Experimental validation in a neutron exposure frame of the MINAS TIRITH for cell damage simulation

Yann Thibaut, Geraldine Gonon, Juan Martinez Guerrero, Michael Petit, Richard Babut, Aurelie Vaurijoux, Gaëtan Gruel, Carmen Villagrasa, S Incerti, Yann Perrot

Experimental validation in a neutron exposure frame of the MINAS TIRITH for cell damage simulation

Y. THIBAUT[1], G. GONON[1], J. S. MARTINEZ[1], M. PETIT[1], R. BABUT[1], A. VAURIJOUX[1], G. GRUEL[1], C. VILLAGRASA[1], S. INCERTI[2], Y. PERROT[1*]

[1] Institut de Radioprotection et de Sûreté Nucléaire (IRSN), PSE-SANTE/SDOS/LDRI, PSE-SANTE/SERAMED/LRAcc, PSE-SANTE/SDOS/LMDN, BP 17, 92262 Fontenay-aux-Roses, France

[2] Université de Bordeaux, CNRS/IN2P3, LP2i, UMR 5797, 33170 Gradignan, France

[*] Corresponding author: yann.perrot@irsn.fr

Abstract

In the domains of medicine and space exploration, refining risk assessment models for protecting healthy tissue from ionizing radiation is crucial. Understanding radiation-induced effects requires biological experimentations at the cellular population level and the cellular scale modeling using Monte Carlo track structure codes.

We present MINAS TIRITH, a tool using Geant4-DNA Monte Carlo-generated databases to study DNA damage distribution at the cell population scale. It introduces a DNA damage location module and propose a method to convert Double-Strand Breaks (DSB) into DNA Damage Response foci. We evaluate damage location precision and DSB-foci conversion parameters.

MINAS TIRITH's accuracy is validated against γ-H2AX foci distribution from cell population exposed to monoenergetic neutron beams (2.5 or 15.1 MeV) under different configurations, yielding mixed radiation fields. Strong agreement between simulation and experimental results was found demonstrating MINAS TIRITH's predictive precision in radiation-induced DNA damage topology.

Additionally, modeling intercellular damage variability within a population subjected to a specific macroscopic dose identifies subpopulations, enhancing realistic fate models. This approach advances our understanding of radiation-induced effects on cellular systems for risk assessment improvement.

State of the Art

Radiation protection aims to minimize exposure to ionizing radiation as much as possible. Although some of exposures are unavoidable in certain contexts such as space exploration or in enclosed areas with high radon concentration, ionizing radiation can have beneficial applications in medicine. Radiation therapy, for example, is a non-invasive treatment that can reduce tumor growth. However, its success depends on a delicate balance between destroying cancer cells and preserving surrounding healthy tissue, known as the therapeutic window [1, 2]. To quantify potentially harmful side effects, risk models have been developed. However, to improve the quality of treatment, it is essential to have a thorough understanding of the effects of ionizing radiation on healthy tissues [3]. This understanding requires fundamental knowledge of the induction of these effects and biological mechanisms they generate [4]. In this context, the DNA molecule has been identified as a critical target of radiation-induced effects [5]. In particular, the DNA double strand break (DSB) has been identified as the critical and more deleterious damage to the cell [4, 6, 7]. Various studies have linked

this type of DNA damage to biological endpoints such as chromosomal aberration [8] or clonogenic death [9] in cell populations that have been irradiated *in vitro*.

Monte-Carlo track structure codes have been pointed out as essential tools to quantify radiation effects and thus improve risk models [10]. Following this idea, modelling tools for calculating radiation-induced DNA damage have been developed [11-14]. Nevertheless, these models are limited to the mean response at the scale of the cell nucleus and therefore are not able to characterize phenomena distributions occurring on a cell population scale. Thus, the description of the damage at the cell population scale is necessary for a better comparison between simulated and experimental results [15], since the correlation between the population response and the variability of imparted energy has been previously highlighted [16]. In this context, the MINAS TIRITH tool has been proposed in order to model the distribution of radiation-induced DNA damage at the scale of a cell population [17]. This tool is based on databases generated by Monte Carlo simulation with the Geant4-DNA tool [18-21] coupled to realistic nucleus geometries [22-24]. The new damage distribution method enabled by MINAS TIRITH allows access to a higher spatial scale in the description of radiation-induced damage compared to the track structure codes typically used. MINAS TIRITH simultaneously considers the variability of the energy imparted in an irradiated population and the variability of the induced damage. To date, the damage distribution method implemented in MINAS TIRITH has only been validated by comparison to Geant4-DNA simulation but not yet experimentally.

Generally, the experimental validation of simulation codes involves comparing simulated DNA $DSBs$ with observation of foci such as γ-H2AX as it has been shown that histone H2AX phosphorylation (γ-H2AX) follows the induction of DNA $DSBs$ [25,26]. This comparison is often made in terms of mean number of DNA $DSBs$ per Gray or per ionizing particle and mean number of γ-H2AX foci per cell for a given absorbed dose or number of ionizing particles [27-29]. However, this comparison may be biaised [30], since the relationship between the number of induced DNA $DSBs$ and the number of γ-H2AX foci observed is not always straightforward [28, 29]. In some studies, correction factors are proposed to convert the induced DNA DSB to observed focus [29, 31]. For instance, the method proposed by Barbieri *et al.* models the γ-H2AX foci signaling based on DNA $DSBs$ clustering and its spatial distribution. On the other hand, Ingram *et al.* propose a code that simulates the γ-H2AX signaling from the DNA $DSBs$ topology taking into account several experimental mechanisms like the microscope resolution or the relative distance between the DSB and the histone [32].

Experimental validation of the simulation codes allows the adjustment of the model parameters and thus a better acuity of the latter. This process is essential in model development [33] as a good agreement between the models and the experiments allows to use the model for predictive purposes. As far as MINAS TIRITH is concerned, its experimental validation in terms of the distribution of radiation-induced damage will ensure its use as a reliable input for developing repair models based on the identification of cell subpopulations linked to a biological effect observed experimentally.

For the validation of MINAS TIRITH, the simulated $DSBs$ are compared to the experimental results obtained in terms of distribution of foci per cell in populations irradiated by monoenergetic neutrons in four different irradiation configurations. To do so, simulated $DSBs$ attributed to each cell are first localized within the cell nucleus and then converted into γ-H2AX foci. The methods developed for damage location and conversion to γ-H2AX foci are presented in the following study. It will also be shown that, for this goal, an accurate dosimetry of the four irradiation configurations is needed as well as a good knowledge of the analysis method of the experimental results.

Materials and Methods

1. Damage location method with the MINAS TIRITH tool

1.1. The damage location data base

The MINAS TIRITH tool [17] is based on two databases. The first database is made of microdosimetric spectra to model the track distribution in the cell population and the second is made of DNA damage spectra to associate a damage yield to each track, considering its microdosimetric characteristics. These databases were built for different monoenergetic particles sampled between 10 keV and 20 MeV for protons, 10 keV and 21 MeV for alphas and 1 keV and 1 MeV for electrons. To constitute the second database, a Geant4-DNA [14] nanodosimetric simulation chain was used, coupled to isochore cell nucleus geometry [24]. This simulation chain [14] allows the calculation of the topology of radiation-induced DNA damage at the nuclear scale. To extend the second database of MINAS TIRITH to damage location, the location of each radiation-induced damage ($DSBs$ and $SSBs$) is recorded. The coordinates of the entry and the stop or exit points (P_{in} and P_{out}) of each track in the nucleus are also extracted (the nucleus being centered on the origin point). The number of simulated particles for each energy bin of the data base was chosen to have enough statistical power: 25,000 for electrons, 1,500 for alphas and between 1,500 and 10,000 for protons depending on the LET. Concerning the irradiation geometry, the particles are emitted from the surface of the nucleus, in an isotropic way.

The data from the damage simulation chain were processed as follows. For each damage (DSB or SSB), the quantities ${}^{//}Damage_{loc}$ and ${}^{\perp}Damage_{loc}$ are defined. P_{Damage} being the position of the damage in the cell nucleus:

$$ {}^{//}Damage_{loc} = \cos\left(\overrightarrow{P_{in}P_{Damage}}, \overrightarrow{P_{in}P_{out}}\right) \frac{\left\|\overrightarrow{P_{in}P_{Damage}}\right\|}{\left\|\overrightarrow{P_{in}P_{out}}\right\|} $$

$$ {}^{\perp}Damage_{loc} = \sin\left(\overrightarrow{P_{in}P_{Damage}}, \overrightarrow{P_{in}P_{out}}\right) \left\|\overrightarrow{P_{in}P_{Damage}}\right\| $$

Thus, ${}^{//}Damage_{loc}$ represents the projection of the damage position onto the line formed by the entry point and the stop (or exit) point of the particle that generated the damage normalized by the chord length of the track. ${}^{\perp}Damage_{loc}$ represents the distance between the damage position and the line formed by the entry point and exit point of the particle that generated the damage. Thanks to these data, spectra in terms of ${}^{//}DSB_{loc}$, ${}^{\perp}DSB_{loc}$, ${}^{//}SSB_{loc}$and ${}^{\perp}SSB_{loc}$ (denoted $F({}^{//}DSB_{loc})$, $F({}^{\perp}DSB_{loc})$, $F({}^{//}SSB_{loc})$ and $F({}^{\perp}SSB_{loc})$) are constructed for each sampled energy of the data base following the same method that those presented in [17] : spectra are spline-smoothed and scaled into 1909 points with ordinates between 0 and 1.

1.2. Determination of the damage location with MINAS TIRITH

Based on the method presented in [17], each distributed track in an irradiated cell nucleus at an absorbed dose D is described with a particle type (p), an initial kinetic energy (E), an entry position (P_{in},), a direction ($\overrightarrow{Dir}$), a chord length (CL) and an imparted energy (ε). Based on all these criteria, a number of damage events is sampled for each track contributing to the total energy deposit. To locate each of these damages, the following method is proposed:

i. The $F\left({}^{//}_{p,E}\widehat{Damage}_{loc}\right)$ spectrum is reconstructed from the database of $F({}^{//}Damage_{loc})$ spectra. This is done by spline interpolating the value ${}^{//}_{p,E}\widehat{Damage}_{loc_k}$

from the values $^{//}Damage_{loc_k}$ for the K ordinates ($K = 1909$) of the database. The same process is followed to reconstruct $F\left(_{p,E}^{\perp}\widehat{Damage}_{loc}\right)$ spectrum.

ii. The two values $_{p,E}^{//}\widehat{Damage}_{loc}$ and $_{p,E}^{\perp}\widehat{Damage}_{loc}$ are calculated by inverse sampling of the $F\left(_{p,E}^{//}\widehat{Damage}_{loc}\right)$ and $F\left(_{p,E}^{\perp}\widehat{Damage}_{loc}\right)$ spectra.

iii. The rotation matrix M_{rot} between unitary vector $\overrightarrow{u_z}$ and the direction of the track $\overrightarrow{Dir}$ is calculated.

iv. A random angle φ_R is sampled on the $[0; 2\pi]$ interval.

v. The damage location P_{Damage} , with $P_{Damage} = \begin{pmatrix} x_{Damage} \\ y_{Damage} \\ z_{Damage} \end{pmatrix}$, $P_{in} = \begin{pmatrix} x_{in} \\ y_{in} \\ z_{in} \end{pmatrix}$, $\overrightarrow{Dir} = \begin{pmatrix} x_{Dir} \\ y_{Dir} \\ z_{Dir} \end{pmatrix}$, is calculated according to the following formula:

$$\begin{pmatrix} x_{Damage} \\ y_{Damage} \\ z_{Damage} \end{pmatrix} = \begin{pmatrix} x_{in} \\ y_{in} \\ z_{in} \end{pmatrix} + {}_{p,E}^{//}\widehat{Damage}_{loc} * CL * \begin{pmatrix} x_{Dir} \\ y_{Dir} \\ z_{Dir} \end{pmatrix} + {}_{p,E}^{\perp}\widehat{Damage}_{loc} * M_{rot} * \begin{pmatrix} \cos(\varphi_R) \\ \sin(\varphi_R) \\ 0 \end{pmatrix}$$

vi. If P_{Damage} is inside the cell nucleus, the location is kept. Else, the steps *ii.* to *v.* are performed again until the condition is accepted.

The method proposed here allows to locate each type of damage associated to each track in all nuclei of the cell population. To validate this method, the reconstruction of the spectra described in steps i. and ii. has been evaluated. For this purpose, the same process of mathematical evaluation of the reconstruction by interpolation as described in [17] was used and the mean absolute percentage error ($MAPE$) was used as indicator [34]. The approach used leads to an overestimation of the error for the reasons described in [17]. The method for determining the saving calculation time was taken from the previous study [17] but including the location of the damage. The time-saving factors of the damage calculation by track was therefore compared between the simulation by a Monte Carlo method using the Geant4-DNA toolkit and the MINAS TIRITH tool including this new damage location module.

2. Cell irradiation experiments performed to validate the MINAS TIRITH tool

The experiments carried out consisted in the irradiation of HUVEC cells grown in monolayers by mono-energetic neutrons of 2.5 or 15.1 MeV on the IRSN's AMANDE facility [35] followed by the study of the topology of the resulting γ-H2AX foci at 30 minutes post-irradiation. Four different irradiation configurations (15.1 MeV-IR1, 15.1 MeV-IR2, 2.5 MeV-IR1 and 2.5 MeV-IR2) were defined depending on the positioning of the SlideFlask on the water phantom (IR1 or IR2, see Figure 1) and the neutron energy (2.5 MeV or 15.1 MeV). The irradiation time was 10 min, as this corresponded to a trade-off between substantial dose delivery and γ-H2AX foci formation kinetics after irradiation. The dosimetry corresponding to each configuration is explained in section 3. The SlideFlask with primary HUVEC cells, filled with pH- and temperature-equilibrated growth medium, was attached to the water phantom using elastic rubber bands. The SlideFlask was positioned as close as possible to the neutron source at a fixed distance of 5.3 cm due to the source-cooling cannula. The SlideFlask was perpendicular to the incident neutron beam.

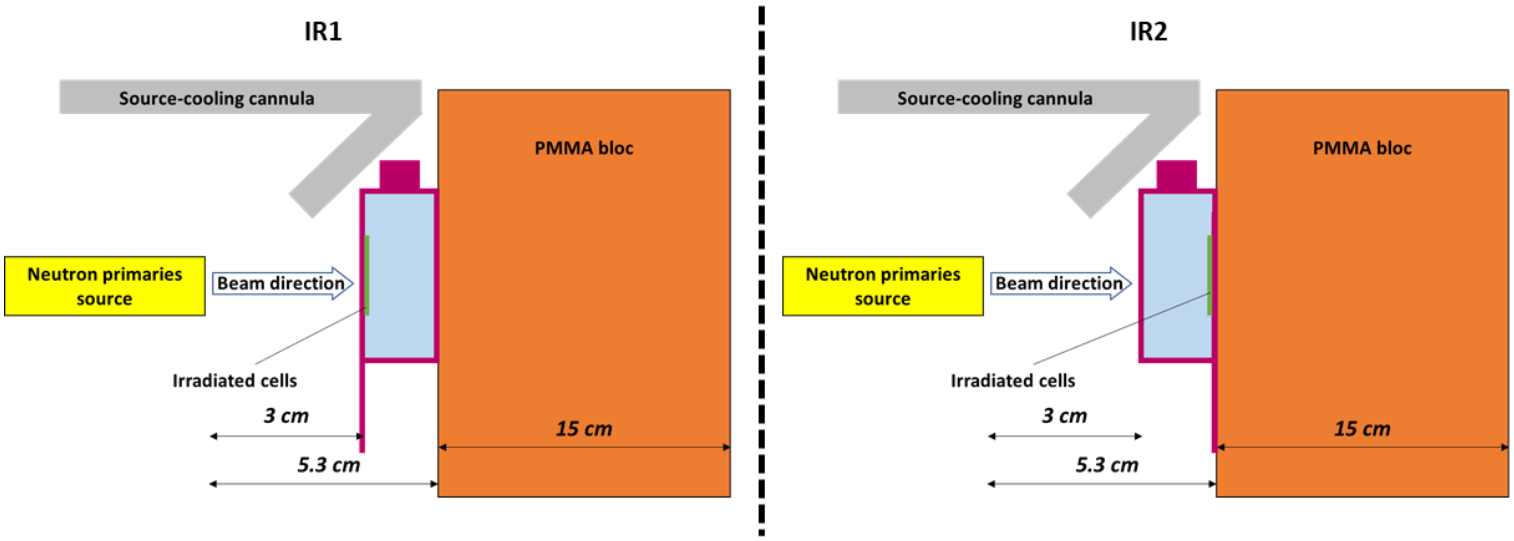

Figure 1: *Schematic description of the experimental set up corresponding to two different irradiation configurations for a given energy of the neutron source (2.5 or 15.1 MeV). On the left, the IR1 configuration with the cells in the Slideflask directly in front of the incoming beam and on the right, the IR2 configuration with cells positioned on the PMMA bloc side.*

After irradiation, the SlideFlask was placed back into the incubator. Thirty minutes post-irradiation, a specific immunofluorescent protocol was performed in order to reveal, *in situ* γ-H2AX foci and the cell nuclei were stained with 4',6-diamidino-2-phenylindole (DAPI).

The images were acquired using an inverted IX83- microscope (Olympus) equipped with a LED light source SpectraX (Lumencor) and a UPLSAPO 100XO oil immersion objective lens (Olympus), then analyzed thanks to a high-throughput imaging software (Olympus, ScanR, v3.3). Detection of cell nuclei and γ-H2AX foci was based on an edge segmentation algorithm and cell nuclei in G_0/G_1 phase of the cell cycle were selected assessed by a "flow cytometry like" analysis [16, 28]. Distributions of γ-H2AX foci formation were examined from at least three identical experiments (replicates) for each irradiation configuration and ~4000 cells analyzed for each sample [27]. Sham-irradiated SlideFlasks were handled in the same way as the test cultures in the irradiation hall, left under the cooling flux, but were not exposed to neutrons.

3. Dosimetry of the experiments

The dosimetry for each of the irradiations was done according to the following protocol:

i. In a first step, AMANDE experimental devices allow a direct time reconstruction of the fluence of primary neutrons emitted in the beam axis at a given distance from the emission source. The fluence values at five meters from the source are measured by a long counter named PLC. The measurement method and reconstruction of the fluence has been detailed in [36]. For each irradiation, the fluence at the entrance of the SlideFlask (three centimeters from the emission source), on the beam axis, integrated over the 10 minutes of irradiation is reconstructed.

ii. In parallel, thanks to a succession of Monte Carlo simulations described in [17], the conversion factors, one for each irradiation configuration, are obtained as shown in Table 1. The scoring volume for the absorbed dose in the population corresponds to the surface experimentally analyzed of dimension 2.08 cm x 0.1 cm by the typical thickness of the monolayer cell population (0.2 µm), centered on the SlideFlask.

Table 1: *Conversion factor expressed in 10^{-12} Gray per primary neutron at the fluence reconstruction distance from the source (3 cm).*

Configurations	Conversion factor (10^{-12} Gy/primary neutron)
15.1 MeV - IR1	37.51
15.1 MeV - IR2	20.13
2.5 MeV - IR1	17.47
2.5 MeV - IR2	8.98

iii. The fluence, the energy and the incidence of the beam of primary neutrons are homogenous on the area of the analyzed cells in all the irradiation configurations. For each configuration, three independent irradiations were performed, thus, the mean dose delivered to the analyzed cells for the twelve irradiations (4 configurations x 3 replicas) according to the fluence, reconstructed from PLC counting, during each irradiation is presented in Table 2.

Table 2: Measured fluences and corresponding calculated absorbed dose in the experiments. *In the middle column, the fluences reconstructed at 3 cm, from PLC counting, from the beam for each irradiation expressed in 10^9 primary neutrons per cm^2 with the associated measurement uncertainty. In the right column, the absorbed dose delivered to the analyzed cells of the irradiated population.*

Configurations	Fluence (10^9 neutron/cm²)	Absorbed dose (mGy)
15.1 MeV - IR1	2.04 +/- 0.27	76.52 +/- 10.13
	3.18 +/- 0.44	119.28 +/- 16.50
	3.05 +/- 0.42	114.41 +/- 15.75
15.1 MeV - IR2	2.07 +/- 0.28	41.67 +/- 5.64
	3.29 +/- 0.45	66.24 +/- 9.06
	3.04 +/- 0.42	61.20 +/- 8.46
2.5 MeV - IR1	4.10 +/- 0.53	71.65 +/- 9.26
	4.01 +/- 0.52	70.07 +/- 9.09
	4.16 +/- 0.55	72.69 +/- 9.61
2.5 MeV - IR2	4.06 +/- 0.53	36.48 +/- 4.76
	4.07 +/- 0.53	36.57 +/- 4.76
	4.16 +/- 0.55	37.38 +/- 4.94

4. Damage modeling

In order to validate the MINAS TIRITH tool, the topology of the radio-induced DNA damage for the different irradiation configurations has been simulated. To do this, a simulation of each of the irradiation conditions was made for the experimentally delivered absorbed dose. As input to these simulations, the energy and incidence spectra of secondary charged particles at the surface of the cell nuclear volumes in the four irradiation configurations were used. These spectra were acquired according to the method described in 2.ii and in [17]. Then, the topology of the cell-by-cell radiation-induced DNA damage for each irradiated population (4 configurations x 3 replicas = 12 irradiated populations at absorbed doses given in Table 2) was simulated. This modeling was limited to the DNA DSB distribution, as the DNA SSB distribution is not observed in the experiments. As an example,

Figure 2 shows the distribution of tracks and DSBs, visualized in 2D using a visualization tool developed on Matlab, for 12 cells for the first replica of the 2.5 MeV-IR1 configuration.

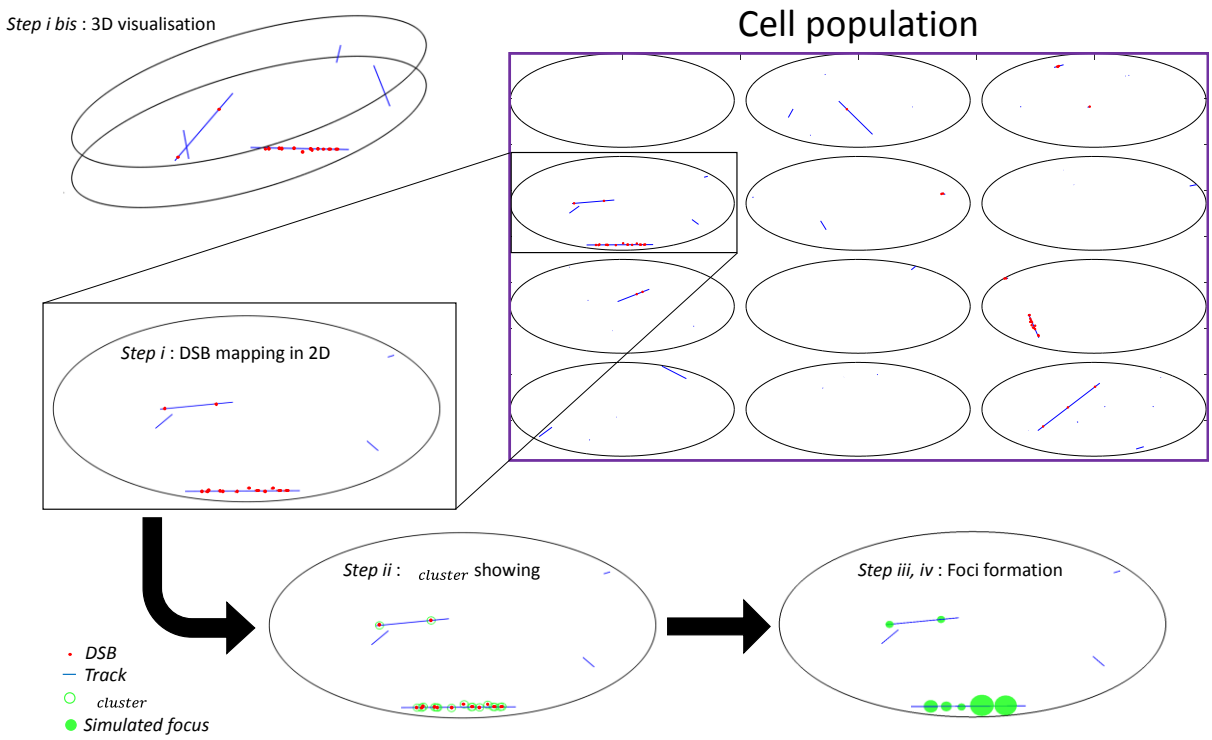

Figure 2: *Example of damage calculated at population level and visualized for 12 cells out of 15,0000 in two dimensions. Focusing on one cell, we illustrate the method of forming simulated foci from DSB topologies.*

In order to have a more direct comparison of the experimental and simulated results, the simulated DSB topology has to be converted into γ-H2AX foci topology. However, in sham-treated samples, a significant number of cells in the G_0/G_1 phase of the cell cycle harbored foci. Thus, in order to include the γ-H2AX foci noise in the simulation, the following protocol was applied for each simulated configuration and each step is illustrated in Figure 2:

i. The DSB topology is collapsed in 2D, in the plane perpendicular to the beam axis. This allows to reproduce the observation configuration in which the experimental images are acquired.

ii. Each DSB forms a circular γ-H2AX focus with a diameter D_{focus} initially equal to a chosen parameter $D_{cluster}$ and a barycenter located at the DSB location. The assumption is made that all induced DSB are signaled at the time of experimental observation of the γ-H2AX foci topology. This assumption is made because the experimental data were acquired at the post-irradiation time on which the maximum of signal is observed, i.e. 30 minutes post-irradiation [28].

iii. If two foci partially overlap or intersect, the two foci are clustered into a single focus with a barycenter juxtaposed to the barycenter of all the DSB that composed each of the two foci. This new focus is also considered as circular, with a radius equal to the distance between the focus barycenter and the most far DSB location added to $\frac{D_{cluster}}{2}$. This operation is repeated until all foci within a cell are distinct.

iv. The area of each focus is estimated to be equal to $\pi \left(\frac{D_{focus}}{2}\right)^2$.

v. Finally, A background number of foci generated by random sampling based on the distribution of (pooled) sham data shown in Figure 3 is randomly added to each cell in the simulation. This step aims at simulating the γ-H2AX foci background that cannot be differentiated from the radio-induced foci experimentally.

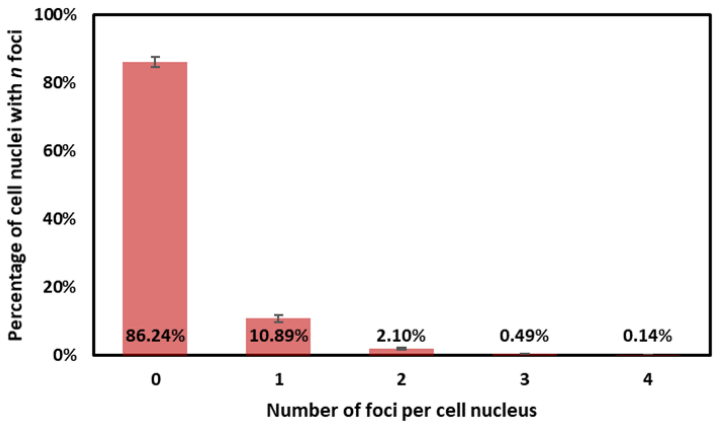

Figure 3: *Distribution of the number of γ-H2AX foci per cell for non-irradiated populations. Error bars represent the mean error on each bin of the distribution induced by variability between pooled shams.*

The value of the clustering distance $D_{cluster}$ has a strong influence on the distributions of the number of foci per cell in the cell population. A study of the impact of this parameter was carried out over a range of values from 0.2 to 0.8 micrometers. As the form of simulated foci distributions is linearly dependant of the value of the $D_{cluster}$ parameter, only the results obtained with the limit values (0.2 and 0.8 micrometers) and the selected value (0.4 micrometers) are presented. In the results section, a comparison between experimental and modeled results is presented for the four irradiation configurations.

Results

1. Evaluation of the reconstruction of spectra by interpolation

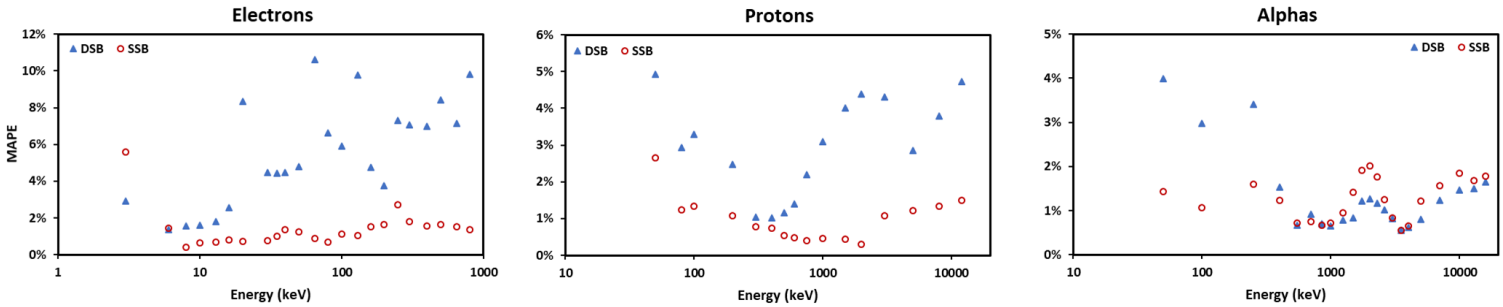

Figure 4: *MAPE indicator values when reconstructing* $F\left(\overset{//}{}_{p,E_l}\widehat{DSB_{loc}}\right)$ *(blue) and* $F\left(\overset{//}{}_{p,E_l}\widehat{SSB_{loc}}\right)$ *(red) spectra for each* p *(electrons, protons, and alpha particles) and* E_l *(x-axis) in the database. Note: scales of each graph are different.*

Evaluation of the interpolation reconstruction error of the $F\left(\overset{//}{}_{p,E_l}\widehat{DSB_{loc}}\right)$ and $F\left(\overset{//}{}_{p,E_l}\widehat{SSB_{loc}}\right)$ spectra shown in Figure 4 was performed using the $MAPE$ indicator. For protons and alpha particles, the indicator is less than 5% over the entire energy range for SSB and DSB. For electrons, the indicator is less than 5% over the entire energy range for SSB, except for the $F\left(\overset{//}{}_{e^-,3\ keV}\widehat{SSB_{loc}}\right)$

spectrum for which it is around 6%, and less than 10% for DSB, except for the $F\left({}_{e^-,65\ keV}^{//}\widehat{DSB_{loc}}\right)$ spectrum for which it is around 11%. It is important to remember that the method used to evaluate the quality of the reconstruction by interpolation of the spectra represents a conservative value of the error for the whole energy range, as explained in [17].

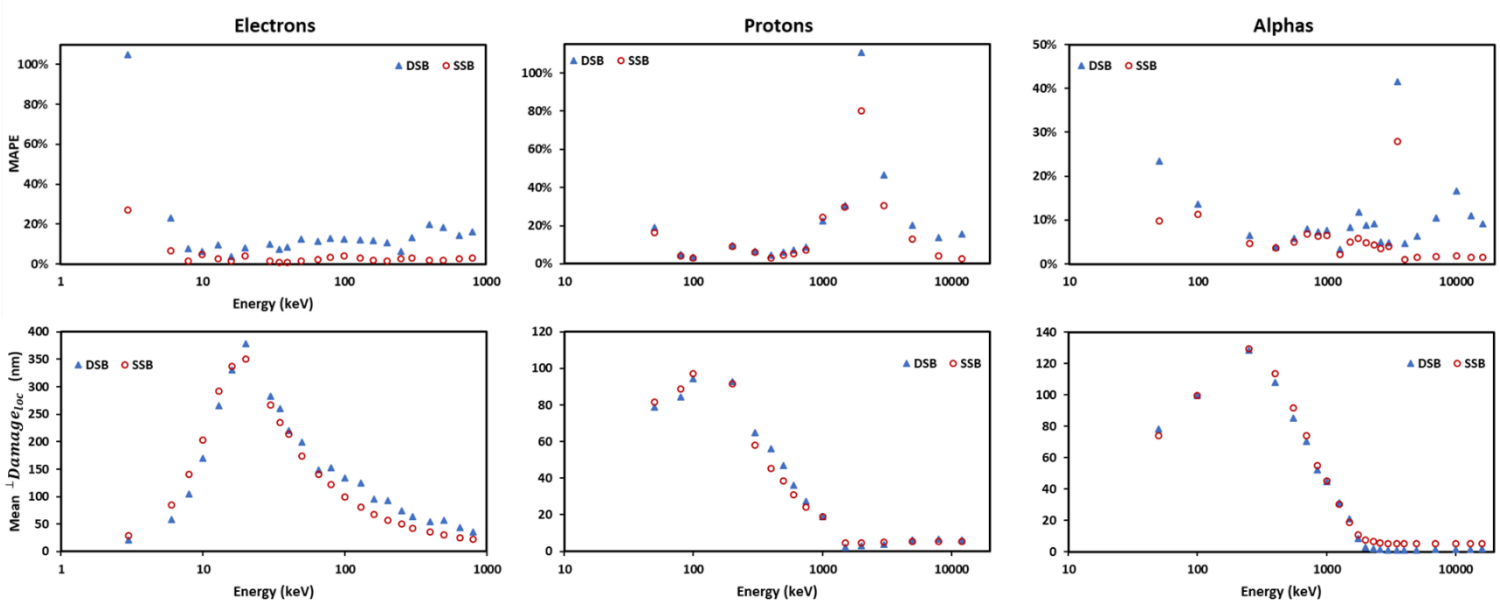

Figure 5: *On the first line, $MAPE$ indicator values when reconstructing $F\left({}_{p,E_l}^{\perp}\widehat{DSB_{loc}}\right)$ (blue) and $F\left({}_{p,E_l}^{\perp}\widehat{SSB_{loc}}\right)$ (red) spectra for each p (electrons, protons, and alpha particles) and E_l (x-axis) in the database. On the second line, the mean values of ${}^{\perp}DSB_{loc}$ and ${}^{\perp}SSB_{loc}$ on the whole energy range. Note: scales of each graph are different.*

For electrons, a $MAPE$ indicator around 5% is observed for $F\left({}_{e^-,E_l}^{\perp}\widehat{SSB_{loc}}\right)$ on the entire energy range, except for $F\left({}_{e^-,3\ keV}^{\perp}\widehat{SSB_{loc}}\right)$ spectrum for which it is around 25%. The same tendency is observed for $F\left({}_{e^-,E_l}^{\perp}\widehat{DSB_{loc}}\right)$ with a $MAPE$ indicator lower than 15% on the whole range and a very high value (up to 100%) for $F\left({}_{e^-,3\ keV}^{\perp}\widehat{DSB_{loc}}\right)$. For protons and alpha particles the $MAPE$ indicator value is globally less than 20% overall E_l range, except for some specific points : $F\left({}_{p^+,1\ to\ 3\ MeV}^{\perp}\widehat{SSB_{loc}}\right)$, $F\left({}_{p^+,1\ to\ 3\ MeV}^{\perp}\widehat{DSB_{loc}}\right)$, $F\left({}_{\alpha,3.5\ MeV}^{\perp}\widehat{SSB_{loc}}\right)$ and $F\left({}_{\alpha,3.5\ MeV}^{\perp}\widehat{DSB_{loc}}\right)$. All the points for which the $MAPE$ indicator value is high are around the inflexion point of the mean of ${}^{\perp}Damage_{loc}$ curves. For all the points with high $MAPE$ highlighted in this paragraph, the absolute value of the ${}^{\perp}Damage_{loc}$ remains below 20 nanometers, as it can be seen in the second row of Figure 5.

2. Time saving factor allowed by MINAS TIRITH in comparison to Geant4-DNA damage simulation

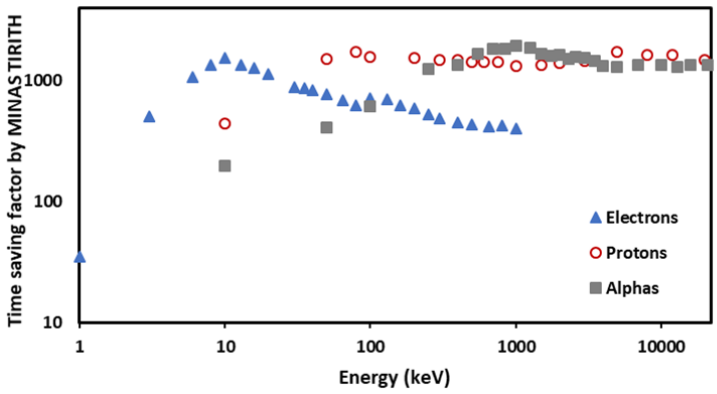

Figure 6: *Time saving by MINAS TIRITH including damage location compared to the Geant4-DNA based simulation chain over the whole energy range and for each particle type.*

Figure 6 shows the time saving factor for damage calculation with location for a single track with the MINAS TIRITH tool compared to the simulation time with the complete Monte Carlo simulation chain developed at the IRSN. MINAS TIRITH offers a time saving factor upper than 1,000 for most of the energy range for protons and alphas. For electrons, this time saving factor is between 100 and 1,000 for most of the energy range (upper than 1,000 around 10 keV and lower than 100 at 1 keV).

3. Influence of the clustering distance on foci number

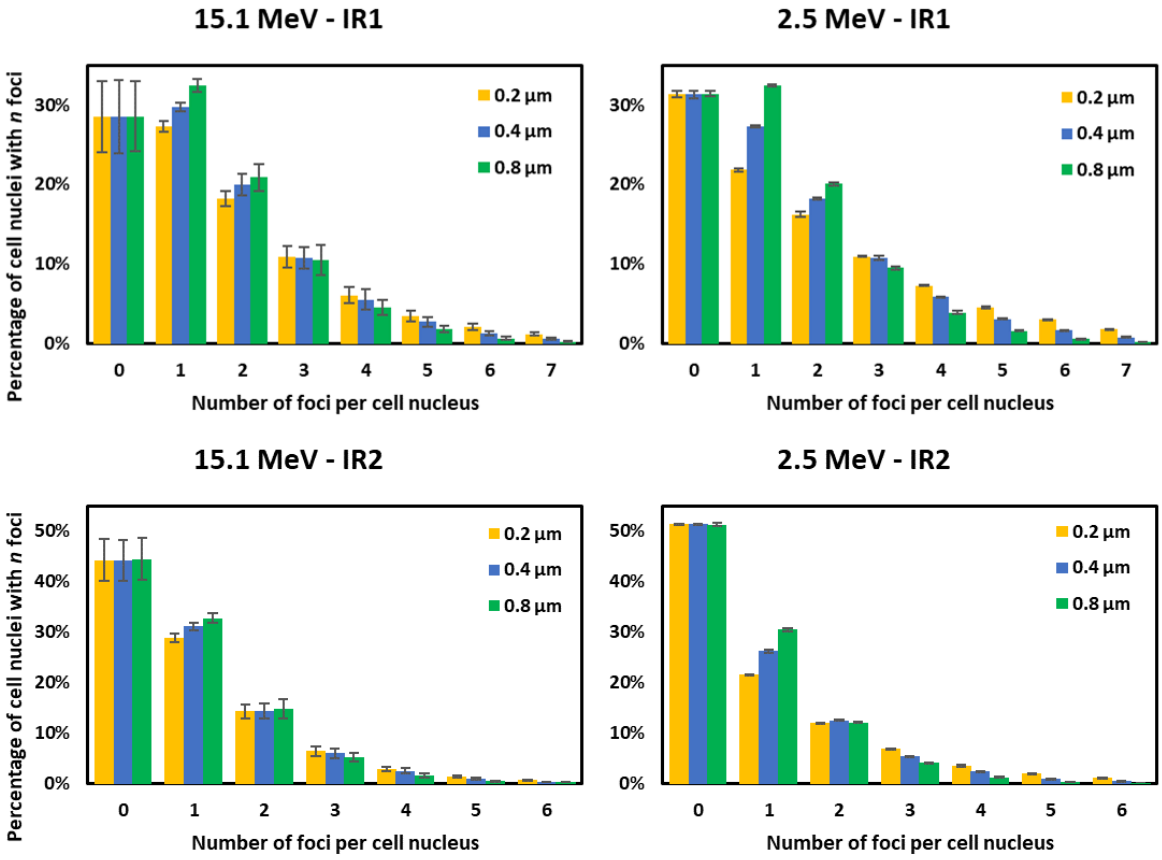

Figure 7: *Proportion of cells in the population (y-axis) with a given number of foci per cell (x-axis) for different clustering distances (0.2 µm in yellow, 0.4 µm in blue and 0.8 µm in green) for*

the four irradiation configurations. Error bars represent the mean error on each bin of the distribution induced by variability between the three simulated doses (replicas).

The distribution of cell nuclei is plotted as a function of the number of simulated foci they contain for different clustering distances in Figure 7. Each bin of these distributions is obtained taking the mean of the three simulated replicas of a given irradiation configuration. The associated uncertainties plotted represent the error on the mean. As expected, the proportion of cells without foci does not depend on the clustering distance. On the other hand, for each irradiation configuration, the proportion of cells with one or two foci increases with clustering distance. Consequently, the proportion of cells with three or more foci decreases with the clustering distance. From a more global point of view, as the clustering distance increases, the mean number of foci per cell decreases. This decrease results in a change in the shape of the distribution of the number of foci per cell in the population. Furthermore, it is important to note that with respect to the uncertainties presented in Figure 7, the differences in the distributions of the number of foci are significant for the 2.5 MeV-IR1 and 2.5 MeV-IR2 configurations. This is not the case for the differences observed in the distributions of the 15.1 MeV-IR1 and 15.1 MeV-IR2 configurations.

4. Comparison of foci distribution and areas between experimental and simulated data

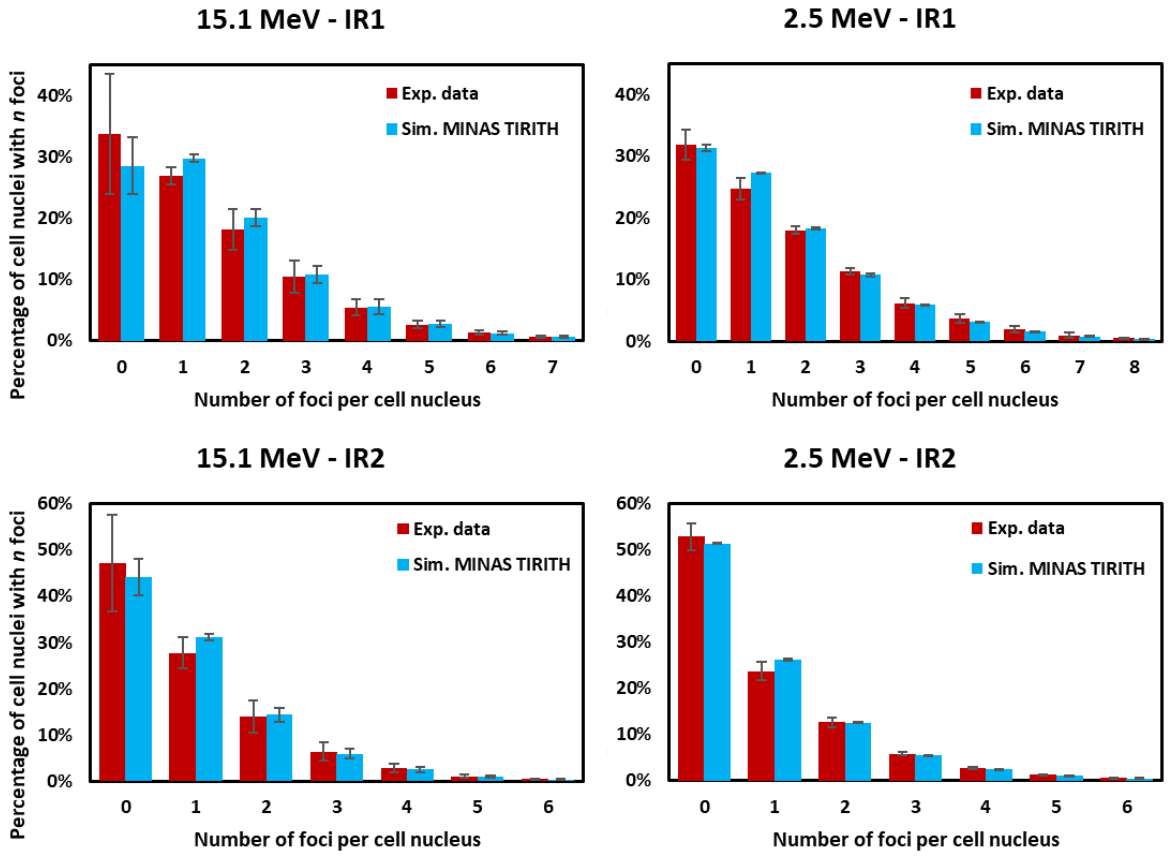

Figure 8: *Comparison of cell proportion in the population (y-axis) with a given number of foci per cell (x-axis) between γ-H2AX experimental data (in red) and simulation data obtained using MINAS TIRITH tool (in blue), with clustering distance equal to 0.4 µm, in the four irradiation configurations.*

The experimental and simulated distributions of foci per cell for the four irradiation configurations are compared in Figure 8. Each bin of these distributions is obtained taking the mean of the three irradiation replicas either performed or simulated for a given irradiation configuration. The associated uncertainties represent the errors of the mean. These uncertainties for the experimental

data are larger at 15.1 MeV than at 2.5 MeV for either the IR1 or IR2 configurations. This is also the case for the simulated data. For the four configurations, the uncertainties are smaller in simulation than in the experimental data. Moreover, an excess of cells of cells with only one focus is observed in simulation compared to the experimental data. In the same way, we observed a lower proportion of cells having no foci in the simulated data than in the experimental data. The proportions of simulated cells with two or more foci are very close to the experimental data in the four irradiation configurations.

Table 3: *Mean foci areas for experimental (Exp.) and simulated with clustering distance equal to 0.4 µm (Sim.) data. Mean errors for the simulated data are lower than 0.01 µm².*

Configuration	Exp. mean areas of foci (µm²)	Sim. mean areas of foci (µm²)
15.1 MeV - IR1	0.77 ± 0.13	0.26
15.1 MeV - IR2	0.69 ± 0.14	0.20
2.5 MeV - IR1	0.54 ± 0.20	0.25
2.5 MeV - IR2	0.53 ± 0.22	0.24

Table 3 presents the mean foci areas for the simulated and the experimental data. The mean foci areas in the experimental data are significantly higher than the foci areas obtained with the simulation in all four irradiation configurations. Those are the result proximity of the simulated DSBs and the chosen clustering distance to convert DSBs into simulated foci. For the configurations at 2.5 MeV, the mean foci areas for both IR1 and IR2 configurations are very close, and this is well reproduced in the simulation even if the absolute value is lower. For the 15.1 MeV configurations, the mean foci area is larger in IR1 than in IR2 in both, simulation and experimental data. Nevertheless, this phenomenon is only a tendency for the experimental data because the difference is not statistically significant.

Discussion

The validation of damage distribution simulation in MINAS TIRITH requires the addition of a damage location module to the tool. In this study, we presented the method developed to this end that is based on the generation of DNA damage spectra on a cell nucleus using the complete Monte Carlo simulation chain [14] based on Geant4-DNA. Prior to validating the location of distributed DNA damage events in MINAS TIRITH by comparing with experimental data, the accuracy of the spectra reconstruction through interpolation over the energy and particle range of the tool was assessed using the $MAPE$ indicator. The position of each damage was analyzed by decomposing it into a parallel and a perpendicular component to the track that caused the damage. Our results demonstrate that the reconstruction of $F\left(^{//}\widehat{Damage}_{loc}\right)$ spectra is highly reliable with a $MAPE$ indicator of less than 10%, as shown in Figure 4, across the entire range of the MINAS TIRITH tool. This indicates that the position of the damage in the parallel component to the track can be sampled from a reconstructed spectrum with a high degree of reliability. However, the reconstruction of spectra for the perpendicular component to the track is less reliable, as demonstrated in Figure 5, with certain energies for each particle, where the $MAPE$ indicator exceeds 100% for electrons and protons, and 40% for alphas. We attribute this discrepancy in the reconstruction of $F\left(^{\perp}\widehat{Damage}_{loc}\right)$ spectra at specific energies for each of these particles to the inflection point of the interpolation curves over the energy range, which leads to large values of the $MAPE$ indicator for each of them, induced by the evaluation method of spectrum reconstruction. The removal of these data points from the analysis of their reconstructions strongly impacts the shape of the interpolation. Moreover, the value of the perpendicular component (Figure 4) is small compared to that of the parallel component modulated by the track chord length which is of the order of magnitude of the size of the

endothelial nucleus. Overall, regarding the accuracy of the reconstruction by interpolation method, the damage location with MINAS TIRITH is consistent with the damage location obtained at the scale of a single cell nucleus using the simulation chain based on Geant4-DNA.

The modeling of damage topology with Geant4-DNA has never been directly validated experimentally in terms of spatial position of the damage. This is primarily due to the limitation that this observable cannot be measured by direct observation through biological experiments. Pulsed field electrophoresis enables the quantification of the size of genome segments resulting from irradiation and allows for the validation of the relative distance between two damage events, but not their absolute spatial distribution within the cell nucleus. The distribution of γ-H2AX foci can be used to infer the spatial position of the Double Strand Breaks damage, even some limitation must be taken into account [31]. However, this technique does not allow for the direct observation of the position of the damage, but instead detects the phosphorylation of the serine 139 of the histone H2AX using immunoassays which spreads up to 1-2 Mbp away around the DSB. This signal is detected via microscopic imaging and is expressed as a bright spot on an image. However, this spot is not punctual, leading to a detection limit in terms of resolution due to the proximity of $DSBs$. Two $DSBs$ that are too close spatially will be detected as a single focus in the analysis of experimental results, mostly in 2D images. Moreover, experimental biases such as the background in terms of γ-H2AX signaling or kinetics of γ-H2AX appearance and disappearance according to the quality of ionizing particles further complicate the interpretation of the detected signal, which may not be entirely attributable to radiation-induced $DSBs$. In conclusion, the validation of the spatial position of damage for modeling damage topology with any Monte Carlo track structure code remains a challenging task due to the limitations of the available experimental techniques. Other work focusing on the relative position of foci within cells will be the subject of a dedicated publication.

Aiming to achieve a high level of agreement with biological observations and experimentally validate the location and number of damages yielded by MINAS TIRITH, the simulated results were post-processed on a similar way to those presented in Gonon *et al.* [28] Barbieri *et al.* [29]. A new fitting parameter $D_{cluster}$ was introduced to simulate the experimental detection limit of distinct $DSBs$ related to non-punctual signaling by a γ-H2AX focus. The impact of this parameter on the simulated results is analyzed in Figure 7. The results indicate that the $D_{cluster}$ distance strongly influences the simulated outcome. Specifically, increasing the $D_{cluster}$ distance resulted in a decreased mean number of foci per cell, which directly affected the distribution of foci per cell increasing the proportion of cells with only one nucleus. To validate the MINAS TIRITH tool, the $D_{cluster}$ distance was set at 0.4 micrometers. This distance was chosen based on the experimental data but remains in agreement with what is reported in the literature [29].

Comparison of experimental and simulated data in four different irradiation configurations showed a very good agreement between the two datasets when the $D_{cluster}$ distance is set to 0.4 µm. Simulated foci distributions were found to be very close to the experimentally observed distributions. However, the simulation systematically exhibited a lower proportion of cells with no foci and inversely, a higher proportion of cells with a single focus, which could be attributed to the possibility that cells with low damage (resulting in a single focus), may have repaired their DNA prior to experimental observation. Additionally, the uncertainty around the simulated distributions was observed to be larger at 15.1 MeV than at 2.5 MeV. This could be due to the heterogeneity of the delivered absorbed doses between the three irradiation replicas of each configuration at 15.1 MeV (Table 2). Furthermore, the larger number of cells analyzed in the simulation (15,000 cell nuclei) than in the experiments (between 3000 and 15,000 cell nuclei) resulted in reduced statistical

uncertainties. Overall, the good agreement between the simulated and experimental data in terms of foci distribution validates the damage location method developed in MINAS TIRITH.

This study provides insights into the accuracy of assumptions made in defining foci in modeling, specifically with regards to the comparison of the mean area (Table 3) of experimentally observed and simulated foci. Our findings indicate that the assumptions made to define foci in modeling are too restrictive as the areas of the simulated foci are smaller. This could be mainly due to the difficulty of modeling the non-uniform intensity profile of each focus observed experimentally. As a result, two foci that partly overlap can be differentiated in experimental analysis using specialized software. Despite this limitation, our simulation faithfully reproduces the experimental observation that the average foci area is higher in the 15.1 MeV-IR1 configuration compared to the 15.1 MeV-IR2 configuration. Furthermore, the average foci area between the 2.5 MeV-IR1 configuration and the 2.5 MeV-IR2 configuration is almost equal in both simulation and experiment. This suggests that the clustering of $DSBs$ in the model is done with a realistic distance $D_{cluster}$ that joins the damage in larger foci, as observed during experiments. This study highlights the importance of carefully considering the assumptions made when modeling foci and the limitations of such procedure. To reach a better comparison between simulated and experimental results, the code proposed by Ingram *et al.* could be used [32].

Importantly, MINAS TIRITH provides a significant time saving factor, with simulation times approximately 1000-fold faster than those based on the Geant4-DNA track structure code, as presented in Figure 6. This significant reduction in simulation time allows for modeling at a higher scale and provides a greater understanding of the mechanisms of DNA damage induction. The ability to model at this higher scale has important implications for the development of new models of cell fate, considering input data that is closer to reality. Such models will provide more accurate predictions of cellular responses to radiation, which is critical for applications such as cancer therapy and radiation protection. Some of these models, such as the Microdosimetric Kinetic Model [37, 38] or the Local Effect Model [39], take dosimetric distributions or damage calculation as input data. It is interesting to note that the scale on which these models operate is of the same order of magnitude as the dimensions on which the formation of foci is calculated in this study. This makes MINAS TIRITH a good candidate for feeding this type of model.

Conclusion

Our study aimed to assess the ability of MINAS TIRITH to accurately model radiation-induced DNA damage distribution at the scale of a cell population. Our results demonstrate that MINAS TIRITH faithfully reproduces the variability of dose distribution and $DSBs$ occurrence, as evidenced by the comparison of experimental and simulated data. The experimental data taken for the validation are the foci distributions resulting from monoenergetic neutron irradiations on endothelial cell populations. The dosimetric characterization of the irradiation is also presented. To perform the validation, a new parameter for the comparison was introduced to get a simulated result closer than the one observed experimentally. This parameter represents the minimum distance between two DSB allowed by microscopic observation. In summary, our findings demonstrate that MINAS TIRITH is a valuable tool for modeling radiation-induced damage at the scale of a cell population, providing significant timesaving and enabling the development of models of cell fate considering the population behavior. This work represents an important step towards improving the accuracy and validity of models used in radiation research and related applications. To expand upon these findings, another article in preparation will focus on the biological effect of the neutrons irradiations taking into account the complexities of mixed-field radiation that enabled MINAS TIRITH to be validated. In

this new study, MINAS TIRITH will be used to gain a deeper understanding of the experimental results analyzed to discuss the different biological effects induced in the different irradiation configurations.

Acknowledgements

This work was partially funded by the French Space Agency (CNES) as part of the ModEll project (grant DAR n°2021/4800001139) and by the European Spatial Agency (ESA) as part of the BioRad III project (grant DAR n°4000132935/21/NL/CRS).

References

[1] **Schneider, U.; *et al.*** Comparative Risk Assessment of Secondary Cancer Incidence after Treatment of Hodgkin's Disease with Photon and Proton Radiation. *Rad. Res.* 2000, 154 (4), 382–388. https://doi.org/10.1667/0033-7587(2000)154[0382:CRAOSC]2.0.CO;2

[2] **Brahme, A.** Optimized radiation therapy based on radiobiological objectives. *Semin. Radiat. Oncol.* 1999, 9 (1), 35–47. https://doi.org/10.1016/S1053-4296(99)80053-8

[3] **Daşu, A.; *et al*.** The use of risk estimation models for the induction of secondary cancers following radiotherapy. *Acta Oncol.* 2005, 44 (4), 339–347. https://doi.org/10.1080/02841860510029833

[4] **Rothkamm, K.; Lobrich, M.** Misrepair of radiation-induced DNA double-strand breaks and its relevance for tumorigenesis and cancer treatment. *Int. J. Oncol.* 2002, 21 (2), 433–440. https://doi.org/10.3892/ijo.21.2.433

[5] **Lomax, M.E.; *et al*.** Biological Consequences of Radiation-induced DNA Damage: Relevance to Radiotherapy. *Clin. Oncol.* 2013, 25 (10), 578–585. https://doi.org/10.1016/j.clon.2013.06.007

[6] **Olive, P.** The Role of DNA Single- and Double-Strand Breaks in Cell Killing by Ionizing Radiation. *Rad. Res.* 1998, 150 (5), S42–S51. https://doi.org/10.2307/3579807

[7] **Goodhead, D.T.** Initial Events in the Cellular Effects of Ionizing Radiations: Clustered Damage in DNA. *Int. J. Radiat. Biol.* 1994, 67 (1), 7–17. https://doi.org/10.1080/09553009414550021

[8] **Iliakis, G.; *et al*.** Mechanisms of DNA double strand break repair and chromosome aberration formation. *Cytogenet. Genome Res.* 2004, 104 (1–4), 14–20. https://doi.org/10.1159/000077461

[9] **Hamada, N.; *et al*.** LET-Dependent Survival of Irradiated Normal Human Fibroblasts and Their Descendents. *Rad. Res.* 2006, 166 (1), 24–30. https://doi.org/10.1667/RR3579.1

[10] **El Naqa, I.; *et al*.** Monte Carlo role in radiobiological modelling of radiotherapy outcomes. *Phys. Med. Biol.* 2012, 57 (11), R75–R97. https://doi.org/10.1088/0031-9155/57/11/R75

[11] **Nikjoo, H.; *et al.*** Computational Approach for Determining the Spectrum of DNA Damage Induced by Ionizing Radiation. *Rad. Res.* 2001, 156 (5), 577–583. https://doi.org/10.1667/0033-7587(2001)156[0577:cafdts]2.0.co;2

[12] **Friedland, W.; *et al.*** Simulation of DNA Damage after Proton Irradiation. *Rad. Res.* 2003, 159 (3), 401–410. https://doi.org/10.1667/0033-7587(2003)159[0401:SODDAP]2.0.CO;2

[13] **Lampe, N.; *et al.*** Mechanistic DNA damage simulations in Geant4-DNA Part 1: A parameter study in a simplified geometry. *Phys. Med.* 2018, 48, 135–145. https://doi.org/10.1016/j.ejmp.2018.02.011

[14] **Meylan, S.; *et al.*** Simulation of early DNA damage after the irradiation of a fibroblast cell nucleus using Geant4-DNA. *Sci. Rep.* 2017, 7, 11923. https://doi.org/10.1038/s41598-017-11851-4

[15] **Chatzipapas, K.P.; *et al.*** Ionizing Radiation and Complex DNA Damage: Quantifying the Radiobiological Damage Using Monte Carlo Simulations. *Cancers* 2020, 12 (4), 799. https://doi.org/10.3390/cancers12040799

[16] **Gruel, G.; *et al.*** Cell to Cell Variability of Radiation-Induced Foci: Relation between Observed Damage and Energy Deposition. *Plos One* 2016, 11 (1), e0145786. https://doi.org/10.1371/journal.pone.0145786

[17] **Thibaut, Y.; *et al*.** MINAS TIRITH: a new tool for simulating radiation-induced DNA damage at the cell population level. *Phys. Med. Biol*. 2023, 68 (3), 034002. https://doi.org/10.1088/1361-6560/acb196

[18] **Incerti, S.; *et al.*** The Geant4-DNA project. *Int. J. Model. Simul. Sci. Comput.* 2010, 1 (2), 157–178. https://doi.org/10.1142/S1793962310000122

[19] **Incerti, S.; *et al.*** Comparison of Geant4 very low energy cross section models with experimental data in water. *Med. Phys.* 2010, 37 (9), 4692–4708. https://doi.org/10.1118/1.3476457

[20] **Bernal, M.; *et al*.** Track structure modeling in liquid water: A review of the Geant4-DNAvery low energy extension of the Geant4 Monte Carlo simulation toolkit. *Phys. Med*. 2015, 31 (8), 861–874. https://doi.org/10.1016/j.ejmp.2015.10.087

[21] **Incerti, S.; *et al*.** Geant4-DNA example applications for track structure simulations in liquid water: A report from the Geant4-DNA Project. *Med. Phys*. 2018, 45 (8), e722–e739. https://doi.org/10.1002/mp.13048

[22] **Meylan, S.; *et al.*** Geant4-DNA simulations using complex DNA geometries generated by the DnaFabric tool. *Comput. Phys. Commun.* 2016, 204, 159–169. https://doi.org/10.1016/j.cpc.2016.02.019

[23] **Tang, N.; *et al.*** Influence of chromatin compaction on simulated early radiation-induced DNA damage using Geant4-DNA. *Med. Phys.* 2019, 46 (3), 1501–1511. https://doi.org/10.1002/mp.13405

[24] **Thibaut, Y.; *et al.*** Nanodosimetric Calculations of Radiation-Induced DNA Damage in a New Nucleus Geometrical Model Based on the Isochore Theory**.** *Int. J. Mol. Sci.* 2022, 22, 3770. https://doi.org/10.3390/ijms23073770

[25] **Rogaku, E.; *et al.*** DNA Double-stranded Breaks Induce Histone H2AX Phosphorylation on Serine 139**.** *J. Biol. Chem.* 1998, 273 (10), 5858–5868. https://doi.org/10.1074/jbc.273.10.5858

[26] **Kinner, A.; *et al.*** γ-H2AX in recognition and signaling of DNA double-strand breaks in the context of chromatin**.** *Nucleic Acids Res.* 2008, 36 (17), 5678–5694. https://doi.org/10.1093/nar/gkn550

[27] **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.* 2019, 20 (24), 6204. https://doi.org/10.3390/ijms20246204

[28] **Gonon, G.; *et al.*** From Energy Deposition of Ionizing Radiation to Cell Damage Signaling: Benchmarking Simulations by Measured Yields of Initial DNA Damage after Ion Microbeam Irradiation. *Rad. Res.* 2019, 191 (6), 566–584. https://doi.org/10.1667/RR15312.1

[29] **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.* 2019, 9, 14019. https://doi.org/10.1038/s41598-019-50408-5

[30] **Penninckx, S.; *et al.*** Quantification of radiation-induced DNA double strand break repair foci to evaluate and predict biological responses to ionizing radiation**.** *NAR Cancer.* 2021, 3 (4), zcab046.
https://doi.org/10.1093/narcan/zcab046

[31] **Rabus, H.; *et al.*** Investigation into the probability for miscounting in foci-based assays. *Radiat. Prot. Dosimetry* 2018, 83 (1–2), 126–131.
https://doi.org/10.1093/rpd/ncy251

[32] **Ingram, S.; *et al.*** A computational approach to quantifying miscounting of radiation-induced double-strand break immunofluorescent foci. *Commun. Biol.* 2022, 5, 700.
https://doi.org/10.1038/s42003-022-03585-5

[33] **Arce, P.; *et al.*** Report on G4-Med, a Geant4 benchmarking system for medical physics applications developed by the Geant4 Medical Simulation Benchmarking Group. *Med. Phys.* 2020, 48 (1), 19–56.
https://doi.org/10.1002/mp.14226

[34] **De Myttenaere, A.; *et al*.** Mean Absolute Percentage Error for regression models. *Neurocomputing* 2016, 192, 38–48.
https://doi.org/10.1016/j.neucom.2015.12.114

[35] **Gressier, V.; *et al.*** AMANDE: a new facility for monoenergetic neutron fields production between 2 keV and 20 MeV. *Radiat. Prot. Dosim.* 2004, 110 (1–4), 49–52.
https://doi.org/10.1093/rpd/nch185

[36] **Gressier, V.; *et al.*** Characterization of a measurement reference standard and neutron fluence determination method in IRSN monoenergetic neutron fields. *Metrologia* 2014*,* 51 (5),431–440.
https://doi.org/10.1088/0026-1394/51/5/431

[37] **Sato, T.; *et al.*** Analysis of cell-survival fractions for heavy-ion irradiations based on microdosimetric kinetic model implemented in the particle and heavy ion transport code system. *Radiat. Prot. Dos.* 2011*,* 143 (2-4),491–496.
https://doi.org/10.1093/rpd/ncq484

[38] **Parisi, A.; *et al*.** The Mayo Clinic Florida Microdosimetric Kinetic Model of Clonogenic Survival: Application to Various Repair-Competent Rodent and Human Cell Lines. *Int. J. Mol. Sci.* 2022, 23, 12491.
https://doi.org/10.3390/ijms232012491

[39] **Elässer, T.; *et al.*** Quantification of the relative biological effectiveness for ion beam radiotherapy: direct experimental comparison of proton and carbon ion beams and a novel approach for treatment planning. *Int. J. Radiation Oncology Biol. Phys.* 2010, 78 (4), 1177–1183.
https://doi.org/10.1016/j.ijrobp.2010.05.014