



OPEN

Fully integrated Monte Carlo simulation for evaluating radiation induced DNA damage and subsequent repair using Geant4-DNA

Dousatsu Sakata^{1✉}, Oleg Belov^{2,3}, Marie-Claude Bordage^{4,5}, Dimitris Emfietzoglou⁶, Susanna Guatelli⁷, Taku Inaniwa¹, Vladimir Ivanchenko^{8,9}, Mathieu Karamitros¹⁰, Ioanna Kyriakou⁶, Nathanael Lampe¹¹, Ivan Petrovic¹², Aleksandra Ristic-Fira¹², Wook-Geun Shin¹³ & Sebastien Incerti¹³

Ionising radiation induced DNA damage and subsequent biological responses to it depend on the radiation's track-structure and its energy loss distribution pattern. To investigate the underlying biological mechanisms involved in such complex system, there is need of predicting biological response by integrated Monte Carlo (MC) simulations across physics, chemistry and biology. Hence, in this work, we have developed an application using the open source Geant4-DNA toolkit to propose a realistic "fully integrated" MC simulation to calculate both early DNA damage and subsequent biological responses with time. We had previously developed an application allowing simulations of radiation induced early DNA damage on a naked cell nucleus model. In the new version presented in this work, we have developed three additional important features: (1) modeling of a realistic cell geometry, (2) inclusion of a biological repair model, (3) refinement of DNA damage parameters for direct damage and indirect damage scoring. The simulation results are validated with experimental data in terms of Single Strand Break (SSB) yields for plasmid and Double Strand Break (DSB) yields for plasmid/human cell. In addition, the yields of indirect DSBs are compatible with the experimental scavengeable damage fraction. The simulation application also demonstrates agreement with experimental data of γ -H2AX yields for gamma ray irradiation. Using this application, it is now possible to predict biological response along time through track-structure MC simulations.

Modeling of ionising radiation induced DNA damage and subsequent biological responses such as DNA repair processes is still a challenging issue due to the complexity of the problem. The response of a living cell when exposed to ionising radiation can be measured by means of biological assays quantifying reproductive cell death and other biological markers such as the cell survival curve^{1,2}, DNA fragment yields using electrophoresis^{3,4} and accumulation of DNA-repair protein through *foci* detection⁵⁻⁷. According to experimental findings, biological endpoints ascribed to the mutation triggered by DNA damage strongly depend on incident particle type and energy. These dependencies are, typically, discussed as a function of Linear Energy Transfer (LET) of the radiation field⁸. From a more microscopic point of view, it is known that the complexity of DNA damage affects its lethality. The damage complexity is related to the microscopic energy deposition pattern in the region surrounding

¹Department of Accelerator and Medical Physics, National Institute of Radiological Sciences, QST, Chiba, Japan. ²Joint Institute for Nuclear Research, Dubna, Russia. ³Dubna State University, Dubna, Russia. ⁴INSERM, UMR 1037, CRCT, Université Paul Sabatier, Toulouse, France. ⁵UMR 1037, CRCT, Université Toulouse III-Paul Sabatier, Toulouse, France. ⁶Medical Physics Laboratory, Medical School, University of Ioannina, 45110 Ioannina, Greece. ⁷Centre For Medical Radiation Physics, University of Wollongong, Wollongong, Australia. ⁸Geant4 Associates International Ltd, Hebden Bridge, UK. ⁹Tomsk State University, Tomsk, Russia. ¹⁰Unaffiliated, Bordeaux, France. ¹¹Unaffiliated, Melbourne, Australia. ¹²Vinca Institute of Nuclear Science, University of Belgrade, Belgrade, Serbia. ¹³Univ. Bordeaux, CNRS, CENBG, UMR 5797, Gradignan 33170, France. ✉email: sakata.dousatsu@qst.go.jp

DNA molecules as discussed for many years in microdosimetric/nanodosimetric studies^{9–12}. Ionising radiation induced DNA damage is created by direct interaction of DNA molecules with particles and indirect interaction of DNA with molecular species created during water radiolysis (typically reactive oxygen species such as hydroxyl radicals)¹³. However, a comprehensive understanding of how ionising radiation leads to biological endpoints, across multi-scale radiobiological phenomena, is yet to be reached. In conventional approaches, to model biological response bridging a gap between DNA damage and biological endpoints, DNA damage yield and its lethality are usually estimated by empirical models^{14,15}. Such models are not directly connected to track structures themselves. Ideally, in order to fully understand radiation induced DNA damage, a mechanistic model associating track structure of radiation with biological endpoints is required. Such a platform should be equipped with abilities to evaluate not only the yield of DNA damage but also the detailed description of the damage, for instance, the fraction of direct/indirect damage induced by physical/chemical interactions respectively as well as the complexity of DNA damage¹⁶.

Monte Carlo (MC) simulation codes such as KURBUC^{10,17,18}, PARTRAC^{19–22} and Geant4-DNA^{23–26}, which is a low energy extension of the MC toolkit Geant4^{27–29}, are useful to evaluate the number of DNA damage events and their complexity induced by an ionising radiation field. These MC simulations are able to calculate physical, physico-chemical and chemical interactions^{30,31}. Recently, using Geant4-DNA, an integrated simulation application has been developed to evaluate proton induced DNA damage in a human cell nucleus model composed of fractally distributed chromatin fibre³² (called hereafter Geant4-DNA_2019) and including an efficient modelling of radiolysis, proposing an alternative to the previous Geant4-DNA works by Meylan et al.³³ (called hereafter Geant4-DNA_SM) and by Rosales et al.³⁴. The simulations using Geant4-DNA_2019 have been validated against experimental DNA damage^{35–40} by means of electrophoresis, e.g. agarose gel electrophoresis (AGE)³ and pulsed-field gel electrophoresis (PFGE)⁴, and against previous MC simulations (KURBUC, PARTRAC)^{10,20,33}. With DNA damage models similar to the ones of PARTRAC, Geant4-DNA_2019 demonstrated almost equivalent simulation results to PARTRAC³² which is considered to be one of the state-of-the-art MC codes for radiation induced DNA damage.

However, until now, track-structure MC simulations are not able to predict biological response and endpoints from their simulated early DNA damage. As the last piece of developing a multi-scale platform for radiobiology, one current challenge in MC simulations is to model and predict biological responses in MC simulations^{41,42}. Hence, in this work, we have upgraded the Geant4-DNA_2019 simulation application to propose a realistic “fully integrated” MC simulation combining physics, chemistry and biology, to predict biological responses in MC simulations.

Materials and methods

We have applied three major upgrades to the Geant4-DNA_2019 simulation application to propose a realistic “fully integrated” MC simulation including the new features (see “[New features for fully integrated MC simulation](#)” *a-c* section). In the following sections, after introduction of the new features, we briefly summarise the simulation configuration (“[Simulation configuration](#)” section), damage scoring/classification (“[Damage scoring and classification](#)” section), biological repair model (“[Biological repair model](#)” section), as well as verification/validation (“[Verification and validation](#)” section). All developments and simulations are based on Geant4 version 10.4.patch2.

New features for fully integrated MC simulation

a. Building a realistic cell geometry by wrapping the cell nucleus in cytoplasm. In the previous MC application³², a naked cell nucleus was simulated for evaluating radiation induced DNA damage, however in reality, the cell nucleus is immersed in cytoplasm and surrounded by the cell membrane. Hence in this work, we have built a more realistic cell model (see “[Simulation configuration](#)” section), placing an ellipsoidal water absorber around the cell nucleus in order to model the cytoplasm.

b. Introducing biological repair model. The previous application could only calculate early DNA damage within a few nano seconds after irradiation. In this work we extended the functionality of the simulation by introducing a semi-empirical model to predict the *foci* accumulation yield along time, up to 25 h after irradiation (see “[Biological repair model](#)” section). The model was developed by Belov et al.¹⁵ and requires the yield of Double Strand Break (DSB) and the fraction of complex DNA damage (see “[Damage scoring and classification](#)” section) as input.

c. Refinement of model parameters for direct damage and indirect damage scoring. The third upgrade is a re-adjustment of the model parameters to calculate direct and indirect damage. For the re-adjustment, in addition to considering DSB yield, we introduce damage patterns of plasmids, the indirect damage fraction and DNA repair (see “[Verification and validation](#)” section) as criteria to determine more realistic model parameters.

As a criterion of the re-adjustment, we performed simulations without the histone scavenging function in order to validate the simulation against the simple DNA fibre of plasmids. Because of experimental difficulties in measuring the yield of Single Strand Breaks (SSBs) for mammalian cells, the previous MC simulations were validated for DSB yields only (and/or fragment distributions separated by a pair of DSBs). Instead, for plasmids, it is possible to measure both the number of SSB and number of DSB for validation, even though the base-pair (bp) density of plasmids is smaller than the bp density of mammalian cell (c.f. $\rho_{\text{plasmid}} \sim 9.4 \times 10^{-6}$ bp/nm³, $\rho_{\text{cell}} \sim 1.5 \times 10^{-2}$ bp/nm^{3,43,44}), and the structure of circular plasmid DNA fibres is twisted forming a supercoil. Hence, in this work, we performed simulations with and without the histone scavenging function in order to

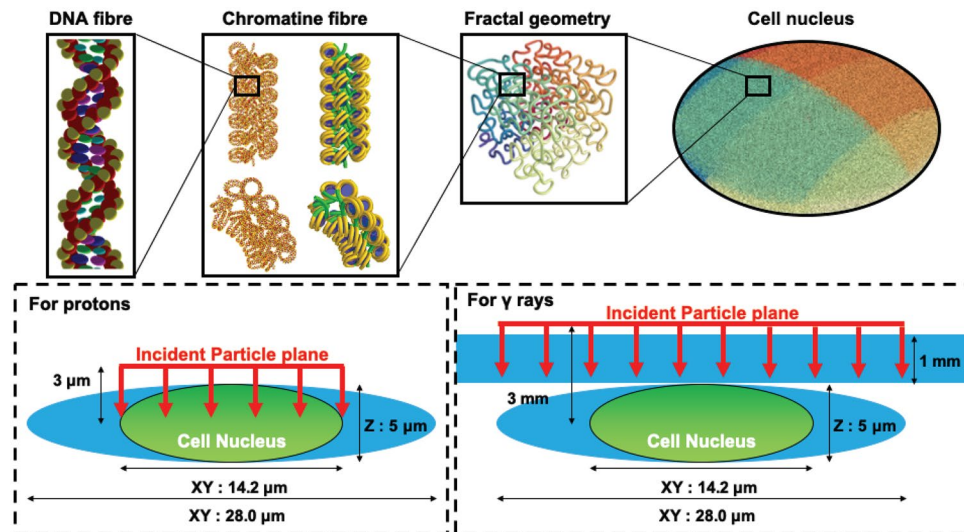


Figure 1. Geometrical model of cell nucleus and its sub-structure. Double helix DNA (left top) is wrapped by histone forming chromatin fibre (2nd left top, left: double helix DNA/right: DNA fibre and histone). Chromatin fibres are assembled as fractal (3rd left top) and confined in ellipsoidal cell nucleus (4th left top). Materials surrounding cell nucleus and beam geometry are shown in bottom panels, for both protons and gamma rays.

obtain a validation against both mammalian chromosome and simple DNA fibre of plasmid. Also, for the validation of indirect damage fraction, we simulate the fraction of indirect DSBs and compare to it the experimental scavengable damage fraction estimated as the maximum scavengable cell death fraction at infinite hydroxyl scavenger concentration.

In accordance with our previous works^{45,46}, to perform more realistic simulations, the chemical stage is modelled for an extended period of time up to 5 ns, limiting the simulation time to reasonable value (see “Simulation configuration” section), which is two times longer than the chemical stage duration considered in Geant4-DNA_2019.

Simulation configuration. The cell nucleus geometry, used in this work, has been developed in Geant4-DNA_2019³². The simulation geometry is presented in Fig. 1. A unit base pair (bp) of DNA is composed of six molecules consisting of two pairs of spherical phosphate molecules (H_3PO_4) and spherical deoxyribose molecules with two ellipsoidal nucleotide bases (guanine, adenine, cytosine and thymine) as a backbone. The combination of bases (adenine-thymine/cytosine-guanine) of each nucleobase pair is chosen randomly. As shown in left top panel of Fig. 1, these DNA units compose DNA fibre forming a twisted structure as known as double-helix structure with a 34° rotation angle per base pair^{47–49}. In the cell nucleus model, the DNA fibre is folded compactly by spherical histones and confined in a cube 75 nm on each side (shown in 2nd left top panel of Fig. 1), forming the fractal base chromatin fibre geometry (shown in 2nd right top panel of Fig. 1). Applying spherical or ellipsoidal mask to the piled up fractal chromatin geometry, it is possible to define any cell nucleus geometrical model (shown in right top panel of Fig. 1). The number of base pairs accommodated in typical human fibroblast cell is around 6 Gbp within a nucleus volume of $\sim 500 \mu\text{m}^3$ ($\rho_{\text{fibroblast}} \sim 0.012 \text{ bp}/\text{nm}^3$)^{50,51}. The typical volume of fibroblast cells is around $2000 \mu\text{m}^3$ ⁵¹. Hence, in this work, a $14.2 \mu\text{m} \times 14.2 \mu\text{m} \times 5.0 \mu\text{m}$ long ellipsoidal cell nucleus ($\sim 528 \mu\text{m}^3$) is surrounded by an ellipsoid of water ($28.0 \mu\text{m} \times 28.0 \mu\text{m} \times 5.0 \mu\text{m}$, $\sim 2052 \mu\text{m}^3$) modelling the cytoplasm. The total number of base pairs is around 6.4 Gbp and the corresponding bp density in the cell nucleus is approximately $0.012 \text{ bp}/\text{nm}^3$. Experimentally, a substrate made of glass or polymer is placed downstream of beam, on which cells are plated. To simulate particle transport in the substrate, we need to know the specific details of the substrate, such as size, composition of material, density and mean excitation energy (so called I-value). In general, it’s difficult to know such details if they are not provided. However, in the literature, the particle energies (excluding gamma rays) and their unrestricted LET at the cell entrance have been already estimated^{35–40}. Likewise, in the corresponding literature about scavengable damage fractions, the particle energies (excluding X-rays) and their dose averaged LET have been estimated^{52,53}. This is the reason why, in this work, we did not place any substrate in the simulations for proton irradiation to fairly compare the simulated results against the experimental results on DSB yields and scavengable damage fraction (as shown in left bottom panel of Fig. 1). For gamma rays, neither particle energy spectra at the cell entrance or material detail of the substrate are provided. In the *foci* experiment⁵⁴, the cells are plated in a Falcon T25 cell culture flask, however, it is difficult to access the details of substrate. Hence, as an alternative approach, we place a 1 mm thick water absorber just before of the cell geometry for gamma ray irradiation (shown in right bottom panel of Fig. 1).

The simulations are performed by evaluating DNA damage in a cell nucleus irradiated by mono-energetic protons. The initial energies of protons are chosen to be 0.3, 0.4, 0.7, 1.0, 1.67, 2.34, 4.0, 7.0 and 50 MeV. They correspond to a wide range of unrestricted linear energy transfer (LET_∞) between 1.2 and 54.41 keV/ μm . The lowest energy of incident protons is higher than the limit of Geant4-DNA_2019 (0.1 MeV), due to the unreliability of

	KURBUC ¹⁸	PARTRAC ²⁰	Geant-DNA_SM ³³	Geant4-DNA_2019 ³²	This Work
R_{dir} (Å)	1.7-3.25 [◇]	2VDWR [◇]	VDWR *	4.5	3.5
E_{min}^{break} (eV)	17.5	5	17.5	5	5
E_{max}^{break} (eV)	17.5	37.5	17.5	37.5	37.5
P_{OH}^{break}	0.13	0.7 [♣]	0.4	0.4	0.405
T_{chem} (ns)	1	10 [♣]	2.5	2.5	5
d_{kill}^{chem} (nm)	4	12.5*	N/A	4.5	9

Table 1. Damage parameters and chemistry limits of this work and previous MC simulations. R_{dir} : Accumulation radius of energy deposition from nucleotide centre. E_{min}^{break} : Minimum energy of direct strand break probability model. E_{max}^{break} : Maximum energy of direct strand break probability model. P_{OH}^{break} : Probability of indirect strand break. T_{chem} : Time limit of chemical diffusion. d_{kill}^{chem} : Production range limit of chemical radiolysis species from nucleotide centre. (VDWR) Summing up of atomic volume with each atomic van der Waals Radius (1.2, 1.7, 1.5, 1.4, 1.9 Å for H, C, N, O, P respectively). (◇) Arch structure with 1.7×3.25 Å section. (◇) To Adjust for cross-section of molecules and take into account hydration shell, VDWR multiplied by 2. (♣) Only damage with deoxyribose is considered as indirect damage. (♣) Additional 2.5 ns hydroxyl radical scavenging is considered. (*) Distance from centre of chromatine fibre. (★) Additionally 24 water molecules considered as hydration shell.

the application in the very high LET domain. The simulations are repeated for gamma rays emitted by a ⁶⁰Co source (energy equal to 1.17 MeV and 1.33 MeV with the same frequency) and a ¹³⁷Cs source (energy equal to 661.7 keV, 32.1 keV and 36.5 keV with the frequency of about 0.92, 0.06 and 0.01, respectively). The radiation source is homogeneously distributed on a 14.2 μm diameter circle plane which corresponds to the 2D-profile of the ellipsoidal cell nucleus for DSB yields and scavengeable fraction (shown as red line in left bottom panels of Fig. 1), and a 28 μm diameter circle plane which corresponds to the 2D-profile of the ellipsoidal cell for foci yield to keep the uniformity of radiation field at the cell nucleus level (shown as red line in right bottom panels of Fig. 1).

DNA geometries are defined as liquid water absorber with scaled density of the corresponding biological material similarly as in our previous work³². G4EmDNAPhysics_option4⁵⁵⁻⁵⁹ has been used to model the particle interactions in the cell for protons, photons, and low energy electrons below 10 keV, since the underlying model provides an improved implementation of the dielectric response of liquid water which is important at low energies. Above 10 keV, for electron interactions, G4EmDNAPhysics_option2 is utilized since the high energy limit of G4EmDNAPhysics_option4 for electrons is 10 keV. These physics lists adopt a track structure approach which is important when describing particle interactions at nano scale level⁶⁰. Geant4-DNA provides an interface for chemistry simulations^{61,62}. The production and reaction schemes of radiolysis for chemical species are utilized in the default configuration of the Geant4-DNA chemistry with a newly developed independent-reaction time (IRT) method^{63,64} which allows for fast-diffusion control of the radiolysis species instead of the default step-by-step method. We model histones as perfect scavengers for all radiolysis species. This means that, in the simulations, any free radiolysis species that enters a histone region (modelled as a 2.5 nm radius sphere) will be stopped and terminated. Additionally, we kill all radiolysis species created more than 9.0 nm away from any DNA molecule centres (d_{kill}^{chem}), in order to reduce computing time since they will likely be scavenged, i.e., the probability of interactions of DNA molecules and the radiolysis species is small. This range cut is equivalent to the maximum diffusion distance of hydroxyl radical (which could cause indirect damage) at 5.0 ns (T_{chem}), since all simulations are stopped at 5.0 ns after the diffusion started. These limits of chemical diffusion are summarized in Table. 1.

Overall, total computing time for 100 incident protons of 10 MeV is about 15 h including geometry building (~5 h) using Intel Xeon CPU E5-2630 v2 (2.60 GHz). It's almost double compared to Geant4-DNA_2019 (~7 h in total), and the time difference is getting longer in the high LET domain. On the other hand, Geant4-DNA_SM requires approximately 10 days³³ for similar simulation conditions using step-by-step chemistry without production limit distance of chemical species (d_{kill}^{chem}).

Damage scoring and classification. The damage scoring model and the classification method of DNA damage were originally proposed by Nikjoo et al.¹⁸. The direct DNA damage model was upgraded in PARTRAC²⁰ using a more realistic damage probability by means of a linearly increasing function of energy deposition, substituting the threshold probability model used in KURBUC. In order to calculate direct/indirect DNA damage in Geant4-DNA_2019, the same models of PARTRAC were used²⁰.

Estimating the model parameters related to probabilities of direct damage and indirect damage is one of the most uncertain part of such simulations. The model parameters of this work are summarized in Table 1 and compared to other MC codes. As listed in the table, the parameters vary covering a wide range. Usually, these parameters are adjusted within a reasonable range. For instance, PARTRAC assumes the total probability of direct damage adjusting the radius of nucleotide (R_{dir} 1.3–1.4 times larger than the nucleotide size)²⁰. Another assumption of PARTRAC is that it considers the semi-bounded 10 water molecules surrounding nucleotide²⁰, so called hydration shell⁶⁵⁻⁶⁷, which allows direct damage by charge transfer on these water molecules. Overall, R_{dir} used in PARTRAC simulations is two times more than the nucleotide size. Compared to Geant4-DNA_2019,

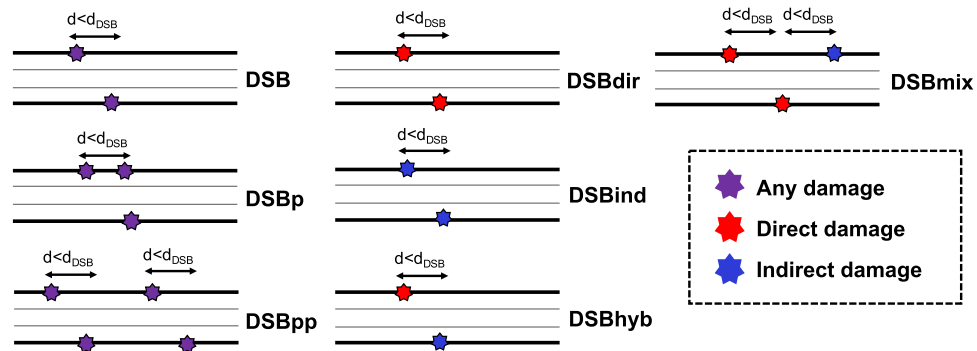


Figure 2. The scheme of classification for complexity of DSB (left) and for source of DSB (centre/right)^{31,47}.

we additionally used plasmid-SSB, plasmid-DSB, scavangeable fraction and *foci* accumulation as criteria for the adjustment of the parameters.

To calculate the probability of direct strand breaks in this work, first of all, all energy deposition was assigned to the closest strand molecule when the position is within 3.5 Å from the centre of a strand molecule in the case of either sugar or phosphate (R_{dir}). This distance is larger than the radius of nucleotide molecules ($R_{phosphate} \sim 2.28$ Å and $R_{deoxyribose} \sim 2.63$ Å) to take into account charge transfer effect in hydration shell, and smaller than the size used in PARTRAC simulations. After each simulation of the tracking of an incident particle and associated secondary radiation field, the total accumulated energy deposit in scoring region is used to determine the probability that a strand break occurred along the DNA fibres. Based on a linearly increasing probability distribution, there is a 0% probability a break occurred when less than 5 eV (E_{min}^{break}) was deposited, but a 100% probability a break occurred when over 37.5 eV (E_{max}^{break}) was deposited in the sugar phosphate moiety, as done in PARTRAC simulations.

The number of indirect damage depends on the value chosen for the likelihood of a chemical reaction between a hydroxyl radical and the sugar phosphate backbone leads to a single strand break (SSB). As the result of adjustment, to increase indirect damage, the probability of indirect damage (P_{OH}^{break}) is set to 0.405 which is almost the same (within $\sim 1\%$) with the value used in Geant4-DNA_2019, and smaller than PARTRAC simulations.

In the present work, all strand breaks (SBs) are classified either as direct SB or indirect SB. Usually, early DNA damage is quantified as number of DSBs which is considered as being two SBs on opposite strands within short distance between each other. The complexity of DSB is also scored when quantifying DNA damage. The complexity model for DSB entails a parameter d_{DSB} which is the maximum separation length between two damage sites on opposite sides of a DNA strand to consider that a DSB has occurred (typically $d_{DSB} = 10$ bp). The complexity types are shown left panel of Fig. 2. A DSBp (DSB plus) requires a DSB and at least one additional break within d_{DSB} , while a DSBpp (DSB plus-plus) requires at least two DSBs along the chromatin fibre segment. Since the lengths of segment are different (KURBUC: 216 bp, This work: straight: 7330 bp; turn and turn-twisted: 5023 bp), we made a constraint with interval of bp length. It will identify a separate damage as occurring when there is a sequence of unbroken DNA strictly greater than 100 bp between two damage events. For DSB complexity, the most complex break type is always selected for the damage cluster. Nikjoo et al.¹⁸ also proposed the classification scheme of DSB breaks by direct/indirect damage source as shown in the middle and the right panel of Fig. 2, since the contribution of chemically scavangeable DNA damage is also of particular interest. When classifying breaks by source, we pay attention to not only damage type, but also to the damage sources. DSBs composed by two indirect damage are classified as DSBind, and those only by direct damage are classified as DSBdir. DSBhyb requires that the DSB doesn't occur in the absence of indirect damage when a segment contains both indirect and direct DSBs. Otherwise, a break caused by indirect and direct sources is classified as DSBmix. Similarly, when a segment contains a DSB composed by both direct and indirect damage, classified as DSBhyb in conjunction with a direct damage for classification.

Biological repair model. In this work, we introduced a semi-empirical theoretical model for predicting biological repair effect presented by Belov et al.¹⁵. The repair model computes yields of accumulated proteins for four principal repair pathways to repair DSBs, namely non-homologous end-joining (NHEJ), homologous recombination (HR), single-strand annealing (SSA), and alternative end-joining mechanism (Alt-NHEJ).

In these considerations, the kinetics of DSB induction can be calculated using the number of complex DSB (N_{cDSB}) and of non-complex DSB (N_{ncDSB}) by¹⁵

$$\frac{dN_0}{dt} = \alpha(L) \frac{dD}{dt} N_{cDSB} - V_{NHEJ} - V_{HR} - V_{SSA} - V_{microSSA} - V_{AltNHEJ}, \quad (1)$$

where $N_0 = N_{ncDSB} + N_{cDSB}$; V_{NHEJ} , V_{HR} , V_{SSA} , $V_{microSSA}$, and $V_{AltNHEJ}$ are the terms characterizing elimination of DSBs by the NHEJ, HR, SSA, micro-SSA, and Alt-NHEJ repair pathways respectively. D is the dose of ionising radiation (Gy) and $\alpha(L)$ is the slope coefficient of linear dose dependence which describes DSB induction per unit of dose (Gy^{-1} per cell) and depends on LET. To simulate the processing of DNA lesions by repair enzymes the mass-action kinetics approach is used. The kinetics of DSB induction and remaining DSB is simulated using

differential equation (1) describing a balance between increase (term $\alpha(L) \frac{dD}{dt} N_{cDSB}$) and decrease of DSB level due to the action of all the repertoire of the repair pathways. This model includes dose rate (dD/dt) effect during irradiation, however in this work, it is assumed that the dose is delivered with very high dose rate, then fixed $dD/dt = D$ at $t = 0$ and $dD/dt = 0$ at $t \neq 0$ as in the study by Belov et al.¹⁵. For each repair path, the DSB elimination terms are expressed as differential equation with the rate constants of the correspond protein attachment and disattachment involved in the repair processes. A dynamic change of intracellular concentrations of main intermediate complexes can be generally expressed by the following differential equation:

$$\frac{dX}{dt} = V_+(X_i, N_0) - V_-(X_i, N_0) \quad (2)$$

where X_i ($i = 1, \dots, n$) is the intracellular level of the i -th attachment/disattachment processes of each pathway, t is time, n is the total number of processes, the functions V_+ and V_- describe the complex attachment and disattachment, respectively. V_+ is the summation of intracellular protein concentrations multiplied by each protein reaction rate constant at the i -th process. Similarly V_- is the summation of intracellular protein degradation. The dimensionless form of the system of differential equations refers to calculation of each pathway as well as to its rate constants of target protein. More details are described in the work by Belov et al.¹⁵ In total, the repair model has 53 rate constants along the repair paths. Most of the rate constants are evaluated by fitting to the experimental data on kinetics of different stages of DSB repair¹⁵. Solving the multi differential equations, the yields of target protein along time are calculated. In this work, five *foci* accumulation yields can be calculated for Ku, DNA-PKcs, RPA, Rad51 and γ -H2AX, although, we mainly discuss in terms of *foci* signal from γ -H2AX using specific immunocytochemistry antibodies raised against γ -H2AX. Using γ -H2AX *foci* as a biomarker of radiation exposure, the repair time-curve is estimated by means of the quantification of *foci* per nucleus at corresponding times after irradiation⁵. When DNA damage arises forming DSB, it is usually followed by the phosphorylation of the variant H2AX of the H2A protein family, which is a component of the histone octamer in nucleosomes. It is phosphorylated by kinases such as phosphatidylinositol-3 kinase-related kinase (PIKK) family of proteins which include DSB repair enzymes, ataxia telangiectasia mutated (ATM), DNA-PKcs and ATM and RAD3-related (ATR)⁶⁸, and the phosphorylated H2AX is called γ -H2AX. According to past studies, major contribution of radiation-induced γ -H2AX *foci* formation is DNA-PKcs and activated ATM⁶⁹. Hence, *foci* induction by Michaelis–Menten kinetics is calculated for the time-curve by summing up all active forms of DNA-PKcs and ATM which remain in defective cells as either DNA-PKcs or LigIV after irradiation.

In this work, it is assumed that the number of complex DSB can be calculated by $N_{cDSB} = N_{DSBp} + 2 \times N_{DSBpp}$, where N_{DSBp} is the number of DSBp and N_{DSBpp} is number of DSBpp, though DSBpp could contain more than two DSBs (the number of complex DSBs composed by three DSBs and more DSBs is neglected by definition). In order to obtain these curves, we used the NHEJ, HR, SSA, and Alt-NHEJ models with zero initial conditions for all intermediate complexes and corresponding values of cDSB.

Verification and validation. The simulation platform is verified against past MC studies and validated against experimental measurements for both protons and gamma rays. Before classifying DNA damage clusters into either SSB or DSB, the total number of strand breaks and their contributions to direct/indirect damage are compared to the past MC simulations for verification. After the DNA damage clusters are classified, both SSB and DSB yields are validated against experimental measurements and benchmarked against the past MC simulations^{18,20,32,33}, followed by validation of scavengeable DNA damage fraction and time-curve of *foci* yield.

Since the cell geometry is modeled imitating a normal human fibroblast cell, the experimental results for normal human fibroblast cells are chosen for validation of the DSB yields and the *foci* yields. In addition, as a reference, we have compared simulated DSB yields in simple DNA fibre against experimental DSB yields of plasmid. For scavengeable damage fraction, normal Chinese hamster fibroblast cell (V79) is selected for the validation, because it is one of the most popular cell lines in radiobiology, and a systematic experimental study of scavengeable damage fraction for human fibroblast cell is not available. Compared to human fibroblast cell, V79 is smaller (average diameter of nucleus $\sim 8 \mu\text{m}$ and cell $\sim 10 \mu\text{m}$ ⁷⁰) and number of base pairs is smaller ($\sim 3.9 \text{ Gbp}$ ⁴⁰), but the base pair density is higher ($\rho_{V79} \sim 0.015 \text{ bp/nm}^3$).

Experimentally, there are mainly two ways to evaluate DSB yields. The first method is based on the separation of DNA fragments through AGE or PFGE. The method can be performed both with plasmid and cell. To compare simulation results to these experimental data both for plasmid and mammalian cell, we have tried to simulate DSB yields with/without histone scavenging function. The simulation without histone scavenging can be regarded as simulation for simple DNA fibre such as plasmid, despite the bp density is different. In addition, unlike in cells, it's possible to quantify not only DSB yield but also SSB yield using plasmids easily. The SSB/DSB ratio provides more insight of the details of radiation induced DNA damage. As an experimental constraint of AGE and PFGE in cells is the difficulty to count DSBs in small fragments (typically shorter than 5–25 kbp), since the measurement is accompanied by the loss of small fragments when cells are embedded in gel plugs³⁵. This is why, we simulate not only the total number of DSBs but also that of “distant” DSBs, those which are separated by at least 10 kbp gaps between two DSBs. The DSB yields measured by electrophoresis are shown as a function of unrestricted linear energy transfer (LET_∞), which is recommended by the ICRU 90⁷¹ excluding the case of gamma rays from ⁶⁰Co. Radiation induced DNA damage is strongly related to the beam quality which is commonly quantified using (the unrestricted) LET. To validate DSB yields with high LET protons, ideally, we need to compare the simulated results against low energy proton irradiation experiments. However, experimentally, it is difficult to perform cellular irradiation experiments with such low energy protons. Hence in this study, as reference, we show the simulated DSB yields compared against the experimental results with helium at high LET.

We note that, in the clinically relevant energy domain, it is well known that DSB yield is depending on the LET, despite that the LET dependence cannot entirely explain the observed biological effects⁸. In the case of gamma rays from ⁶⁰Co, we assume LET = 0.5 keV/μm which is the upper-limit of LET for the ⁶⁰Co photons³⁹. Likewise, the scavengeable damage fractions are also shown as a function of LET, since the experimental data is not for protons but for carbons and X-rays. In the literature of experiments measuring scavengeable damage fractions, incident energy and dose averaged LETs have been described, hence in this study, we converted the incident particle energy to LET_∞ excluding the case of X-rays. For X-rays, we set the dose-averaged LET at 9.4 keV/μm as listed in the literature⁵³.

The second method is based on detection of *foci* representing the accumulation of proteins related to DSB repair process⁵⁻⁷. *Foci* measurement has become a standard to evaluate radiation induced DNA damage, however, the relation between *foci* yield and number of DSBs is still unclear⁷²⁻⁷⁴. Hence in this work, only experimental data by AGE and PFGE are compared to the simulated number of DSBs as benchmark, because of the difficulty of evaluating absolute values of DSB yields by *foci* measurements. Whilst *foci* measurements cannot be used to determine the number of DSBs, they can be used to evaluate DNA repair processes with respect to time. The calculated time-curve of protein accumulation which is newly implemented into the simulation application is compared with experimental γ-H2AX yield irradiated with gamma rays from a ¹³⁷Cs source.

The scavengeable damage fraction can be measured from the maximum degree of protection (*DP*) at infinite dimethylsulfoxide (DMSO) concentration. DMSO has been used as scavenger of radiolysis free radical species, in particular hydroxyl radical. DMSO is permeated among the Chinese hamster cell V79 before irradiation with different concentration up to 1.0 mol. The *DP* is calculated as follows:

$$DP = \frac{\ln SF_0 - \ln SF_x}{\ln SF_0}, \quad (3)$$

where SF_0 and SF_x are the survival fraction at 0 and x mol of DMSO, respectively. A regression line can be plotted over the reciprocal *DP* via fitting to the experimental reciprocal *DP*s with several concentrations of DMSO by Eq. (4) to estimate intercept of the fitting function.

$$\frac{1}{DP} = k \cdot \frac{1}{x} + y_\infty, \quad (4)$$

where k is the slope of the equation and y_∞ is the intercept of the function giving the inverse of the maximum *DP* at infinite DMSO concentration ($1/x = 0$). To compare the scavengeable damage fraction obtained by the maximum *DP*, we calculate the fraction of scavengeable DSBs which could not be classified as DSB without indirect damage by:

$$\frac{N_{DSB_{ind}} + N_{DSB_{hyb}}}{N_{DSB_{dir}} + N_{DSB_{mix}} + N_{DSB_{ind}} + N_{DSB_{hyb}}}, \quad (5)$$

where $N_{DSB_{dir}}$, $N_{DSB_{ind}}$, $N_{DSB_{mix}}$ and $N_{DSB_{hyb}}$ are the numbers of DSB_{dir}, DSB_{ind}, DSB_{mix} and DSB_{hyb}, respectively. The calculations are presented as a function of unrestricted linear energy transfer (LET_∞). By using Nikjoo's classification, we neglect the contribution of *DSB_{mix}* towards the scavengeable damage fraction, since it cannot be clearly classified as scavengeable DSB or not.

Results

The total number of strand breaks (SBs) is plotted in Fig. 3 as a function of unrestricted LET, as well as the number of direct and indirect damage events. The yield of both direct and indirect damage is close to the yield of Geant4-DNA_SM in the low LET domain. The yield of indirect damage is gradually decreasing with LET for Geant4-DNA_SM. With the increase of LET, the total number of radiolysis species is increasing. Unlike indirect damage, there is no significant LET dependence in the case of direct damage.

When neglecting histone scavenging, the yield of indirect damage is strongly enhanced across the simulated LET range. For instance, the yield without histone scavenging functionality for radiolysis species at 10 keV/μm is decreasing from around 350 down to 200 Gy⁻¹Gbp⁻¹, since more hydroxyl radicals can survive if the radiolysis species are not scavenged by histone.

Figure 4 shows the simulated SSB, DSB yields and their ratio as a function of unrestricted LET, including previous simulations^{18,20,32} and experimental results³⁵⁻³⁹. Similarly to the results deriving from the PARTRAC simulations, the simulated SSB yield is found to decrease with LET, although higher than PARTRAC at low LET (below 30 keV/μm) and smaller at high LET (above 40 keV/μm). If the histone scavenging effect is turned off, the SSB yield is getting larger, in particular at low LET, and the order of magnitude of SSB yield is reaching the yield level of the experimental SSB yield of plasmid.

As explained in “Verification and validation” section, to take into account experimental biases on the DSB yield, we calculate both the total DSB yield (solid red line) and the 10 kbp distant DSB yield (dotted red line) for proton irradiation. Also the 5 kbp distant DSB yield of gamma rays from ⁶⁰Co is presented as open red square where the limit of bp length is estimated from the experiment³⁹. The total DSB yields are continually increasing with LET, while the difference between total DSBs and distant DSBs is also increasing because of the loss of small fragments in experimental procedure. The distant DSB yields for protons are approaching experimental data obtained by electrophoresis with a 13.3% difference on average. At the same time, the distant DSB yield for gamma rays is in agreement with the corresponding experimental data within 0.6%.

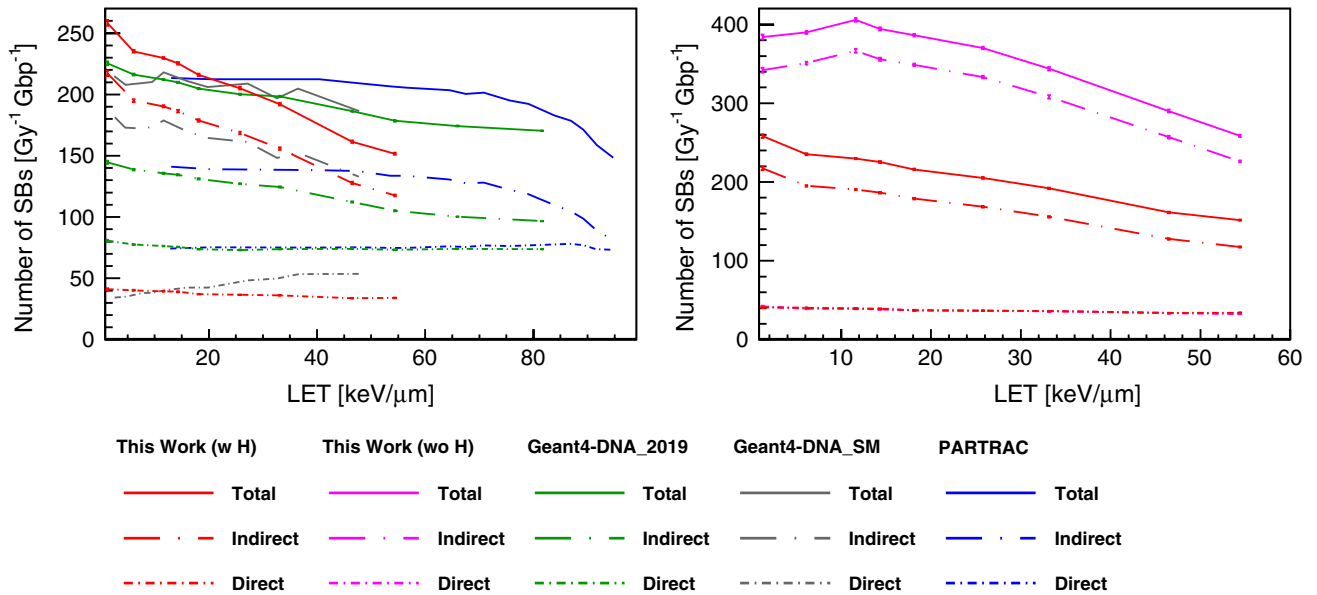


Figure 3. Number of strand breaks (SBs) per Gy and per Gbp induced by protons in a cell nucleus as a function of unrestricted LET. (Left: Comparison with the previous simulations. Right: Comparison with and without histone scavenging functionality). The solid lines show the total SB yield; the long dot-dashed lines show the indirect SB yield; the short lines show the direct SB yield. For this work, two types of histone scavenging conditions have been applied, simulation results with perfect scavenging (w H) are shown as red lines, and without scavenging functionality (wo H) are shown as magenta lines. The direct damage yields of this work with histone and without histone are overlapped with each other.

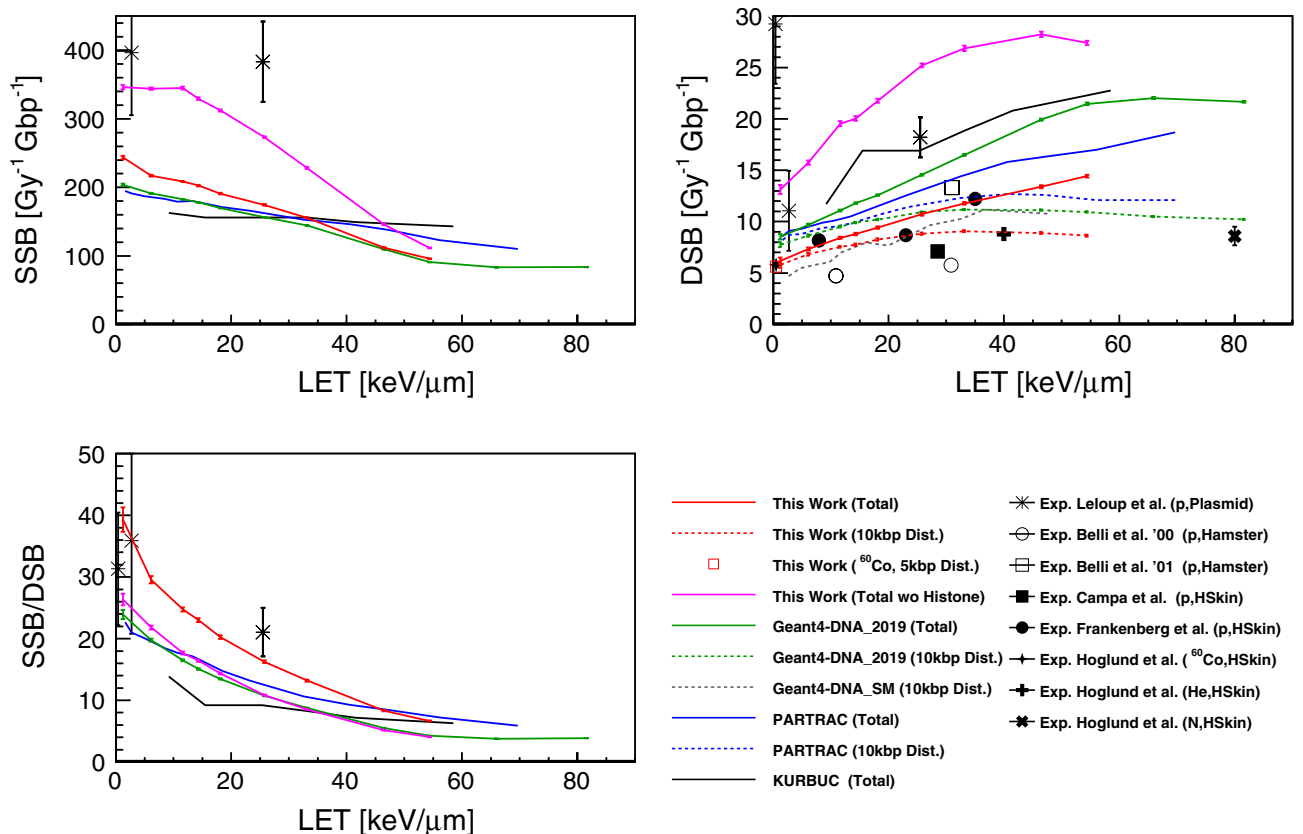


Figure 4. Left top: SSB yield per Gy and per Gbp as a function of unrestricted LET. Right top: DSB yield per Gy and per Gbp as a function of unrestricted LET. Left bottom: SSB/DSB ratio as a function of unrestricted LET. (HSkin: Human skin fibroblast cell). The simulated results are compared to the experimental data^{35–40}.

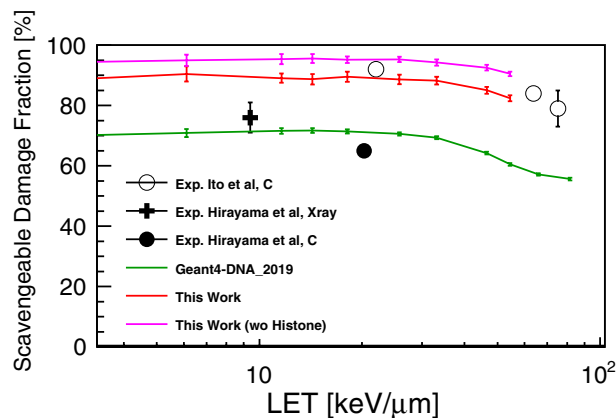


Figure 5. The scavengeable damage fraction by means of the ratio of scavengeable DSBs compared to the total number of DSBs. The simulated fraction is compared to the experimental scavengeable degree of V79 cell estimated at the infinite DMSO concentration^{52,53}.

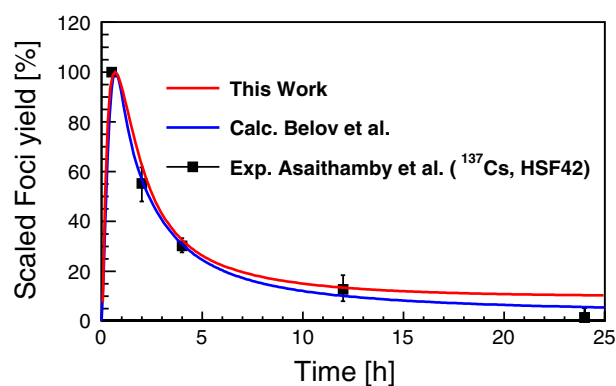


Figure 6. γ -H2AX yield as a function of time from irradiation by gamma rays from ^{137}Cs at the dose of 1 Gy. The predicted foci yield is calculated with simulated number of DSBs and complex DSBs as inputs. The experimental data is for normal human skin fibroblasts HSF42⁵⁴.

In terms of SSB/DSB ratio, the ratio decreases as a function of LET. This means that for higher LET values (above $40 \text{ keV}/\mu\text{m}$), radiation induced DNA damage becomes more significant because the DSB yield becomes bigger than the SSB yield.

The proton induced scavengeable damage fraction is shown in Fig. 5. At low LET, about 90% of scavengeable damage fraction has been simulated, and with the increase of LET, the simulations indicate that the scavengeable damage fraction is decreasing. As in the simulations, the scavenging effect of histone reduces the predictable fraction by around 5% if the histone is regarded as a perfect scavenger. The simulated indirect DSBs are good agreement with the experimental scavengeable damage fraction⁵². Using previous damage parameters set used in Geant4-DNA_2019, it is possible to reproduce the other data set of the experimental scavengeable damage fraction⁵³.

Figure 6 shows the calculation results of scaled γ -H2AX yield as a function of time after irradiation up to 25 h with the simulated number of DSBs and the irreparable fraction as input parameters compared to the experimental data for normal human skin fibroblasts HSF42 exposed to gamma-rays from ^{137}Cs at the dose of 1 Gy ⁵⁴. The calculated γ -H2AX yield with optimized rate constants leads to a very good agreement with the experimental results (within 1.6% difference on average). The decreasing speed of γ -H2AX yield reasonably matches with the experimental result at 10 h.

Discussion

The indirect SB yields of this work become lower than Geant4-DNA_SM with increasing LET, as shown in Fig. 3. With the increase of LET, the total number of radiolysis species is increasing, however, the yield of unstable species decreases quickly with time because of a higher reaction probability due to higher density around particle tracks. This is the reason why indirect SB yields become lower with increasing LET. However, in terms of direct SBs, the difference between the results of this work and Geant4-DNA_SM becomes larger with the increase of LET. As we discussed in the previous work³² the difference is mainly due to the probability model of direct damage and the size of the scoring volume³². It is more evident when comparing the direct yield calculated with

this work and the yield deriving from PARTRAC simulations²⁰, since the probability model is the same and the scoring volume of PARTRAC is larger than the volume adopted here (see Table 1).

The simulated yields of distant DSBs for protons and gamma rays from a ⁶⁰Co source are in good agreement with the experimental results (within 13.3% for protons on average and 0.7% for gamma rays) as shown in Fig. 4. In a previous study⁴⁷, the systematic uncertainty from the physics particle transport model was estimated to be around 20% on electron induced DSB yields for simple DNA fibre. Similarly, using a cell nucleus model, physics model and chemistry model can cause uncertainties up to 34% and 16% in the simulated DSB yields respectively as presented by Zhu et al.⁷⁵. Also, direct damage model, chemistry time duration, and hydroxyl damage probability can cause differences of up to 28%, 51%, and 71%, respectively⁷⁵. Therefore, the deviation from experimental data and the other simulations can be wrapped by the systematic uncertainties of the particle transport and it is not surprising that the DSB yields of this work are almost 20% lower than the yields of PARTRAC simulation. Compared to Geant4-DNA_2019, two improvements have been applied, firstly for geometry and secondly for simulation parameters. The difference between this work and Geant4-DNA_2019 mainly comes from the parameter refinement, the geometrical improvement having not much influence, (see Supplementary Figure S1 online which shows DSB yields of this work and Geant4-DNA_2019 with/without cytoplasm geometry.) Turning off the histone scavenging effect makes the SSB/DSB yield decrease. This suggest that histone scavenging has a role of protecting against radiation induced DSBs. PARTRAC simulates smaller SSB/DSB ratios compared to this work when considering histone scavenging. Hence, PARTRAC should simulate much smaller ratio if the histone scavenging effect is turned off. Compared to previous simulations³², this work is characterised by a better agreement to the experimental results in plasmid. It ought to be noted that we need to do further investigations of radiation induced DNA damage on plasmids with more realistic geometrical models^{76,77} to have more fair comparisons (at least, comparison of simulations with more realistic plasmid bp density, since bp density affects number of indirect damage⁴⁵).

When the scavengeable damage fraction is compared to the experimental data by Ito et al.⁵², the simulated fraction shows a good agreement in the LET range 20–54.41 keV/μm. However, since systematic differences can be found between the experiments by Ito et al.⁵² and by Hirayama et al.⁵³, it's difficult to get agreement with both experimental results at low and moderate LET (below a few tens keV/μm). With the increase of base pair density, fraction of indirect damage is increasing⁴⁵. Since the base pair density of V79 cell ($\rho_{V79} \sim 0.015$ bp/nm³) is higher than the density of simulated cell nucleus model ($\rho_{fibroblast} \sim 0.012$ bp/nm³), the experimental scavengeable damage fractions can be a few percent higher than the simulated scavengeable damage fractions. We note that further studies are required for simulations of scavengeable damage fraction with heavier ions for validation in the high LET domain. Also, a similar study for normal human fibroblast cells would be useful for a more comprehensive validation of this platform.

Across this work, the calculation results of scaled γ -H2AX yield have been validated for gamma-rays from ¹³⁷Cs at the dose of 1 Gy⁵⁴. Both this work and calculation by Belov et al. are close to the experimental data below 5 h. After 5 h, this work predicts *foci* yield larger than the yield calculated by Belov et al. The main difference comes from estimation of irreparable fraction. In this work, the irreparable fraction is simulated as fraction of complex DNA damage (~ 0.12), on the other hand, the fraction used in the calculated yield was estimated directly as *foci* yield fraction at 24 h measured by Asaithamby et al. (0.01). Hence, it's not surprising that the previous calculations are close to the experimental data at 24 h. It ought to be noted that, at such low yield, signal to noise ratio of the measurement is not high, since the intensity of signal light cannot be clearly separated from the background. Hence, experimentally, the systematic uncertainty of the measured *foci* yield at 24 h can be large. And due to background subtraction, detected *foci* yield can be lower than actual yield level, especially at such low yield. In the previous work by Belov et al.¹⁵, the model is able to predict yield of γ -H2AX foci in human skin fibroblasts (HSF42) irradiated with ¹⁶O, ²⁸Si, and ⁵⁶Fe. However, further investigations are required to make sure whether the application can predict yields of γ -H2AX foci for the other irradiation sources (including proton). In addition, the model parameters of biological repair, in this work, have been optimized for a human fibroblast cell. However, biological response against radiation and kinetics of DNA repair protein accumulation depend on cell type, since the repair path frequency, repair speed and its consequence varies from one cell line to another. For instance, application of the quantitative model to reconstruct the repair kinetics in cancer cells like breast carcinoma MCF-7 cells and NSCLC (lung) HTB177 cells⁷⁸, requires a deeper understanding of the particular stages of DSB repair in such cells and, consequently, another set of parameters are needed. It is clear that in order to explore the mechanism of radiation induced DNA damage, the biological prediction model will need to be optimized cell line by cell line. We note, to simulate particle transport in a substrate, the properties of the substrate should be known, such as composition of material, density and mean excitation energy. Hence, for more accurate simulations, these specific properties should be provided to compare the simulations with the experimental studies. Nevertheless in this work, we place a water absorber substituting a substrate as an alternative approach.

Conclusion

As a new important step for radiobiological simulation using Geant4-DNA, the first fully integrated simulation chain adopting IRT has been upgraded across the physical, chemical and biological stages of cellular radiation action in a single application. We have re-optimized parameters in the simulations including a more realistic cell geometry, whereby the cell nucleus is covered by a water absorber acting as cytoplasm.

To show the reliability of the application, we have simulated not only yield of DSBs for a mammalian cell but also the yield of SSB/DSB for plasmid simultaneously, as well as scavengeable DSB yields by switching off the histone scavenging functionality. Using this application, it now becomes possible to predict biological response along time through track-structure MC simulations. The simulation results for mammalian cells are in agreement

with experimental data in terms of DSB yields (within 13.3% for protons on average and 0.6% for gamma rays from a ^{60}Co source) and γ -H2AX yields (within 1.6% on average for gamma rays from a ^{137}Cs source).

Received: 16 April 2020; Accepted: 15 October 2020

Published online: 27 November 2020

References

- Puck, T. T. & Marcus, P. I. Rapid method for viable cell titration and clone production with HeLa cells in tissue culture: the use of X-irradiated cells to supply conditioning factors. *Proc. Natl. Acad. Sci.* **41**, 432–437 (1995).
- Puck, T. T. & Marcus, P. I. Action of X rays on mammalian cells. *J. Exp. Med.* **103**, 653–666 (1956).
- Kryndushkin, D. S., Alexandrov, I. M., Ter-Avanesyan, M. D. & Kushnirov, V. V. Yeast [PSI⁺] prion aggregates are formed by small Sup35 polymers fragmented by Hsp104. *J. Biol. Chem.* **278**, 49636–49643 (2003).
- Iliakis, G. E., Cicilioni, O. & Metzger, L. Measurement of DNA double strand breaks in CHO cells at various stages of the cell cycle using pulse field gel electrophoresis: Calibrations by means of 125I decay. *Int. J. Radiat. Biol.* **59**, 343–357 (1991).
- Paul, T. T. *et al.* A critical role for histone H2AX in recruitment of repair factors to nuclear foci after DNA damage. *Curr. Biol.* **10**, 886–895 (2000).
- Rogakou, E. P. *et al.* DNA double-stranded breaks induce histone H2AX phosphorylation on serine 139. *J. Biol. Chem.* **273**, 5858–5868 (1998).
- Rogakou, E. P., Boon, C., Redon, C. & Bonner, W. M. Megabase chromatin domains involved in DNA double-strand breaks in vivo. *J. Cell Biol.* **146**, 905–916 (1999).
- Sørensen, B. S., Overgaard, J. & Bassler, N. In vitro RBE-LET dependence for multiple particle types. *Acta Oncol.* **50**(6), 757–762. <https://doi.org/10.3109/0284186X.2011.582518> (2011).
- Goodhead, D. T. Initial events in the cellular effects of ionizing radiations: clustered damage in DNA. *Int. J. Radiat. Biol.* **65**, 7–17 (1994).
- Nikjoo, H., O'Neill, P., Goodhead, D. T. & Terrissol, M. Computational modeling of low-energy electron-induced DNA damage by early physical and chemical events. *Int. J. Radiat. Biol.* **71**, 467–483 (1997).
- Liamsuwan, T., Emfietzoglou, D., Uehara, S. & Nikjoo, H. Microdosimetry of low-energy electrons. *Int. J. Radiat. Biol.* **88**, 899–907 (2012).
- Sato, T. & Furusawa, Y. Cell survival fraction estimation based on the probability densities of domain and cell nucleus specific energies using improved microdosimetric kinetic models. *Radiat. Res.* **178**, 341–356 (2012).
- Hemnani, T. & Parihar, M. S. Reactive oxygen species and oxidative DNA damage. *Indian J. Physiol. Pharmacol.* **42**, 440–452 (1998).
- Stewart, R. D. Two-Lesion kinetic model of double-strand break rejoining and cell killing. *Radiat. Res.* **156**, 365–378 (2001).
- Belov, O. V. *et al.* A quantitative model of the major pathways for radiation-induced DNA double-strand break repair. *J. Theor. Biol.* **366**, 115–130 (2015).
- Nikjoo, H., Taleei, R., Liamsuwan, T., Liljequist, D. & Emfietzoglou, D. Perspectives in radiation biophysics: from radiation track structure simulation to mechanistic models of DNA damage and repair. *Radiat. Phys. Chem.* **128**, 3–10 (2016).
- Uehara, S., Nikjoo, H. & Goodhead, D. T. Cross-section of water vapour for the Monte Carlo electrons track structure code from 10 eV to the MeV region. *Phys. Med. Biol.* **38**, 1841–1858 (1993).
- Nikjoo, H., O'Neill, P., Wilson, W. E. & Goodhead, D. T. Computational approach for determining the spectrum of DNA damage induced by ionizing radiation. *Radiat. Res.* **156**, 577–583 (2001).
- Friedland, W. *et al.* Monte Carlo simulation of the production of short DNA fragments by low-linear energy transfer radiation using higher-order DNA models. *Radiat. Res.* **150**, 170–182 (1998).
- Friedland, W. *et al.* Simulation of DNA damage after proton irradiation. *Radiat. Res.* **159**, 401–410 (2003).
- Friedland, W., Dingfelder, M., Kundrát, P. & Jacob, P. Track structures, DNA targets and radiation effects in the biophysical Monte Carlo simulation code PARTRAC. *Mutat. Res.* **711**, 28–40 (2011).
- Friedland, W. *et al.* Comprehensive track-structure based evaluation of DNA damage by light ions from radiotherapy-relevant energies down to stopping. *Sci. Rep.* **7**, 45161 (2017).
- Incerti, S. *et al.* The Geant4-DNA project. *Int. J. Model. Simul. Sci. Comput.* **1**, 157–178 (2009).
- Incerti, S. *et al.* Comparison of Geant4 very low energy cross section models with experimental data in water. *Med. Phys.* **37**, 4692–4708 (2010).
- Bernal, M. A. *et al.* Track structure modeling in liquid water: a review of the Geant4-DNA very low energy extension of the Geant4 Monte Carlo simulation toolkit. *Phys. Med.* **31**, 157–178 (2015).
- Incerti, S. *et al.* Geant4-DNA example applications for track structure simulations in liquid water: a report from the Geant4-DNA Project. *Med. Phys.* **45**, e722–e739 (2018).
- Agostinelli, S. *et al.* Geant4—a simulation toolkit. *Nucl. Instrum. Methods A* **506**, 250–303 (2003).
- Allison, J. *et al.* Geant4 developments and applications. *IEEE Trans. Nucl. Sci.* **53**, 270–278 (2006).
- Allison, J. *et al.* Recent developments in Geant4. *Nucl. Instrum. Methods A* **835**, 186–225 (2016).
- Nikjoo, H. *et al.* Track structure in radiation biology: theory and applications. *Int. J. Radiat. Biol.* **73**, 355–364 (1998).
- Nikjoo, H. *et al.* Radiation track, DNA damage and response—a review. *Rep. Prog. Phys.* **79**, 116601 (2016).
- Sakata, D. *et al.* Evaluation of early radiation DNA damage in a fractal cell nucleus model using Geant4-DNA. *Phys. Med.* **62**, 152–157 (2019).
- Meylan, S. *et al.* Simulation of early DNA damage after the irradiation of a fibroblast cell nucleus using Geant4-DNA. *Sci. Rep.* **7**, 11923 (2017).
- Rosales, L. F., Incerti, S., Francis, Z. & Bernal, M. A. Accounting for radiation-induced indirect damage on DNA with the Geant4-DNA code. *Phys. Med.* **51**, 108–116 (2018).
- Campa, A. *et al.* DNA DSB induced in human cells by charged particles and gamma rays: Experimental results and theoretical approaches. *Int. J. Radiat. Biol.* **81**, 841–854 (2005).
- Frankenberg, D. *et al.* Induction of DNA Double-Strand Breaks by 1H and 4He Ions in Primary Human Skin Fibroblasts in the LET Range of 8 to 124 keV/um. *Radiat. Res.* **151**, 540–549 (1999).
- Belli, M. *et al.* DNA DSB induction and rejoining in V79 cells irradiated with light ions: a constant field gel electrophoresis study. *Int. J. Radiat. Biol.* **76**, 1095–1104 (2000).
- Leloup, C. *et al.* Evaluation of lesion clustering in irradiated plasmid DNA. *Int. J. Radiat. Biol.* **81**, 41–54 (2005).
- Höglund, E., Blomquist, E., Carlsson, J. & Stenerlöw, B. DNA damage induced by radiation of different linear energy transfer: initial fragmentation. *Int. J. Radiat. Biol.* **76**, 539–547 (2000).
- Belli, M. *et al.* DNA Fragmentation in mammalian cells exposed to various light ions. *Adv. Space Res.* **27**, 393–399 (2001).
- Barbieri, S. *et al.* Modelling γ -H2AX foci induction to mimic limitation in the scoring technique. *Rad. Protec. Dos.* **183**, 121–125 (2019).

42. Barbieri, S. *et al.* Predicting DNA damage foci and their experimental readout with 2D microscopy: a unified approach applied to photon and neutron exposures. *Sci. Rep.* **9**, 14019 (2019).
43. Suzuki, S. *et al.* Quantitative analysis of DNA orientation in stationary AC electric fields using fluorescence anisotropy. *IEEE Trans. Ind. Appl.* **34**, 75–83 (1998).
44. Ghosh, S. K. & Jost, D. How epigenome drives chromatin folding and dynamics, insights from efficient coarse-grained models of chromosomes. *PLOS Comput. Biol.* **14**, e1006159 (2018).
45. Tang, N. *et al.* Influence of chromatin compaction on simulated early radiation-induced DNA damage using Geant4-DNA. *Med. Phys.* **46**, 1501–1511 (2019).
46. Tang, N. *et al.* Assessment of radio-induced damage in endothelial cells irradiated with 40 kVp, 220 kVp, and 4 MV X-rays by means of micro and nanodosimetric calculations. *Int. J. Mol. Sci.* **20**, 6204 (2019).
47. Lampe, N. *et al.* Mechanistic DNA damage simulations in Geant4-DNA part 1: a parameter study in a simplified geometry. *Phys. Med.* **48**, 135–145 (2018).
48. Lampe, N. *et al.* Mechanistic DNA damage simulations in Geant4-DNA part 2: Electron and proton damage in a bacterial cell. *Phys. Med.* **48**, 146–155 (2018).
49. Lampe N. The Long Term Impact of Ionising Radiation on Living Systems. Ph.D thesis from University Clermont-Auvergne (2017).
50. Lamond, A. I. & Sleeman, J. E. Nuclear substructure and dynamics. *Curr. Biol.* **13**, R825–R828 (2003).
51. Milo, R. & Phillips, R. Cell Biology by the Numbers *New York, NY and Abingdon, UK. Garland Science, Taylor and Francis Group (US)*, (2016). online version available (confirmed 8 July 2020)) at <http://book.bionumbers.org/>
52. Ito, A. *et al.* Contribution of indirect action to radiation-induced mammalian cell inactivation: Dependence on photon energy and heavy-ion LET. *Radiat. Res.* **165**, 703–712 (2006).
53. Hirayama, R. *et al.* Contributions of direct and indirect actions in cell killing by high-LET radiations. *Radiat. Res.* **171**, 212–218 (2009).
54. Asaithamby, A. *et al.* Repair of HZE-particle-induced DNA double-strand breaks in normal human fibroblasts. *Radiat. Res.* **169**, 437–446 (2008).
55. Emfietzoglou, D. Inelastic cross-sections for electron transport in liquid water: a comparison of dielectric models. *Radiat. Phys. Chem.* **66**, 373–385 (2003).
56. Kyriakou, I., Šeň, M., Nourry, V. & Incerti, S. The impact of new Geant4-DNA cross section models on electron track structure simulations in liquid water. *J. Appl. Phys.* **119**, 194902 (2016).
57. Kyriakou, I., Incerti, S. & Francis, Z. Technical note: improvements in Geant4 energy-loss model and the effect on low-energy electron transport in liquid water. *Med. Phys.* **42**, 3870–3876 (2015).
58. Bordage, M. C. *et al.* Implementation of new physics models for low energy electrons in liquid water in Geant4-DNA. *Phys. Med.* **32**, 1833–1840 (2016).
59. Kyriakou, I. *et al.* Microdosimetry of electrons in liquid water using the low-energy models of Geant4. *J. Appl. Phys.* **122**, 024303 (2017).
60. Lazarakis, P. *et al.* Investigation of track structure and condensed history physics models for applications in radiation dosimetry on a micro and nano scale in Geant4. *Bio. Phys. Eng. Exp.* **4**, 2 (2018).
61. Karamitros, M. *et al.* Modeling radiation chemistry in the Geant4 toolkit. *Prog. Nucl. Sci. Technol.* **2**, 503–508 (2011).
62. Karamitros, M. *et al.* Diffusion-controlled reactions modeling in Geant4-DNA. *J. Comput. Phys.* **274**, 841–882 (2014).
63. Green, N. J. B., Pilling, M. J., Pimblott, S. M. & Clifford, P. Stochastic modeling of fast kinetics in a radiation track. *J. Phys. Chem.* **94**, 251–258 (1990).
64. Karamitros, M. *et al.* Implementing the Independent Reaction Time method in Geant4 for radiation chemistry simulations. [arXiv:2006.14225](https://arxiv.org/abs/2006.14225) (2020).
65. Falk, M., Poole, A. G. & Goymour, C. G. Infrared study of the state of water in the hydration shell of DNA. *Can. J. Chem.* **48**, 1536 (1970).
66. Swarts, S. G. *et al.* Mechanisms of direct radiation damage in DNA, based on a study of the yields of base damage, deoxyribose damage, and trapped radicals in d(GCACGCGTGC)₂. *Radiat. Res.* **168**, 367–381 (2007).
67. Alizadeh, E., Sanz, A. G., García, G. & Sanche, L. Radiation damage to DNA: the indirect effect of low-energy electrons. *J. Phys. Chem. Lett.* **4**, 820–825 (2013).
68. Mah, L. J., El-Osta, A. & Karagiannis, T. C. γ H2AX: a sensitive molecular marker of DNA damage and repair. *Leukemia* **24**, 679–686 (2010).
69. Derheimer, F. A. & Kastan, M. B. Multiple roles of ATM in monitoring and maintaining DNA integrity. *FEBS Lett.* **584**, 3675–3681 (2010).
70. Howell, R. W. *et al.* The question of relative biological effectiveness and quality factor for Auger emitters incorporated into proliferating mammalian cells. *Radiat. Res.* **128**, 282–292 (1991).
71. Report 90: Key Data for Ionizing-Radiation Dosimetry: Measurement Standards and Applications. *J. Int. Com. Rad. Units Meas.* **14**, 1 (2016).
72. Cucinotta, F. A. *et al.* Biochemical kinetics model of DSB repair and induction of γ -H2AX foci by non-homologous end joining. *Radiat. Res.* **169**, 214–222 (2008).
73. Nikitaki, Z. *et al.* Non-DSB clustered DNA lesions. Does theory colocalize with the experiment?. *Radiat. Phys. Chem.* **128**, 26–35 (2016).
74. Nikitaki, Z. *et al.* Measurement of complex DNA damage induction and repair in human cellular systems after exposure to ionizing radiations of varying linear energy transfer (LET) *Free Rad. Res.* **50**, S64–S78 (2016).
75. Zhu, H. *et al.* A parameter sensitivity study for simulating DNA damage after proton irradiation using TOPAS-nBio. *Phys. Med. Biol.* **65**, 085015 (2020).
76. Fujimoto, B. S. & Schurr, J. M. Monte Carlo simulations of supercoiled DNAs confined to a plane. *Biophys. J.* **82**, 944–962 (2002).
77. McNamara, A. L. *et al.* Geometrical structures for radiation biology research as implemented in the TOPAS-nBio toolkit. *Phys. Med. Biol.* **63**, 175018 (2018).
78. Petković, V. D. *et al.* Biological outcomes of γ -radiation induced DNA damages in breast and lung cancer cells pretreated with free radical scavengers. *Int. J. Radiat. Biol.* **95**, 274–285 (2019).

Acknowledgements

We would like to acknowledge the financial support from the CNRS PICS #7340 and PICS #8235 France - Greece, from the CNRS PICS #8070 France - Serbia, the IdEx Bordeaux University - France - International Post-doctorates program in the framework of the “France-Japan Particle Physics Laboratory” International Associated Laboratory, and the IdEx Bordeaux University - France - International Doctorates program in the framework of the “France-Korea Particle Physics Laboratory” International Associated Laboratory (2017–2020) S. Guatelli, D. Sakata, S. Incerti, I. Kyriakou and D. Emfietzoglou acknowledge financial support from the Australian Research Council, ARC DP170100967. I. Kyriakou and D. Emfietzoglou acknowledge financial support from ESA (Contract No. 4000126645/19/NL/BW). O. Belov acknowledges the funding provided by RFBR, project number

17-29-01005-ofi-m. I. Petrovic and A. Ristic Fira acknowledge financial support from the Ministry of Education, Science and Technological Development of Serbia (grants 173046 and 171019). D. Sakata acknowledges financial support from JSPS KAKENHI, grant number JP20K16840.

Author contributions

D.S. developed the integrated application.; D.S. also analysed the results and drafted the manuscript.; M.C.B., D.E., I.K. and S.I have developed the particle transport models.; M.K. developed the models used in the chemical stage.; O.B. developed the biological prediction model, and discussed the validity of the biological model with I.P and A.R.F.; D.S, N.L, V.I. and S.I. contributed to the integration of all features.; This manuscript has been improved and concluded by D.S., S.G., D.E., I.K., T.I. and S.I.; This work is originally proposed by S.I.; All authors reviewed and approved the final manuscript.

Competing Interests

The authors declare no competing interests.

Additional information

Supplementary information is available for this paper at <https://doi.org/10.1038/s41598-020-75982-x>.

Correspondence and requests for materials should be addressed to D.S.

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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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 licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020