A simulation approach for determining the spectrum of DNA damage induced by protons

Mojtaba Mokari1, Mohammad Hassan Alamatsaz1, Hossein Moeini1 and Reza Taleei2,3
1 Department of Physics, Isfahan University of Technology, Isfahan 84156-83111, Iran
2 Department of Radiation Oncology, University of Virginia, Charlottesville, VA 22908, United States of America
3 Author to whom any correspondence should be addressed.
E-mail: reza.taleei@virginia.edu
Keywords: DNA damage, Geant4-DNA, Monte Carlo simulations, double-strand break

Abstract
To study the molecular damage induced in the form of single-strand and double-strand breaks by ionizing radiation at the DNA level, the Geant4-DNA Monte Carlo simulation code for complete transportation of primary protons and other secondary particles in liquid water has been employed in this work. To this aim, a B-DNA model and a thorough classification of the complexity of the DNA damage were used. Strand breaks were assumed to have primarily originated by direct physical interactions via energy depositions, assuming a threshold energy of 17.5 eV, or indirect chemical reactions of hydroxyl radicals, assuming a probability of 0.13. The simulation results on the complexity and frequency of various damages are computed for proton energies of 0.5–20 MeV. The yield results for a cell (Gy cell)−1 are presented, assuming 22 chromosomes per cell and a mean number of 245 Mbp per chromosome. The results show that for proton energies below 2 MeV, more than 50% of the energy depositions within the DNA volume resulted in strand breaks. For double-strand breaks (DSBs), there is considerable sensitivity of DSB frequency to the proton energy. A comparison of DSB frequencies predicted by different simulations and experiments is presented as a function of proton linear energy transfer (LET). We show that our yield results (Gy Gbp)−1 are generally comparable with various experimental data and there seems to be a better agreement between our results and a number of experimental studies when compared to other simulations.

Introduction
When ionizing radiation interacts with cells, it will cause both early and late biophysical effects. Early effects could last for a few femtoseconds to a few days, while the late effects would last for years. The initial effects of radiation include the effects from physical processes due to ionization and excitation interactions as well as the effects of chemical radicals (Nikjoo and Uehara 2004). The interaction of ionizing radiation with cells that would instigate DNA damage has a major role in cancer therapy. DNA damage includes single-strand and double-strand breaks (SSB and DSB). Double-strand breaks can cause the death of the cell, if they result in mis-repair or unrepaired damage (Nikjoo et al 2016).

Until now an accurate quantitative study of various types of damage, e.g. complex type DSBs, has not been possible (Nikjoo et al 2012). As such, people try to use simulation methods to study the biophysical interactions of radiation. The most common Monte Carlo simulation codes use the spatial history of the particles inside the matter that are implemented in codes such as FLUKA (Ferrari et al 2005), Geant4 (Agostinelli et al 2003, Incerti et al 2010), MCNP (Goorley et al 2012), MCEP (Uehara 1986), PITS (Wilson and Nikjoo 1999), PENELOPE (Fernández-Varea et al 2012), or Monte Carlo Track Structure (MCTS) codes like PARTRAC (Friedland et al 2011) and KURBUC (Taleei et al 2013, Nikjoo et al 2016). When calculating the damage, parameters such as the type of incident radiation, radiation energy, interaction cross sections, threshold energy $E_{th}$ (Nikjoo et al 1997), and the probability of indirect interactions of chemical radicals with the DNA (Nikjoo et al 2002) appear to be important. With the MCDS code, which has a quasi-phenomenological algorithm, one avoids the initial
simulation of the physical and chemical processes (Semenenko and Stewart 2005). To calculate damage frequency using codes such as KURBUC or Geant4-DNA, one first simulates the ionizing radiation in the environment and subsequently calculates the damage considering the energy deposits and the probabilistic effect of the chemical radicals. KURBUC is a comprehensive code for simulating various types of particles such as electrons, protons, neutrons, and heavy ions. KURBUC-Proton was the first MCTS code released in 2001 for simulating proton interactions, which covers the proton energy ranges of 1 keV to 1 MeV with the unique feature of including a charge-exchange model for $\text{H}^+$, $\text{H}^0$, and $\text{H}^-$ in the Bragg peak region (Uehara et al 2001). Geant4-DNA is a subproject of Geant4—a general purpose particle-matter Monte Carlo simulation toolkit—that is extended with processes for modelling the early biological damage induced by ionizing radiation at the DNA scale. The implemented models and physics processes in Geant4-DNA allow for step-by-step simulation of the interaction of particles in few materials, including liquid water, down to the eV energy scale (Bernal et al 2015). Meylan et al (2017) simulated a fibroblast cell nucleus using Geant4-DNA with protons. Lampe et al (2017) effectively simulated the bacterial nucleus and studied the DNA damage from electrons and protons in a modelled full genome of Escherichia coli cell using Geant4-DNA.

In this work, we have used Geant4-DNA (Geant4 version 10.3) to simulate primary protons, with energies ranging from 0.5 MeV to 20 MeV, and the resulting secondary particles in liquid water. The main objective of this article is to calculate the direct and indirect DNA damage by incident protons and secondary particles cumulatively, taking into account both physical and chemical interactions. We also compared our yield results to the published experimental data and as such conducted a benchmarking of the Geant4-DNA performance taking an alternative approach to calculation when compared to the existing Geant4-DNA simulations. For this purpose, the frequency of simple and clustered damages of different complexity, as well as the yield for SSB and DSB, are determined. Yield values are calculated per nucleotide pair (base pair) and per cell and the results are compared with other simulation works and experimental studies of Frankenberg et al (1999), Belli et al (2000), Belli et al (2001), and Campa et al (2005). Calculation of the yield per nucleotide pair is conducted in two ways. One is based on the total number of SSBs and DSBs, whereas the other method uses the frequency of the deposited energies.

**Methods**

When a cell is irradiated, the initial damage to the DNA molecules would be the result of direct physical interactions of the ionizing radiation with the DNA or indirect chemical reactions of the produced free radicals in the surrounding water. The simulation study in this work has been performed using the Geant4-DNA toolkit for particle transport through water. Geant4-DNA offers a variety of models to simulate the physical interactions of electrons in liquid water. With this code, interactions are grouped into three categories: elastic interactions, inelastic interactions (ionization and electronic excitation) and inelastic sub-excitation interactions (vibrational excitation and molecular attachment) (Incerti et al 2018). A physical interaction is explained by a physics process that can extract several models, which can be either complementary or alternative. Therefore, several cross section models could be taken into account (Incerti et al 2010). We selected default physics interactions implemented in Geant4-DNA for electron transport in liquid water to collect all physics processes and required lists of particles. The electron models (in this work) for elastic scattering (7.4 eV–1 MeV), electronic excitation (9 eV–1 MeV), ionization (11 eV–1 MeV), vibrational excitation (2 eV–100 eV), and molecular attachment (4 eV–13 eV) interactions are the G4DNAChampionElasticModel, G4DNABornExcitationModel, G4DNABornIonisationModel, G4DNASancheExcitationModel and G4DNAMeltonAttachmentModel. Also, the interactions for protons and neutral hydrogen atoms for all physical references are the same, including the models for elastic scattering (100 eV–1 MeV), electronic excitation (10 eV–100 MeV), ionization (100 eV–100 MeV), and electron capture or loss (100 eV–100 MeV) (Incerti et al 2010, de la Fuente Rosales et al 2018). Taking into account both physical and chemical interactions, the toolkit allows extraction of information, such as energy deposition, position and time of flight of various particles and radiations at different stages of their propagation through matter. On their path through matter, the primary protons and secondary particles can undergo various processes. The cross sections used in Geant4-DNA account for elastic scattering, ionization, excitation, and Auger cascades, which are of importance in our DNA damage simulations. The energy cut-off threshold for tracking electrons is 7.4 eV (the default of the Geant4-DNA), below which the tracking would cease and the remaining kinetic energy of the electron would be deposited at once.

Different representations of DNA have been employed in earlier studies to model DNA damage. Charlton et al (1989) and Nikjoo et al (2001) used the B-DNA model. This model consists of a cylinder divided into sugar–phosphate and base regions without considering the details of atomic structures in oligonucleotides. The sugar–phosphate chains surround the center of a cylinder with a 10 Å diameter and a 36 degree helical rotation. The DNA molecule diameter is 23 Å. Another common DNA model is the phosphodiester group (PDG) which consists of prisms with circular center bases used in the works of Bernal et al (2011). Friedland et al (2002) also used...
the PDG model and defined the position of phosphor, oxygen, hydrogen and carbon atoms with the van der Waals radius (Pater et al 2014). However, the atomic structure of the DNA double helix is generally considered as well-known at the smallest scale (Branco et al 2007) but DNA structures of higher levels are described differently by several models. In most studies, the models are simplified because they represent a unique chromatin structure. Recently, the DnaFabric project was initiated to manipulate complex DNA geometrical models for creation of adaptable and complex DNA models (Meylan et al 2016). Using a protein data bank file, we extracted the position of the atoms of a 216 bp long double helix B-DNA (equivalent to 73.44 nm and consisting of 432 nucleotides). A B-DNA molecule is one of the common DNA structures in living creatures (Dickerson et al 1982, Meylan et al 2016). Each nucleotide comprises sugar–phosphate groups and a base group (see figure 1 for an artist’s view of such DNA shapes). We then sampled the DNA molecules within the volume of a water sphere of 100 nm radius. Our isotropic point source was located at the center of the water sphere that emitted protons of desired energy. The sampling of DNA within the spherical volume was performed based on the $\mu$-randomness method (Kellerer 1975). We found an optimized number of DNA molecules which could satisfy the two sampling accuracy criteria explained by Nikjoo et al (1989, 1991). For the first criterion, we compared the ratio of the energy deposition in the original sphere to its volume with the ratio of energy deposition in the DNA molecules to their volumes. For the second criterion, the mean of the inverse frequency-averaged mean specific energy ($\bar{Z}_f^{-1}$) was calculated and compared with the frequency of hits of any size ($f(>0)$) (Nikjoo et al 1991). For both tests, in order to conduct a good sampling, the ratios were targeted to be equal to within 5% uncertainty.

The calculations for this study comprised three stages. The physical stage was the first stage, where the simulation of physical interactions of various particles in water was pursued until they reached the energy or geometrical cut-off. In the second stage, hereafter referred to as the chemical stage, the simulation of physico-chemical and chemical processes up to 1 ns was performed. In the third stage, referred to as the damage formation stage, a written algorithm determined the damage types according to the definition of damage spectra given by Nikjoo et al (1997).

At the end of the physical stage, when the tracking of various particles is completed, the positions and deposited energies, at the end of each step involving ionization or excitation, were derived. For the sum of such energy depositions in one sugar–phosphate volume that was more than a threshold value $E_{\text{ssb}}$, a strand break (SB) was registered at the corresponding DNA segment (Pater et al 2014). In accordance with the experimental studies by Martin and Haseltine (1981), Terrissol (1994), and Kandaiya et al (1996), the threshold value was assumed to be $E_{\text{ssb}} = 17.5$ eV.

For the chemical stage, we have exploited the TimeStepAction class of the Geant4-DNA, where we recorded the positions of the produced radicals in the environment 1 ns after throwing the primary proton (hereafter referred to as the chemical stage simulation time). Among the various radicals and molecules that are produced...
within the water environment, including $H_2O_2 \cdot H_2 \cdot e^{-}_{aq} \cdot OH^- \cdot OH^- \cdot H^+ \cdot OH^-$, hydroxyl radicals (OH•), would occur more commonly due to their capability to interact with the DNA segments (Roots and Okada 1975, Milligan et al 1996). Table 1 makes a comparison between the simulation and experimental outcomes of the reaction rates for a few chemical reactions involving the production of different radicals. It can be seen that overall, in Geant4-DNA, more hydroxyl radicals would react in the environment and hence the share of indirect damages is assumed to be higher when compared to experimental data. The derived positions of the hydroxyl radicals were then checked in our algorithm to see whether they would fall within the volume of any imaginary cylinder of $(8 + 2.3)$ nm diameter, with its longitudinal axis coinciding with the axis of the DNA cylinder of 2.3 nm diameter. For the corresponding DNA molecules of those cylinders that can pass this condition, the closest sugar or phosphate nucleotide to the hydroxyl radical would then be found and an SB at that DNA segment would be registered. The probabilities of a hydroxyl radical interacting with the base and sugar–phosphate are 80% and 20%, respectively. Therefore, the sugar radicals produced due to the interaction of hydroxyl with sugar–phosphate lead to SBs with a 65% probability. Consequently, the probability of SB damage (indirect damage) due to the interactions of hydroxyl radicals with DNA nucleotides is equal to 0.13 (Nikjoo et al 1997). Following the approach of Nikjoo et al (1997), we increased the number of events from $10^3$ to $5 \times 10^4$.

As stated earlier, each physical and chemical interaction (in this work, corresponding to satisfying the $E_{ab}$ and $P_{OH}$ conditions) is considered to cause one SB. The classification of the DNA SBs is performed according to either the complexity of the clustered damage or the origin of the breaks (direct and indirect). The categorization of the DNA damage types is performed according to Nikjoo et al (2016) and shown in figure 2, using an algorithm written in Python programming.

### Results

Table 2 shows the calculated relative yields of different types of SBs as a function of proton energies. In this table, complex damage has been defined as SSBc (=SSB + 2SSB) and DSBc (=DSB + DSB++) (Nikjoo et al 2001). These data show that at low energies the majority of proton hits in DNA do lead to damage in the form of SBs. Specifically, this is about more than 50% of hits, for proton energies less than 2 MeV. It also shows that the frequency of simple SSBs is independent of primary proton energy, comprising about one third of hits. It is interesting to note that 2SSB appears to be the second most frequent type of damage and has a frequency of more than three times that of the SSB+ for all proton energies. This is contrary to the damage frequency from primary electrons Nikjoo et al (1997), where the frequency of 2SSB is substantially less than SSB+ for various energies. It is also observed that the yields of DSB+ and DSB++ are comparable to simple DSBs at proton energies lower than 2 MeV. Furthermore, the yields of DSBs are about two times lower than those of 2SSB. This is especially interesting considering the definitions of 2SSB and DSB damage, based on the 10 bp distance between the two breaks. It can be seen that, for high proton energies above 2 MeV, they decrease with energy relatively close to each other; whereas for lower primary energies, SSB appears to be about 10% less than DSB+. The data in this table do not account for the role of base damage in inducing DNA strand breaks or adding to their complexities. Hence, these data can only represent a lower limit of the complexity of damages.

Figure 3 compares our results for DSB yield with other simulation and experimental data, as a function of the proton linear energy transfer (LET) in water. Although there is a considerable difference between our results and the simulation works of Nikjoo et al (2001) and Friedland et al (2003), the estimated yields (for LET values below about 30 keV $\mu m^{-1}$) are reasonably close to those estimated by the simulations of Meylan et al (2017). Also, the trends of the damage yields for various simulations and experimental data are similar. Our results show a better agreement with the experimental data of Belli et al (2000) and Campa et al (2005), and reproduce the data of Frankenberger et al (1999) at LETs close to 25 keV $\mu m^{-1}$.

Figure 4 compares the frequency of the direct and indirect damages (as the percentage of the total number of breaks), as a function of the number of damage sites. For example, for 0.5 MeV protons, the probability for direct interactions capable of inducing one break per DNA segment is about 31.2% of the total breaks. In the case of

### Table 1. Chemical interactions and reaction rates in the production of radicals according to Geant4-DNA (Karamitros et al 2014) and experimental data (Buxton et al 1988). For a more comprehensive list of reactions, refer to Karamitros et al (2014), Buxton et al (1988).

| Reaction                      | Reaction rate (Geant4-DNA) ($dm^3 \text{mol}^{-1} s^{-1}$) | Reaction rate (experiment) ($dm^3 \text{mol}^{-1} s^{-1}$) |
|-------------------------------|------------------------------------------------------------|------------------------------------------------------------|
| $H_2 + OH \rightarrow H^+ + H_2O$ | $4.17 \times 10^7$ | $4.5 \times 10^7$ |
| $OH^- + OH \rightarrow H_2O_2$ | $0.44 \times 10^{10}$ | $0.6 \times 10^{10}$ |
| $H_2O_2 + e^-_{aq} \rightarrow OH^- + OH$ | $1.41 \times 10^{10}$ | $1.3 \times 10^{10}$ |
| $H^+ + OH \rightarrow H_2O$ | $1.44 \times 10^{10}$ | $2.0 \times 10^{10}$ |
| $OH^- + e^-_{aq} \rightarrow OH^-$ | $2.95 \times 10^{10}$ | $2.5 \times 10^{10}$ |
0.5 MeV protons with a single damage site, the obtained value for the frequency of All breaks appears to be less than the corresponding values for Direct and Indirect breaks. This can be explained as follows. The frequency of All breaks with a single damage site is the sum of the frequencies of Direct and Indirect breaks with a single damage site minus two times the frequency of damages that are labeled as both single Direct and single Indirect breaks. Hence for 0.5 MeV protons, assuming that the frequencies of All, Direct, and Indirect breaks with a single damage site are $\text{All}_{\text{single}}$, $\text{Direct}_{\text{single}}$, and $\text{Indirect}_{\text{single}}$, respectively, the frequency of All breaks is

\[ \text{All}_{\text{single}} = \text{Direct}_{\text{single}} + \text{Indirect}_{\text{single}} - 2 \times \text{both single Direct and single Indirect breaks}. \]

This relationship allows for a more accurate estimation of the frequency of All breaks, especially when dealing with single damage sites.

---

**Figure 2.** Types of DNA damage induced by direct energy deposition and the reactions of OH radicals. For simplicity, the DNA is shown as four parallel lines. The solid lines represent the sugar–phosphate (S–P) backbone and the dashed lines represent the bases. A ·×· represents an SB in DNA. If only one ·×· is on either of the strands, it is labeled as SSB. If two ·×· are on opposite strands within 10 bp of each other, it is labeled as DSB. If two SSBs (one on each strand) are more than 10 bp apart, it is labeled as SSB+. A DSB accompanied by additional SSBs within 10 bp separation is labeled as DSB+. More than one DSB in the DNA is labeled as DSB++. The NB (no break) category refers to a DNA molecule without any SBs (Nikjoo et al. 2016).

**Table 2.** Relative yield of SBs classified by complexity as a function of proton energy.

| Energy (MeV) | LET (keV/µm) | NB (%) | SSB (%) | SSB+ (%) | 2SSB (%) | DSB (%) | DSB+ (%) | Y_{SSB} (Gy Gbp⁻¹) | Y_{DSB} (Gy Gbp⁻¹) |
|-------------|---------------|--------|---------|----------|----------|---------|----------|-------------------|-------------------|
| 0.5         | 39.7          | 26.59  | 29.26   | 4.06     | 17.42    | 10.27   | 7.70     | 4.68              | 42.33             |
| 1           | 24.2          | 37.92  | 31.16   | 3.52     | 13.75    | 7.59    | 4.90     | 1.18              | 35.66             |
| 2           | 13.9          | 50.44  | 30.34   | 2.59     | 9.93     | 4.67    | 1.76     | 0.28              | 29.19             |
| 10          | 3.4           | 63.29  | 27.35   | 1.57     | 5.19     | 2.10    | 0.47     | 0.03              | 19.81             |
| 20          | 1.9           | 67.92  | 25.27   | 1.10     | 3.95     | 1.50    | 0.25     | 0.01              | 16.67             |

**Figure 3.** Comparison of our simulation results for $Y_{DSB}$ (×) with the simulation and experimental results of Nikjoo et al. (2001), Friedland et al. (2003), Meylan et al. (2017), Frankenberg et al. (1999), Belli et al. (2000) (diamonds), Belli et al. (2001) (pentagon), and Campa et al. (2005).
damage sites are about 29.7%, 31.2%, and 32.4%, respectively, the frequency of the damage sites that are tagged as both single Direct and single Indirect breaks is found to be about 16.9% ($=\frac{31.2\% + 32.4\% - 29.7\%}{2}$).

Thus, for low enough proton energies, e.g. 0.5 MeV, it is more likely to create both single Direct and single Indirect breaks (16.9%) than either single Direct (31.2% − 16.9% = 14.3%) or single Indirect breaks (32.4% − 16.9% = 15.5%). When compared with other available data, the results for the total number of breaks and 1 MeV protons agree closely with those in Nikjoo et al (2001). As a general feature of these diagrams, the probability of creating direct or indirect single-site damage (All breaks) significantly increases with increasing proton energy. For the probability of creating direct or indirect double (and to some extent triple) site damage, there is, however, not much sensitivity observed to the proton energy. For a greater amount of site damage there is a slow gradual decrease in the frequency with energy. The probability of creating both single Direct and single Indirect breaks for other proton energies 1, 2, and 10 MeV are obtained from figure 4 to be about 18.4%, 17.9%, and 14.9%, respectively.

Table 3 lists the number of different types of breaks as a function of the deposited energy in the DNA segment. The two columns on the left represent the energy deposition intervals and the corresponding number of events falling within each interval (total hits). The body of the table indicates the number of events within an energy interval, which produced a specific type of break in our simulations. Thus, of the total 1913 events which deposited between 100 and 150 eV, 739 (39%) produced an SSB. The total number of SSBs created by energy depositions in this energy interval is calculated to be $739 + 168 + 2 \times (563 + 261 + 52 + 2) = 2663$, in which each DSB amounts to two SSBs (SSB$_{all} = \text{SSB} + \text{SSB}^+ = 2 \times (2 \text{SSB} + \text{DSB} + \text{DSB}^+ + \text{DSB}^{++})$ and DSB$_{all} = \text{DSB} + \text{DSB}^+ + \text{DSB}^{++}$ (Charlton et al 1989). Hence, the deposited energy in this interval creates 1.39 SSB all ($=2663/1913$) per event of 216 bp size. Similarly, for DSBs, a total of 0.16 DSB all ($=315/1913$) per event are produced in this range of energy deposition.

Figure 5 presents the results of table 3 plus the ones for other proton energies. The damage yield is calculated for 1000 events and is normalized to the number of hits. One general feature of these graphs is that as the complexity of the breaks increases, the frequency patterns extend and peak at a higher deposited energy. Also, for each break type, the lower the proton energy, the higher the number of extensions and the number of breaks appears to be (corresponding to the integral of the frequency graph). Table 4 shows the frequency distribution of the total number of hits (see table 3, left column) per unit dose per DNA segment, as a function of the total energy deposited in the DNA segment.

Figure 6 shows the average total number of SSB and DSB breaks per DNA segment, for different primary proton energies. The remarkable feature of the data is that for a given energy deposited in the DNA segment,
there seems to be no significant dependence on the proton energy, particularly for the results of the SSBs. Also to be noted is the saturation effect in the pattern of the SSBs at high LET values, where the probability of producing a break per unit energy deposition decreases with increasing energy deposition.

Table 5 compares the yield of the damage sites (Gy Gbp)$^{-1}$, obtained from the total number of SSBs and DSBs, with the values obtained from the frequency of the deposited energies that are calculated as follows. Let \( n(E, y) \) be the number of breaks of a particular type, e.g. SSB or DSB, produced when energy \( E \) is deposited in a DNA segment of length \( y \), and let \( P(E, y) \) be the number of energy depositions of size \( E \) in length \( y \) per unit dose. Hence, \( n(E, y) \) would correspond to the values in figure 6, whereas \( P(E, y) \) would correspond to the values in table 4 (for example, for 10 eV deposited energy and 0.5 MeV protons, it is \( 0.105 \times 10^{-5} \)). Following the procedure in Charlton et al (1989), the yield can thus be calculated as \( \text{Yield} = \sum n(E, y) \times P(E, y) / y \), in which the summation runs over the deposited energy values and \( y = 216 \) is expressed in bp. The yield results of this calculation are shown in the 4th and 5th columns of table 5. As can be seen, the results of these two columns reproduce the corresponding results of the 2nd and 3rd columns reasonably well. The reason for such similarity in the outcomes is that in principle the two methods of calculation are the same; however, in employing the above-mentioned formula of Charlton et al (1989), where we used the numbers in table 4, we have chosen the mid values of each interval of deposited energy (e.g. 10 eV for the first interval) as \( E \) in the yield formula. The yield values (Gy cell)$^{-1}$

| Deposited energy (eV) | Total hits | NB | SSB | SSB$^+$ | 2SSB | DSB | DSB$^+$ | DSB$^{++}$ |
|----------------------|------------|----|-----|--------|------|-----|--------|---------|
| 0–20                 | 9933       | 8649 | 1211 | 9       | 56   | 8    | 0      | 0       |
| 20–40                | 4627       | 2867 | 1569 | 26      | 144  | 21   | 0      | 0       |
| 40–60                | 2744       | 979  | 1471 | 37      | 203  | 53   | 1      | 0       |
| 60–80                | 1911       | 430  | 1075 | 64      | 275  | 61   | 6      | 0       |
| 80–100               | 1351       | 173  | 725  | 72      | 291  | 84   | 6      | 0       |
| 100–150              | 1913       | 128  | 739  | 168     | 563  | 261  | 52     | 2       |
| 150–200              | 1076       | 19   | 266  | 105     | 383  | 215  | 81     | 7       |
| 200–250              | 636        | 2    | 78   | 77      | 225  | 172  | 76     | 6       |
| 250–300              | 369        | 0    | 21   | 28      | 124  | 111  | 70     | 15      |
| 300–350              | 198        | 0    | 4    | 14      | 43   | 68   | 60     | 9       |
| 350–400              | 80         | 0    | 0    | 4       | 12   | 23   | 29     | 12      |
| 400–450              | 35         | 0    | 0    | 1       | 5    | 8    | 15     | 6       |
| >450                 | 128        | 0    | 37   | 8       | 30   | 22   | 22     | 9       |

Figure 5. The number of breaks per hit for 1000 events, as a function of the deposited energy in the DNA segment, for various proton energies.
are also presented in this table. In the derivation of the latter results, the average molecular weight of a chromosome was calculated by assuming 22 chromosomes per cell and a mean number of 245 Mbp per chromosome. The relative chromosome size distribution is taken from the cytometry techniques to range from 610 Mbp to 40 Mbp with a number-averaged chromosome size of 245 Mbp (Nusse et al 1987, Taleei et al 2013).

Discussion

Introducing a primary proton source, various damage types in the DNA sample were calculated, based on the Geant4-DNA simulations for physical and chemical stages. For the physical stage, the threshold energy for recording a hit as a break was considered to be 17.5 eV. The same value has been used in the simulations by Meylan et al (2017) where they simulated a fibroblast cell nucleus using Geant4-DNA, and Nikjoo et al (2001) where they simulated B-DNAs using KURBUC. Using the PARTRAC code, Friedland et al (2003) investigated a threshold variation between 5 and 37.5 eV, implementing a linear acceptation probability (a linear increase in the probability from zero, for a deposited energy less than 5 eV, to 1 when it exceeds 37.5 eV) for direct damage. In their simulations, they implemented a basic chromatin fiber element, including 30 nucleosomes and an ideal arrangement of chromatin fiber rods in rhombic loops forming a rosette-like structure of 0.5 Mbp genomic length. Whereas the interaction probability of the hydroxyl radicals is considered in Meylan et al (2017) to be 0.4, we have adopted the value 0.13, which is the same as in Nikjoo et al (2001) and Friedland et al (2003). As in Nikjoo et al (2001), we limited the chemical stage simulation time to 1 ns for the interaction of hydroxyl radicals with DNA—to be compared with 2.5 ns in Meylan et al (2017). In Geant4-DNA, the chemical stage is simulated

| Deposited Energy (eV) | 0.5 MeV protons | 1 MeV protons | 2 MeV protons | 10 MeV protons | 20 MeV protons |
|-----------------------|-----------------|---------------|---------------|----------------|----------------|
| 0–20                  | 0.105           | 0.314         | 0.857         | 2.399          | 3.543          |
| 20–40                 | 0.092           | 0.209         | 0.399         | 0.718          | 0.841          |
| 40–60                 | 0.074           | 0.138         | 0.237         | 0.368          | 0.406          |
| 60–80                 | 0.065           | 0.106         | 0.165         | 0.233          | 0.231          |
| 80–100                | 0.051           | 0.078         | 0.117         | 0.138          | 0.132          |
| 100–150               | 0.093           | 0.133         | 0.165         | 0.177          | 0.154          |
| 150–200               | 0.06            | 0.081         | 0.093         | 0.058          | 0.032          |
| 200–250               | 0.041           | 0.053         | 0.055         | 0.019          | 0.002          |
| 250–300               | 0.031           | 0.038         | 0.032         | 0.004          |                |
| 300–350               | 0.024           | 0.03          | 0.017         | 0.002          |                |
| 350–400               | 0.021           | 0.02          | 0.007         |                |                |
| 400–450               | 0.016           | 0.016         | 0.003         |                |                |
| >450                  | 0.074           | 0.030         | 0.010         |                |                |
over several time steps during which the movement of all the molecules is governed by their diffusion coefficients (Karamitros et al 2014). In our simulations, we did not specifically model the scavenging reactions that decrease the number of the existing hydroxyl radicals for damaging the DNA, whereas Friedland et al (2003) have taken into account the scavenging of the chemical species at each time step due to random absorption of the radicals, and as such considered an appreciably longer chemical stage simulation time of 10 ns.

It is already known that the cross sections used in the Geant4-DNA for the elastic, ionization, and excitation processes are less than those implemented in other codes (Bordage et al 2016, Famulari et al 2017). For example, for electron energies higher than 100 eV, the difference between the excitation cross sections of CPA100 and Geant4-DNA amounts to about an order of magnitude (Bordage et al 2016). Although the ionization cross sections of CPA100, compared to the other codes, agree better with the experimental ones, the ionization cross sections of Geant4-DNA and CPA100 are in reasonable agreement with each other, for electrons of more than 100 eV energy (produced numerously as secondary particles, having a primary proton source), where the ionization is the most important process (Bernal et al 2015, Bordage et al 2016). Furthermore, the maximum total excitation cross section in Geant4-DNA is smaller than in PARTRAC (Dingfelder 2008). Table 1 shows that the chemical reaction rates of the hydroxyl radicals with other molecules and radicals (including hydroxyl) are lower in Geant4-DNA when compared to the experimental values (see, for example, the first row of table 1), whereas the production rates of the hydroxyl radicals are greater in Geant4-DNA when compared to the experimental results (see the third row of table 1). Hence, in Geant4-DNA, the share of indirect damage is greater and more hydroxyl radicals are subject to reactions.

Despite the above differences between various simulations, the results of figure 3 with PARTRAC and KUR-BUC are comparable to our results and those presented in Meylan et al (2017): the latter two using Geant4-DNA with different geometries. The reason for such agreement is partly due to using similar methodology in terms of the definition of SB with regards to the physical and chemical processes. Our simulations are in good agreement with the experimental data of Frankenberg et al (1999) up to around 25 keV μm−1 as well as Belli et al (2000) and Campa et al (2005) over the whole range of LET shown in figure 3 (although there are only three data points available from the latter two experiments). This is especially true with regards to the data of Frankenberg et al (1999) around a LET of 25 keV μm−1, Belli et al (2000) around 11 keV μm−1, and Campa et al (2005) around 28 keV μm−1. As with the general trend of the experimental data, our DSB yield increases with the LET of the primary protons up to 7.8 (Gy Gbp)−1 for a LET of 39.7 keV μm−1. Figure 3 also shows that the DSB yield increases with the LET according to various simulations. With the exception of the DSB yield results of Nikjoo et al (2001), which appear to show a linear increase with LET, our simulations, as well as the simulations of Friedland et al (2003) and Meylan et al (2017), show a less pronounced increase in the DSB yield above a certain value of LET (about 35 keV μm−1 for the latter two works and 25 keV μm−1 in our simulations). It is also interesting to note that among different simulations presented in figure 3, our results seem to show a better agreement of the trend and proximity to the experimental data points of Belli et al (2000) and Campa et al (2005).

Although we have considered the energy deposition contributions of the primary and secondary particles cumulatively in the SSB and DSB calculations, a better understanding of the spectrum of DNA damage could be obtained by separating the contributions of protons and secondary delta electrons. Such differentiation of the processes of high LET ion tracks of low fluence and low LET electrons of high fluence may lead to a more detailed evaluation of the biological effects of radiation than with the LET approach. By calculating the ionization and excitation event rates induced by secondary electrons and charge-changing reaction processes one can estimate the effective production rate of ionization and excitation around the Bragg peak per unit distance traveled by the proton and investigate the implicit effects of the secondary electrons on the stopping power of protons (Date et al 2006).

| Energy (MeV) | $Y_{SSB}^*$ (Gy Gbp)$^{-1}$ | $Y_{DSB}^*$ (Gy Gbp)$^{-1}$ | $Y_{SSB}^*$ (Gy Gbp)$^{-1}$ | $Y_{DSB}^*$ (Gy Gbp)$^{-1}$ | $Y_{SSB}^*$ (Gy cell)$^{-1}$ | $Y_{DSB}^*$ (Gy cell)$^{-1}$ |
|-------------|-----------------|-----------------|-----------------|-----------------|-----------------|-----------------|
| 0.5         | 39.05           | 7.80            | 38.97           | 7.76            | 210.05          | 41.83           |
| 1           | 50.92           | 7.77            | 50.66           | 7.68            | 273.06          | 41.39           |
| 2           | 62.74           | 6.36            | 62.72           | 6.38            | 338.06          | 34.39           |
| 10          | 73.45           | 4.3             | 73.46           | 4.27            | 395.95          | 23.01           |
| 20          | 75.15           | 3.50            | 75.02           | 3.58            | 404.36          | 19.30           |
Conclusions

In this article, using Geant4-DNA simulations and Python programming for data analysis, we presented our results for the frequency of simple and complex damage caused by primary protons in a B-DNA model. The simulation toolkit allowed us to calculate the energy depositions from physical interactions of protons and secondary particles, and to account for the role of the hydroxyl radicals in producing strand breaks through chemical reactions. Simulations were performed introducing a point source of protons at the center of a spherical liquid water medium (Nikjoo et al. 1989, 1997, 1999, 2001), isotropically thrown in full phase space with energies ranging from 0.5 MeV to 20 MeV. As such, the probability of simple and complex damages as well as SSB and DSB yields were calculated. The results were compared with other simulation works as well as with the data obtained from experiments (Frankenberg et al. 1999, Belli et al. 2000) that used pulsed field gel electrophoresis. Overall, there is reasonable agreement between our results and the presented experimental data. In particular, our results, when compared with other presented simulations, provide better similarities with the trends and values of the experimental data of Belli et al. (2000) and Campa et al. (2005). The discrepancies observed between our DSB yield results and the corresponding results in other simulation works can be assigned to the exploited DNA geometries, the type of chemical processes considered in the simulations, and different parameter adjustments and criteria for SB registration caused by physical or chemical processes.

Acknowledgments

The authors gratefully acknowledge the Sheikh Bahaei National High Performance Computing Center (SBNHPC) for providing computing facilities and time. The SBNHPC is supported by Isfahan University of Technology (IUT).

ORCID iDs

Mojtaba Mokari https://orcid.org/0000-0001-6282-1486

References

Agostinelli S et al 2003 Geant4—a simulation toolkit Nucl. Instrum. Methods Phys. Res. A 506 250–303
Barrios R, Skurski P and Simons J 2002 Mechanism for damage to DNA by low-energy electrons J. Phys. Chem. B 106 7991–4
Belli M et al 2001 DNA fragmentation in mammalian cells exposed to various light ions Adv. Space Res. 27 393–9
Belli M, Cherubini R, Dalla Vecchia M, Dini V, Moschini G, Signoretti C, Simone G, Tabocchini M A and Tiveron P 2000 DNA DSB induction and rejoining in V79 cells irradiated with light ions: a constant field gel electrophoresis study Int. J. Radiat. Biol. 76 1095–104
Bernal M A et al 2015 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. 31 661–74
Bernal M A, deAlmeida C E, Sampaio C, Incerti S, Champion C and Nieminen P 2011 The invariance of the total direct DNA strand break yield Med. Phys. 38 4147–53
Bordage M C, Bordes J, Edel S, Terrissol M, Franckes X, Bardies M, Lampe N and Incerti S 2016 Implementation of new physics models for low energy electrons in liquid water in Geant4-DNA Phys. Med. 32 1833–40
Branco M R and Pombo A 2007 Chromosome organization: new facts, new models Trends Cell Biol. 17 127–34
Buxton G V, Greenstock C L, Helman W P and Ross A B 1988 Critical review of rate constants for reactions of hydrated electrons, hydrogen atoms and hydroxyl radicals (·OH/·O)1 in aqueous solution J. Phys. Chern. Ref. Data 17 513–86
Campa A et al 2005 DNA DSB induced in human cells by charged particles and gamma rays: experimental results and theoretical approaches Int. J. Radiat. Biol. 81 841–54
Charlton D, Nikjoo H and Humm J L 1989 Calculation of initial yields of single- and double-strand breaks in cell nuclei from electrons protons and alpha particles Int. J. Radiat. Biol. 56 1–19
Date H, Sutherland K L, Hayashi T, Matsuzaki Y and Kiyanagi Y 2006 Inelastic collision processes of low-energy protons in liquid water Radiat. Phys. Chem. 75 179–87
De la Fuente Rosales L, Incerti S, Franchin C and Bernal M A 2018 Accounting for radiation-induced indirect damage on DNA with the Geant 4-DNA code Phys. Med. 31 108–16
Dickerson R E, Drew H R, Conner B N, Wing R M, Fratini A V and Kopka M L 1982 The Anatomy of A- B- and Z-DNA Science 216 475–85
Dingfelder M, Ritchie R H, Turner J E, Friedland W, Paretzke H G and Hamm R N 2008 Comparisons of calculations with PARTRAC and NOREC: transport of electrons in liquid water Radiat. Res. 163 105–12
Ferrari A, Sala P, Fasso A and Ranft J 2005 FLUKA: A Multi-Particle Transport Code (Geneva: CERN)
Frankenberg D, Brede H J, Schrewe U J, Steinmetz M, Frankenberg-Schweiger M, Kasten G and Pralle E 1999 Induction of DNA double-strand breaks by 1 and 4he ions in primary human skin fibroblasts in the LET range of 8 to 124 keV·µm−1 Radiat. Res. 151 504–9
Friedland W, Bernhardt P, Jacob P, Paretzke H G and Dingfelder M 2002 Simulation of DNA damage after proton and low LET irradiation Radiat. Prot. Dosim. 99 99–102
