. Mol. Model.* **24**, 319 (2018).

21. Rastogi, R. P., Richa Kumar, A., Tyagi, M. B. & Sinha, R. P. Molecular mechanisms of ultraviolet radiation-induced DNA damage and repair. *J Nucleic Acids* **2010**, 592980 (2010).
22. Karschau, J. *et al.* A matter of life or death: modeling DNA damage and repair in bacteria. *Biophys. J* **100**, 814–821 (2011).
23. Beggs, C. B. A quantitative method for evaluating the photoreactivation of ultraviolet damaged microorganisms. *Photochem. Photobiol. Sci.* **1**, 431–437 (2002).
24. Kowalski, W. J., Bahnfleth, W. P., Witham, D. L., Severin, B. F. & Whittam, T. S. Mathematical modeling of ultraviolet germicidal irradiation for air disinfection. *Quant. Microbiol.* **2**, 249–270 (2000).
25. Aguiar, A. *et al.* Modeling UV-damage to *E. coli* bacteria. *IMA preprints*, 1–22 (1996).
26. Jagger, J. (Prentice-Hall, 1967).
27. Kato, T., Rothman, R. H. & Clark, A. J. Analysis of the role of recombination and repair in mutagenesis of *Escherichia coli* by UV irradiation. *Genetics* **87**, 1–18 (1977).
28. Zavilgelsky, G., Gurzadyan, G. & Nikogosyan, D. Pyrimidine dimers, single-strand breaks and crosslinks induced in DNA by powerful laser UV irradiation. *Photobiochem. Photobiophys.* **8**, 175–187 (1984).
29. Lavoie, M. & Mathieu, L. G. Isolation and partial characterization of an *Escherichia coli* mutant resistant to colicin A. *Can. J. Microbiol.* **21**, 1595–1601 (1975).
30. 30Trudel, L., Arriaga-Alba, M. & Lavoie, M. C. Survey of drug and phage resistance and colicin and hemolysin production among coliforms isolated in the Ivory Coast. 905–907 (1984).
31. Bachmann, B. J. Pedigrees of some mutant strains of *Escherichia coli* K-12. *Bacteriol. Revs* **36**, 525–557 (1972).
32. Scheutz, F. & Strockbine, N. A. 1–49 (Wiley, 2015).
33. Espinosa-Aguirre, J. J., Aroumir, C., Meza, M. T., Cienfuegos, E. & Cortinas de Nava, C. Genotoxicity of amebicide and anthelmintic drugs in *Escherichia coli* pol A⁺/pol A. *Mutat. Res.* **188**, 111–120 (1987).
34. Manrique-Suárez, V. *et al.* A thymidineless *Escherichia coli* strain useful for in vivo DNA labeling. *Rev. CENIC Cienc. Biol.* **35**, 67–70 (2004).
35. Nikogosyan, D., Angelov, D. & Oraevsky, A. Determination of parameters of excited states of DNA and RNA bases by laser UV photolysis. *Photochem. Photobiol.* **35**, 627–635 (1982).
36. Joklik, W. K., Willet, H. P., Amos, D. B. & Wilfret, C. M. *Zinsser Microbiology* (Appleton and Lange, Norwalk, 1992).
37. Avery, H. E. *Basic reaction kinetics and mechanisms* (Macmillan International Higher Education, London, 1974).
38. Zimmerman, S. B. & Murphy, L. D. Release of compact nucleoids with characteristic shapes from *Escherichia coli*. *J. Bacteriol.* **183**, 5041–5049 (2001).
39. Howard-Flanders, P. DNA repair. *Ann. Rev. Biochem.* **37**, 175–200. <https://doi.org/10.1146/annurev.bi.37.070168.001135> (1968).
40. Walker, J., Halliday, D. & Resnick, R. (ed J. Walker) 558–560 (Wiley, 2014).
41. Avers, C. J. *Cell biology*. (Van Nostrand, 1976).
42. NIH. *Genome* <https://www.ncbi.nlm.nih.gov/genome>.
43. Hecht, R. M., Taggart, R. T. & Pettijohn, D. E. Size and DNA content of purified *E. coli* nucleoids observed by fluorescence microscopy. *Nature* **253**, 60–62 (1975).
44. Hobot, J. A. *et al.* Shape and fine structure of nucleoids observed on sections of ultrarapidly frozen and cryosubstituted bacteria. *J. Bacteriol.* **162**, 960–971 (1985).
45. Kellenberger, E. Les nucléoïdes de *Escherichia coli* étudiés à l'aide du microscope électronique. *Experientia* **8**, 99–101 (1952).
46. Pettijohn, D. E. *Escherichia coli and salmonella: cellular and molecular biology* (eds F. C. Curtiss R. I. I. Ingraham J. L. Lin E. C. C. Low K. B. Magasanik B. Reznikoff W. S. Riley M. Schaechter M. Neidhardt & H. E. Umbarger) 158–166 (ASM Press, 1996).
47. Valkenburg, J. A. C., Woldringh, C. L., Brakenhoff, G. J., Van Der Voort, H. T. M. & Nanninga, N. Confocal scanning light microscopy of the *Escherichia coli* nucleoid: comparison with phase-contrast and electron microscope images. *J. Bacteriol.* **161**, 478–483 (1985).
48. Harm, W. *Biological effects of ultraviolet radiation* (University Press, Cambridge, 1980).
49. Blattner, F. R. *et al.* The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1462 (1997).
50. Capaldo-Kimball, F. & Barbour, S. D. Involvement of recombination genes in growth and viability of *Escherichia coli* K-12. *J. Bacteriol.* **106**, 204–212 (1971).
51. Miller, R. V. & Kokjohn, T. A. General microbiology of recA: environmental and evolutionary significance. *Ann. Rev. Microbiol.* **44**, 365–394 (1990).
52. Washburn, B. K. & Kushner, S. R. Construction and analysis of deletions in the structural gene (*uvrD*) for DNA helicase II of *Escherichia coli*. *J. Bacteriol.* **178**, 2569–2575 (1991).
53. Sanchez-Navarrete, J., Arriaga-Alba, M., Ruiz-Perez, N. J. & Toscano-Garibay, J. D. Antimutagenic activity of vitamin B1 against damages induced by chemical and physical mutagens in *Salmonella typhimurium* and *Escherichia coli*. *Toxicol. In Vitro* **45**, 202–206 (2017).
54. Rupp, W. D. Early days of dna repair: discovery of nucleotide excision repair and Homology-dependent recombinational repair. 499–505 (2013).
55. Hyman, J., Leifer, Z., Rosenkranz, H. S. & The, E. The *E. coli* Pol A-1 assay: a quantitative procedure for diffusible and non-diffusible chemicals. *Mutat. Res.* **74**, 107–111 (1980).
56. Youngs, D. A. & Smith, K. C. Single-strand breaks in the DNA of the *uvrA* and *uvrB* strains of *Escherichia coli* K-12 after ultraviolet irradiation. *Photochem. Photobiol.* **24**, 533–541 (1976).

Acknowledgements

Authors acknowledge Dra. Myriam Arriaga Alba for her support and multiple intellectual contributions.

Author contributions

J.S.N. performed acquisition of data and mathematical modeling. N.J.R.P. made substantial contributions to conception, experimental design and execution. J.S.N., N.J.R.P. and J.D.T.G. contributed to the analysis and interpretation of data. A.G.T. participated in the critical review of the mathematical modeling and statistical analysis. J.D.T.G. participated in drafting the article and revising it critically for intellectual content and technical writing. All authors gave final approval of the version to be submitted and any revised version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-67838-1>.

Correspondence and requests for materials should be addressed to J.D.T.-G.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020