Introduction

With the wide application of radiation therapy, the effect of ionizing radiation at the (sub-)cellular level has been the subject of substantial research. This is due to the strong evidence that DNA is the critical target for most of the biologic effects of radiation, including cell killing, carcinogenesis and mutation (1). Incident ionizing particles can induce DNA damage through the interaction of primary and secondary particles (direct) as well as through the interaction of radiolysis products with the DNA target volume (indirect) within a time scale of femtoseconds (10^-15 s) to microseconds (10^-6 s). DNA double-strand breaks (DSBs) are recognized as the most important lesions caused by radiation and can, for example, be experimentally investigated with pulsed-field gel electrophoresis (PFGE), single-cell gel electrophoresis (also known as the comet assay) and DNA damage-induced nuclear protein foci (radiation-induced foci assay) (1).

The initial DNA damages will undergo different repair pathways in the nucleus and can be repaired accurately, misrepaired or unrepaired (residual) after irradiation. The misrepaired and residual DSBs will lead to genomic rearrangements within or between chromosomes, resulting in different types of chromosome aberrations that can drive cell death. Chromosome aberrations can be experimentally measured by staining mitotic chromosomes with Giemsa or fluorescent probes (2).

Theoretically, DNA damage can be simulated with the Monte Carlo (MC) method using track structure simulations. Several Monte Carlo tools, e.g., PARTRAC (3), KURBUC (4), PITS (5), NASIC (6), RITRACKS (7, 8), compared with experimental results. © 2020 by Radiation Research Society.
Geant4-DNA (9, 10), TOPAS-nBio (11) and GPU-based gMicroMC (12, 13), were developed to model physical, pre-chemical and chemical reactions in water, which can be regarded as a surrogate of biological tissue (14). The comprehensive track structure Monte Carlo code systems support simulation for photons, electrons, protons, neutrons and heavy ions (3, 9, 15). The combination of track structure simulations with nucleus geometry models is considered as the “gold standard” for predicting radiation-induced DNA damage spectrum (16). In addition to track structure Monte Carlo tools, the Monte Carlo damage simulation (MCDS) algorithm was developed for fast calculation of DNA lesions induced by various types of particles under different oxygen conditions (16–18). Such simulation tools are complementary to experimental studies, as they help interpret results and aid in generating hypotheses for further investigations.

To connect the initial DNA damage to chromosome aberrations, several in silico mechanistic DNA repair models were developed to describe the cellular processes related to DNA damage repair. The repair kinetics can be modeled with a set of differential equations or stochastic simulations. The biochemical kinetics model developed by Cucinotta et al. is one of the pioneering works that mathematically describes the binding of repair components (including Ku70/80, DNA-PKcs and the ligase IV-XRCC4 heterodimer) to DSBs with several intermediate repair complexes leading to DNA rejoining via the non-homologous end joining (NHEJ) pathway (19). Taleei et al. proposed a mathematical kinetic model that considers more identified proteins involved in the NHEJ pathway (20, 21). Friedland et al. modeled the NHEJ repair pathway with stochastic simulation of attachment and dissociation of involved repair enzymes and step-by-step diffusion of DNA ends; this model was incorporated into PARTRAC and presented the first simulation framework for track structure, DNA damage, repair and chromosome aberration kinetics (22–24). The DNA Mechanistic Repair Simulator (DaMaRiS) framework (25, 26), developed at the University of Manchester, describes the scheme of the NHEJ pathway with stochastic simulation, and this model was applied to predict the dependence of DNA damage repair fidelity on damage density (26). This model was further modified by considering the homologous recombination (HR) pathway, and the results suggested an entwined relationship between the repair choice of NHEJ and HR (27). The Mechanistic DNA Repair and Survival (MEDRAS) model developed by McMahon et al. (28) is another milestone of mechanistic modeling of cellular radiation response. The MEDRAS model is based on temporal and spatial modeling of DNA repair incorporating a Gaussian spatial dependence of DSB misrejoining; it is the first single model that considers three key repair processes [NHEJ, HR and microhomology mediated end joining (MMEJ) pathways] with different repair fidelities and underlying mechanisms across the cell cycle in mammalian cells. MEDRAS is able to predict key biological end points across a range of cell types, including repair kinetics, chromosome aberrations and cell survival. To facilitate cross-code comparison, both DaMaRiS and MEDRAS can read the standard DNA damage (SDD) data format (29), which is a standardized scoring method to link DNA damage induction simulations to the modeling of repair.

In this study, the cellular response after proton irradiation was investigated using track structure Monte Carlo simulations for damage induction and repair modeling with MEDRAS. A fibroblast nucleus model with a fractal geometry was built and implemented in TOPAS-nBio (11) for initial DNA damage yield simulations. The DNA damage yield results were scored in the SDD format and quantified by the yields of strand breaks (SBs), single-strand breaks (SSBs) and DSBs, and compared with published simulated and experimentally measured data. Furthermore, the dicentric and acentric fragment yields and micronuclei formation after proton irradiation were calculated and compared with experimental measurements by Edwards et al. (2), Wakatsuki et al. (30) and Yang et al. (31).

MATERIALS AND METHODS

DNA Model

A model of a human fibroblast nucleus at G0/G1 phase was built within the TOPAS-nBio framework. TOPAS-nBio is an extension to the TOPAS Monte Carlo application (32), which is based on Geant4 (33–35) to support simulation studies at the patient (macro and micro) scale (36, 37). The TOPAS-nBio extension is an advanced, yet user-friendly, Monte Carlo track structure simulation tool based on Geant4-DNA (9, 10). It offers an extensive library of advanced biological geometries ranging from the micrometer scale (e.g., cells and organelles) down to the nanometer scale (e.g., DNA molecules and proteins) (38) to support simulation studies at the cellular and sub-cellular levels with the goal of furthering our understanding of radiobiological effects at the DNA (nanometer) scale (38–40).

A full nucleus model was developed based on folding chromatin fibers in a fractal as indicated by Lieberman-Aiden et al. (41); the chromatin conformation is consistent with a knot-free fractal globule at the megabase scale. The fractal geometry of DNA was also adopted in other published work (42, 43). The DNA model was organized in a hierarchical structure including folding of the DNA double helix, nucleosomes, chromatin fibers and chromosomal territories.

DNA double helix. The DNA double helix (Fig. 1A) was formed with a half-cylindrical base volume with a radius of 0.5 nm and a quarter-cylindrical sugar-phosphate backbone volume with a radius of 1.15 nm based on the geometry in Henthorn et al. (44). The nucleotide pair has a height of 0.34 nm and was rotated by 36° for each subsequent pair. Additionally, a hydration shell with a thickness of 0.16 nm was built around the sugar-phosphate backbone. Ionizations within the shell can directly react with the DNA, causing direct DNA damage.

Nucleosome. The nucleosome (Fig. 1B) consists of a histone protein complex, modeled as a cylinder with a diameter of 6.6 nm and a length of 5.7 nm. The histone cylinder is wrapped with 1.65 left-handed turns (200 pairs of nucleotides) of the DNA double helix. The nucleosomes are connected to each other with 99 pairs of nucleotides, including linker DNA following a Bezier curve to ensure a smooth connection (44), forming a chromatin fiber.

Chromatin fiber, fiber loop and voxel. The chromatin fiber (Fig. 1C) has a radius of 37.1 nm and a length of 120 nm. The helix of...
nucleosomes was built around the central axis of the fiber and set to repeat every six nucleosomes. Each fiber consists of 51 nucleosomes and contains 15,150 base pairs (bp) of DNA. Chromatin fibers were folded according to a continuous 3D Hilbert space-filling curve (38) to form chromatin fiber loops and fill a voxel (Fig. 1D). One chromatin fiber loop has seven chromatin fibers and was formed by a single iterated Hilbert curve, and four chromatin fiber loops were placed next to each other to use the space of the voxel efficiently. Each voxel has a side length of 0.3 μm and contains 0.42 Mbp of DNA.

Chromosomes and nucleus. The nucleus (Fig. 1E) has a spherical shape with a diameter of 9.3 μm and contains 46 chromosomes. The number of voxels and DNA content comprising each chromosome are summarized in Table 1. The full nucleus contains 14,328 voxels and 6.08 Gbp DNA, giving a final DNA density of 14.4 Mbp/μm$^3$. The material of the nucleus is set to G4_WATER with a density of 1 g/cm$^3$, except for the backbone volume for which the density of G4_WATER is set to 1.407 g/cm$^3$ ($\frac{29}{38}$m) to support simulations of protons solely up to 100 MeV (9). The ionization cross section in the TsEmDNAPhysics module is extrapolated to support the simulation of protons up to 500 MeV (11). Another advantage of the TsEmDNAPhysics module is that users can easily select the physics models of each interaction process with several command lines in the TOPAS-nBio parameter file. The recommended physics list of TOPAS-nBio for use in water radiolysis simulations (11) is contained in the physics module TsEmDNAPhysics. Most of the direct DNA damages after proton irradiation are caused by secondary electrons (>80%; see Appendix Fig. A1). The TsEmDNAPhysics module uses the following Geant4-DNA models to describe the interaction processes for electrons: G4DNACrystalBallModel for elastic scattering processes (7.4 eV–1 MeV); G4DNABornExcitationModel for electronic excitation processes (9 eV–1 MeV); G4DNABornAttachmentModel for ionization processes (11 eV–1 MeV); and G4DNASancheExcitationModel for vibrational excitation processes (2 eV–100 eV); and G4DNAMeltonAttachmentModel for molecular attachment processes (4 eV–13 eV). The ionization cross section in the TsEmDNAPhysics module was extrapolated to support the simulation of protons up to 500 MeV while the default Geant4-DNA physics model (G4EmDNAPhysics) supports simulations of protons solely up to 100 MeV (9). The TsEmDNAPhysics module uses G4DNABornAttachmentModel and cross-section data-adopted G4DNAPhysics for protons within 100 MeV and uses G4DNARuddAttachmentModel and published cross-section data from RITRACKS (8) to support simulations of protons up to 500 MeV. Another advantage of the TsEmDNAPhysics module is that users can easily select the physics models of each interaction process with several command lines in the TOPAS-nBio parameter file.

The initial DNA damage after proton irradiation (0.5–500 MeV, corresponding to the LET region of 60–0.2 keV/μm) was simulated with TOPAS-nBio (11). The nucleus model was placed at the center of a cubic world with a side length of 14 μm, large enough to contain the nucleus, which was filled with water with a density of 1 g/cm$^3$. To avoid a systematic alignment of proton tracks with the chromatin fiber, primary protons were initiated randomly on the surface of the nucleus and propagated inside the nucleus with a random direction. The induced DNA damage caused by direct and indirect interactions in the physical and chemical stages was quantified as either SBs, SSBs or DSBs (see below) and was output in the standard DNA damage (SDD) data format (29). To obtain sufficient statistics, 100 runs were performed for each energy point. Each simulation run had a fixed number of primaries and deposited a dose of 1 Gy within the nucleus. The statistical uncertainty of dose and DSB yield were smaller than 2%.

The track-length averaged LET was recorded as an index of radiation quality, and it was calculated with:

$$\text{LET} = \frac{\varepsilon}{d},$$

where $d$ is the mean proton cord length within the nucleus and $\varepsilon$ is the energy deposition of all the primary and secondary particles within the nucleus.

Physical stage simulation. Direct DNA damage was induced from physical interactions between the primary and secondary particles within the DNA target. The recommended physics list of TOPAS-nBio for use in water radiolysis simulations (11) is contained in the physics module TsEmDNAPhysics. Most of the direct DNA damages after proton irradiation are caused by secondary electrons (>80%; see Appendix Fig. A1). The TsEmDNAPhysics module uses the following Geant4-DNA models to describe the interaction processes for electrons: G4DNACrystalBallModel for elastic scattering processes (7.4 eV–1 MeV); G4DNABornExcitationModel for electronic excitation processes (9 eV–1 MeV); G4DNABornAttachmentModel for ionization processes (11 eV–1 MeV); and G4DNASancheExcitationModel for vibrational excitation processes (2 eV–100 eV); and G4DNAMeltonAttachmentModel for molecular attachment processes (4 eV–13 eV). The ionization cross section in the TsEmDNAPhysics module was extrapolated to support the simulation of protons up to 500 MeV while the default Geant4-DNA physics model (G4EmDNAPhysics) supports simulations of protons solely up to 100 MeV (9). The TsEmDNAPhysics module uses G4DNABornAttachmentModel and cross-section data-adopted G4DNAPhysics for protons within 100 MeV and uses G4DNARuddAttachmentModel and published cross-section data from RITRACKS (8) to support simulations of protons up to 500 MeV. Another advantage of the TsEmDNAPhysics module is that users can easily select the physics models of each interaction process with several command lines in the TOPAS-nBio parameter file.

Only interactions with the backbone led to DNA damage, i.e., base damage did not contribute to the recorded DNA damage. While interactions between primary and secondary particles on the backbone volumes were the main cause of direct SBs, it was assumed that
**TABLE 1**

<table>
<thead>
<tr>
<th>Chromosome ID</th>
<th>Voxel Number</th>
<th>DNA base pairs (Mbp)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1, 2</td>
<td>596</td>
<td>253</td>
</tr>
<tr>
<td>3, 4</td>
<td>585</td>
<td>248</td>
</tr>
<tr>
<td>5, 6</td>
<td>481</td>
<td>204</td>
</tr>
<tr>
<td>7, 8</td>
<td>461</td>
<td>196</td>
</tr>
<tr>
<td>9, 10</td>
<td>436</td>
<td>185</td>
</tr>
<tr>
<td>11, 12</td>
<td>412</td>
<td>175</td>
</tr>
<tr>
<td>13, 14</td>
<td>383</td>
<td>162</td>
</tr>
<tr>
<td>15, 16</td>
<td>352</td>
<td>149</td>
</tr>
<tr>
<td>17, 18</td>
<td>338</td>
<td>143</td>
</tr>
<tr>
<td>19, 20</td>
<td>326</td>
<td>138</td>
</tr>
<tr>
<td>21, 22</td>
<td>324</td>
<td>137</td>
</tr>
<tr>
<td>23, 24</td>
<td>319</td>
<td>135</td>
</tr>
<tr>
<td>25, 26</td>
<td>275</td>
<td>117</td>
</tr>
<tr>
<td>27, 28</td>
<td>256</td>
<td>109</td>
</tr>
<tr>
<td>29, 30</td>
<td>242</td>
<td>103</td>
</tr>
<tr>
<td>31, 32</td>
<td>214</td>
<td>91</td>
</tr>
<tr>
<td>33, 34</td>
<td>189</td>
<td>80</td>
</tr>
<tr>
<td>35, 36</td>
<td>183</td>
<td>78</td>
</tr>
<tr>
<td>37, 38</td>
<td>154</td>
<td>65</td>
</tr>
<tr>
<td>39, 40</td>
<td>150</td>
<td>64</td>
</tr>
<tr>
<td>41, 42</td>
<td>113</td>
<td>48</td>
</tr>
<tr>
<td>43, 44</td>
<td>119</td>
<td>50</td>
</tr>
<tr>
<td>Y</td>
<td>139</td>
<td>59</td>
</tr>
<tr>
<td>X</td>
<td>373</td>
<td>158</td>
</tr>
<tr>
<td>Total</td>
<td>14,328</td>
<td>6,078</td>
</tr>
</tbody>
</table>

*Based on data from the Genome Reference Consortium (GRC) www.ncbi.nlm.nih.gov/grc. The nucleus contains 14,328 voxels and 6.08 Gbp DNA (1.02 times the reference data); therefore, the DNA content of each chromosome was scaled by a factor of 1.02 to assign voxels to chromosomes.

ionizations in the hydration shell could also lead to direct damage (47–49). A direct strand break (SBdir) was formed if at least 17.5 eV of deposited energy was accumulated within the backbone volume and neighboring hydration shell within a single history.

**Pre-chemical and chemical stage simulation.** The TsEmDNA-Chemistry model (40) was used to simulate the chemical interactions in the pre-chemical stage and chemical stage. The TsEmDNAChemistry model inherited the capabilities to simulate water radiolysis of Geant4-DNA (50, 51), but used reviewed and updated chemistry parameters. These parameters included thermalization distance of solvated electron, diffusion coefficients and reaction rates in the G4EmDNAChemistry model. The TsEmDNAChemistry model was validated by calculating time-dependent G values (number of G4EmDNAChemistry model. The TsEmDNAChemistry model was validated with experimental data. The model aims at predicting key biological end points, including repair kinetics, chromosome aberrations and cell survival, among others, across a range of cell types. Recently, this model was updated by separating the DNA fragment end rejoining and foci clearance and refitted with additional datasets for more realistic description of DNA repair processes. The updated end rejoining coefficients of NHEJ (fast) and HR (slow) repair processes were 2.07 ± 0.17/h and 0.26 ± 0.01/h, respectively. The newest version of the model is available as open source on GitHub: https://github.com/sjmcmahon/Medras-MC.

The model starts from the initial DNA damages recorded in the SDD data format (29). Free DNA ends interact with other free DNA ends with a Gaussian separation-dependent probability (56):

$$\zeta(d) \propto e^{-\frac{d^2}{2R^2}},$$

(2)

where d is the separation between the two ends, σ is a characteristic rejoining range with a fitted value of σ = 0.046Rmax, and Rmax is the radius of the nucleus. The characteristic rejoining range (~0.21 μm) is relatively small compared to other models that focus on chromosome aberrations as a main end point [1 μm or more (56, 57)]. This is...
In this study, the number of dicentric fragments (dicentric (multi-centromere) and acentric (centromere free) fragments) result in different types of chromosome aberrations, including genomic rearrangements within or between chromosomes and may without a centromere. The misrepair of DNA fragments leads to then the centromere was at the 50 Mbp position. One DSB splits the MEDRAS model, i.e., if the chromosome had a length of 100 Mbp to the binary misrepair events to give the total yield of misrepair. This simulation considered all events until a given repair time, at which point DSB free ends may have been correctly repaired, misrepaired (misrejoin with free ends from other DSBs) or remained unrepaired.

Because DNA repair processes such as NHEJ do not have perfect fidelity, in addition to these binary misrepair events involving two DSBs, there is also a small chance of even isolated DSBs being incorrectly misrepaired. Previously published analytic models, in which this effect was quantified, obtained a fidelity of 98.54% for NHEJ, or equivalently a 1.46 ± 0.2% chance that such a break will be misrepaired (28). These additional misrepair events can then be added to the binary misrepair events to give the total yield of misrepair. This misrepair makes a negligible contribution at high doses and LETs, but is potentially more significant in low-dose low-LET regions.

Chromosome aberration. Each of the chromosomes had a centromere which was placed at the center of each chromosome in the MEDRAS model, i.e., if the chromosome had a length of 100 Mbp then the centromere was at the 50 Mbp position. One DSB splits the chromosome into two DNA fragments, one with a centromere and one without a centromere. The misrepair of DNA fragments leads to genomic rearrangements within or between chromosomes and may result in different types of chromosome aberrations, including dicentric (multi-centromere) and acentric (centromere free) fragments. In this study, the number of dicentric fragments (N_{DF}) and acentric fragments (N_{AF}) was counted after the repair process, based on the number of centromere(s) in the fragments after calculation of binary misrepair events. The impact of centromere position on the formation of acentric and dicentric fragments was investigated by modeling the placement of the centromere at different positions along the chromosome (10%, 30% and 50%). Our results show that the impact is less than 5% (data not shown).

Micronuclei (MN) mainly originate from acentric fragments or whole chromosomes that fail to be included in the daughter nuclei at the completion of telophase during mitosis and are correlated to cell killing (59–61). Experimentally, the formation of MN is usually scored in cytochalasin-blocked binucleated (BN) cells, which are cells attempting their first mitosis after irradiation. The percentage of BN cells with MN is usually used to describe the MN response to dose. The radiation dose response of BN cells with MN in the range 0–4 Gy is generally considered linear-quadratic for low-LET X rays and becomes linear for high-LET particles like protons and other heavy charged particles (59). MN can also be found in nonirradiated cells causing a background value of the percentage of BN cells with MN.

According to the origination of the MN, the number of MN can be analytically calculated with:

\[ N_{MN} = p_1 N_{AF} + p_2 N_{WC}, \]

where, \( N_{AF} \) and \( N_{WC} \) are the number of acentric fragments and whole chromosomes, and \( p_1 \) and \( p_2 \) are the probability of acentric fragments and whole chromosomes that escape from the nucleus and form MN.

Because the MEDRAS model considers all forms of misrepair, not only events that lead to chromosome aberrations, and many of these other types of misrepair occur over shorter ranges. The rejoining range in this case is in line with those of other models that focus on similar results [e.g., 70 nm (26) to 0.3 μm (58)]. For DSB breaks distributed uniformly within a spherical nucleus, the correct rejoin probability is:

\[ P_{\text{correct}} = \frac{1 - e^{-\eta}}{\eta}, \]

where \( \eta = \sum_{i=1}^{N-1} \frac{1}{2} \gamma_i \) is the sum of the interaction rates between a free end in a given DSB with the remaining \( N - 1 \) DSBs (28). This expression is appropriate for low-LET irradiations such as X rays.

For the non-uniform damage distributions created by ions in this work, the misrejoin process was stochastically simulated by randomly sampling interactions between break ends, weighted according to the process-specific repair rates and the distance-dependent scaling in Eq. (2). This simulation considered all events until a given repair time, at which point DSB free ends may have been correctly repaired, misrepaired (misrejoin with free ends from other DSBs) or remained unrepaired.

In this study, it was assumed that \( N_{AF} = 0 \) and \( N_{WC} = 46 \) for nonirradiated nuclei, i.e., only MN originating from whole chromosomes contribute to the background MN formation. Experimentally measured background MN levels vary over different studies (30, 31, 62). To better compare with experimental data, \( p_2 \), was set to 0 for irradiated nuclei and only MN originating from acentric fragments were considered. Predictions were compared to experimentally measured data after subtraction of background MN data.

Additional fragments resulting from the limited fidelity of NHEJ were incorporated in the calculation of MN yields after low-LET irradiation, where this contribution is expected to be relatively large. The formation of micronuclei was scored as the percentage of BN cells with MN, meaning that the number of cells with at least one MN was calculated. When cells are exposed to low-LET radiation at low dose, most of the initial DSBs can be correctly repaired and most of the cells may not contain any aberration (acentric, dicentric or MN); the number of BN cells with more than one MN can be underestimated if the repair fidelity of the NHEJ pathway is not considered, leading to significant difference with experimental data. The aberration yield per cell increases with LET as well as with dose, and the impact of repair fidelity gradually disappears when each cell has at least one MN on average. The number of additional acentric fragments can be empirically calculated with:

\[ N_{AF} = p_3 p_4 N_{DSB} D, \]

where \( N_{AF} \) is the number of additional acentric fragments, \( p_1 = 1.46 \) ± 0.2% is the base misrepair rate (28), and \( p_3 \) is the probability that one misrepair event leads to an acentric fragment above the chosen size threshold, for single-DSB misrepair events which are associated with the limited fidelity of NHEJ. \( N_{DSB} \) is the number of DSBs per nucleus per Gy, and \( D \) is the dose in the nucleus.

RESULTS AND DISCUSSION

Proton LET as a Function of Energy

Figure 3 shows the relationship between proton LET and particle energy simulated with TOPAS-nBio and digitalized from references. In low-energy regions, the discrepancy between different published data [maximum relative difference is 32.5% at 0.5 MeV between our calculation and Meylan et al. (47)] is mainly due to the size of the scoring volume, since low-energy protons probably cannot transverse the entire nucleus (e.g., the range of 0.5 MeV...
protons is \( \approx 8.7 \, \mu m \) in liquid water). The impact of scoring volume on track-averaged LET was tested by scoring with the same nucleus geometry adopted by Meylan et al. (47) (an elliptical nucleus with a half axis of 2.5 \( \mu m \) along particle incident direction) and the relative difference was reduced to 2%. For the high-energy region (>100 MeV), the LET values calculated with TOPAS-nBio agree by 4% with Friedland et al. (48).

**Initial DNA Damage**

Figure 4 shows the DNA damage results as a function of proton LET simulated with TOPAS-nBio. Simulated direct SB and SSB yields (Fig. 4A and B) are generally flat in the low-LET region. The direct SB yield slightly increases with LET in the higher LET region (>10 keV/\( \mu m \)) and the direct SSB yield decreases slightly in high-LET regions because more SBs are categorized as DSBs. For indirect breaks, SB and SSB yields both initially increase with LET, before flattening at approximately 5 keV/\( \mu m \) and then falling at higher LETs. For DSB damages (Fig. 4C), the direct DSB yield increases with the proton LET due to the denser local energy depositions of primary protons and secondary particles. The indirect DSB yield first increases with proton LET and then saturates at higher LET due to more chemical species being generated locally and interacting with each other, rather than with biological targets.

Figure 4D shows the relative contribution of indirect and hybrid damages as a fraction of each break type. Our simulations predicted that most of the DSB damages were classified as hybrid type, caused by the combined effect of direct and indirect damage.

**DNA Damage Repair**

Figure 6 shows characteristics of DNA damage repair kinetics using the MEDRAS model (28, 55), after simulations of the physical and chemical stages. Figure 6A shows the residual DSB fraction as a function of repair time. It was predicted that more than 95% of the DSBs would be repaired within the first 24 h. The repair processes for DNA damages caused by high-LET protons are slower due to more complex damage (DSBs located close to each other) that are more difficult to repair. Figure 6B shows the predicted misrepair and residual DSB fractions after a 24-h repair process period as a function of the proton LET.
residual DSB fraction is initially flat (~1%) in low-LET regions and increases slightly at higher LETs, reaching up to 3.3% at 60.1 keV/µm (0.5 MeV proton). This delay in repair arises because rapid DSB repair requires other available free DSB ends within a relatively short separation. At high rates of misrepair, some DSB ends can become isolated and will take longer to rejoin with other distant DSB ends. For calculation of misrepair fractions, different DNA fragment detection thresholds were applied to exclude short DNA fragments for better quantification of misrepair fractions in experiments, since many assays are not sensitive to the tiny inversions or deletions caused by misrepair of short DNA fragments. The predicted misrepair fractions increase rapidly with LET and reach 63.7% at 60.1 keV/µm without a fragment detection threshold. This value dropped to 15.8% with an estimated chromosome aberration detection threshold of 3 Mbp (69).

**Chromosome Aberrations**

The prediction of dicentric and acentric fragment yields in cells irradiated with 8.7 MeV protons and cultured 48 h postirradiation were compared with experimentally measured results in Fig. 7. It is important to note that a detection

![FIG. 5. Comparison of DNA damage obtained with TOPAS-nBio (black squares, dots, solid and dashed lines) and other simulated (diamonds, right triangles, solid and dashed lines) or experimentally measured data (up triangles). Panel A: SB yield per Gy per Gbp DNA. Panel B: SSB yield per Gy per Gbp DNA. Panel C: DSB yield per Gy per Gbp DNA. The experimental values measured by Frankenberg et al. (67), Campa et al. (64) and Belli et al. (66, 68) were generally measured using the PFGE method with different detection thresholds across different cell lines. Panel D: Ratio of SSB yield to DSB yield.](image)

![FIG. 6. Characteristics of DNA damage repair. Panel A: Residual DSB fractions as a function of repair time postirradiation with 0.5/1/20 MeV protons. Panel B: Misrepair or residual DSB fractions as a function of proton LET after 24 h of the repair process.](image)
A threshold of 3 Mbp was applied to calculate the number of detectable acentric fragments. Acentric fragments have a wide range of sizes, from tens of bp to Mbp, and many, particularly the smaller acentrics, may be undetectable. Acentric fragments in experiments were defined as excess acentric fragments, which were acentrics after allowing one for each dicentric fragment and centric ring (70). Therefore, the number of excess acentric fragments was calculated with the number of detectable acentric fragments subtracted by the number of dicentric fragments and the number of centric rings.

Predicted dicentric fragment yields (solid squares in Fig. 7A) and excess acentric fragment yields (solid squares in Fig. 7B) show a consistent quasi-linear trend with experimental data (open squares) throughout the investigated dose region, but with higher absolute values: dicentric fragment yields were overestimated by a factor of approximately 2 and excess acentric fragment yields were overestimated by a factor of approximately 10. There are a number of possible explanations for this: Experimental measurements can underestimate absolute values due to a detection threshold or resolution limitation, and the sensitivity of the Giemsa staining method ranges widely and depends on technical and examiner skill (71). Unlike dicentric fragments, acentric fragments have a wide range of sizes and many may be undetectable; this could lead to higher overestimation of excess acentric fragment yields than dicentric fragment yields.

Furthermore, TOPAS-nBio predicted a higher DSB yield than experimental results (Fig. 5C), which subsequently resulted in higher predicted chromosome aberration yields since the dicentric and acentric fragments originated from the misrepair of DSBs. The impact of DSB yields on chromosome aberration yields was quantified by calculating dicentric and excess acentric fragment yields with only direct DSBs (which were ~30% of total DSBs, as shown in Fig. 4C), i.e., initial DNA damages were simulated with physics stage only. The resulting dicentric and excess acentric fragment yields (solid upside-down triangles in Fig. 7) were decreased by a factor of 10.

Additionally, the whole nucleus model as well as the repair model were both developed for fibroblasts, while experimental measurements were performed with lymphocytes, which are more prone to radiation-induced apoptosis (72, 73). In experiments, cells were assayed 48 h postirradiation, analyzing first division metaphases. Thus, cells that died in interphase (e.g., by apoptosis) or were arrested and had not reached mitosis would not be counted. Many of the large-sized acentric fragments are due to residual damages; cells, especially lymphocytes, with such damages may not make it to mitosis, leading to potentially significant differences in counts. In this work, the fibroblast nucleus model was used to represent lymphocytes as a simplifying assumption, while, actually, lymphocytes have different geometry characteristics. For instance, Friedland et al. developed different models for fibroblasts and lymphocytes (74). We have performed preliminary investigations (data not shown) of the impact of nucleus shape, and found that total rates of misrepair are not significantly affected until one dimension of the nucleus is very small (<~1 μm), which is not the case for cell types considered here.

Figure 8 shows the comparison between predicted dose response of BN cells with MN and experimental data. The experiments were performed delivering 1 GeV protons with a reported LET of 0.2–0.24 keV/μm. Our prediction was based on protons of 500 MeV (LET of 0.2 keV/μm), which is the highest proton energy that can be simulated with the current TOPAS-nBio version. The probability \( p_f \) of acentric fragments that escape from the nucleus to form a MN was set to 0.5 (75). The predicted dose response of BN
cells with MN with a detection threshold of 3 Mbp (solid triangles connected with solid line) generally shows a linear trend throughout the whole dose region, which is consistent with that indicated by Durante et al. (59). The predictions from the binary misrepair model are lower than experimental data, especially in the low-dose region. A potential reason for the lower predicted level in the lower dose region is that the experiments by Yang et al. (31) were performed at very low-dose levels (down to $6.25 \times 10^{-8}$ Gy), where the MN response may be dominated by bystander signaling or the cells may fail to detect low levels of DNA damage and activate repair processes (76). Additionally, micronuclei may be more visible compared to acentric fragments, as they are wrapped in membrane; thus, the detection threshold for micronuclei may be lower.

Micronuclei yields from all acentric fragments, including binary misrepair and additional single-DSB misrepair, were considered with a detection threshold of 3 Mbp (solid triangles connected with dashed line in Fig. 8) and a lower detection threshold, 10 kbp (solid squares connected with dashed line in Fig. 8). For this work, $p_4$, was estimated to be 0.24 for acentric fragments larger than 3 Mbp and 0.41 for acentric fragments larger than 10 kbp after irradiation with 500 MeV protons. As shown, the dose response of BN cells with MN in the low-dose region was significantly corrected with a detection threshold of 10 kbp, agreeing within statistical uncertainties for all doses above 0.1 Gy (MN yield above 1%).

**SUMMARY AND CONCLUSION**

The cellular response after proton irradiation was studied by simulating the initial DNA damage in a whole nucleus model with TOPAS-nBio and using the MEDRAS mechanistic repair model to predict the DNA damage repair processes and subsequent chromosome aberrations.

A whole nucleus model with a fractal DNA geometry, including the hierarchical structure of the genome from the DNA double helix, nucleosomes, chromatin fibers and chromosomal territories, was built and implemented in TOPAS-nBio. The spherical nucleus had a diameter of 9.3 μm and consisted of 46 chromosomes with a total DNA content and density of 6.08 Gbp and 14.4 Mbp/μm³, respectively. The straight chromatin fibers and voxelized geometry used in this work were not as complex as the models developed in other published studies (15, 48). However, our simplified geometry significantly reduced the required computational memory and accelerated the navigation in Geant4 (as well as TOPAS); the initialization of the whole nucleus geometry can be finished within several minutes and a single simulation run with a 1 Gy dose in the nucleus generally can be completed within 10 h using a multi-thread (10 threads) mode [with Intel(R) Xeon(R) CPU L5640/X5660/E5450@2.27–3.00GHz]. In our current nucleus model the chromatin fibers were not smoothly connected. However, the impact of unlinked chromatin fibers should be negligible since the missing geometries are only a few base pairs every $\sim 150$ kbp.

The initial DNA damage yields induced by incident protons were simulated by modeling both the physical and chemical interactions within the nucleus with the default process models available in TOPAS-nBio. The initial SBs were found to be mostly caused by indirect damages (60–75%) and most of the DSBs were caused by the combined effect of direct and indirect damages. The total DSB yield increased from 6.5 DSB/Gy/Gbp in low-LET regions (0.2
keV/μm) to 21.2 DSB/Gy/Gbp in the high-LET region (60.1 keV/μm). With an increase in proton LET, more complex DNA damages and short DNA fragments were formed, with DSBs leading to detectable DNA fragments (>10 kbp) showing a saturation in the high-LET region.

By applying the mechanistic repair model, the characteristics of DNA damage repair and the dose response of chromosome aberrations were obtained. It was found that more than 95% of the DSBs are repaired within the first 24 h, and the highest misrepair fraction was 15.8% at 60 keV/μm with an estimated chromosome aberration detection threshold of 3 Mbp. The number of chromosome aberrations after 48 h of repair time was also obtained and the predicted dicentric and excess acentric fragment yields shared consistent quasi-linear trends with experimental data, but with higher absolute values which can be attributed to the limitation of detection resolution in experiments, overestimation of initial damages and count loss due to experimental procedures. The dose responses of BN cells with MN were calculated and our results show that the non-zero misrepair rate has a significant impact on the predicted yield at low doses and LETs, and that a 10 kbp detection threshold may be more appropriate for predictions of micronuclei formation.

Our simulations show that Monte Carlo tools can predict DNA damage, and by incorporating a repair model, the DNA damage repair characteristics and chromosome aberration yields can be obtained. Such simulations can be employed to interpret experimental data and design new hypotheses. All simulations were done with the user-friendly framework of TOPAS-nBio, which does not require programming knowledge and has been made available as open-source code. Our results were recorded in the SDD data format, which facilitates inter-model comparisons, and can be easily used for related research benchmarks.

**APPENDIX**

Table A1 provides a summary of simulation setup, and Table A2 shows the damage yield (per Gy per Gbp) after proton (0.5–500 MeV) irradiation obtained using TOPAS-nBio. Figure A1 shows the energy deposition contribution in the DNA backbone and hydration shell from secondary electrons in the physics stage. Simulations were performed using the TsEmDNAPhysics model for different primary proton energies.

**ACKNOWLEDGMENTS**

HZ was sponsored by the China Scholarship Council for a year-long study abroad program at Massachusetts General Hospital. NH is supported by the European Union’s Horizon 2020 Research and Innovation program under grant agreement no. 730983. This work was supported by the National Cancer Institute (grant no. R01 CA187003: TOPAS - nBio, a Monte Carlo Tool for Radiation Biology Research). Simulations were performed on the O2 High Performance Compute Cluster, supported by the Research Computing Group at Harvard Medical School, and the ERISOOne cluster, supported by Partners Information Services at Partners Healthcare Systems.

Received: September 30, 2019; accepted: April 11, 2020; published online: May 13, 2020

**REFERENCES**


59. Durante M, Formenti SC. Radiation-induced chromosomal...
aberrations and immunotherapy: Micronuclei, cytosolic DNA, and interferon-production pathway. Front Oncol 2018; 8:192.


### TABLE A1

<table>
<thead>
<tr>
<th>Simualtion tool</th>
<th>TOPAS-nBio</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source</td>
<td>0.5–500 MeV protons. Randomly sampled on the surface of the nucleus, pointing inside the nucleus with a random direction</td>
</tr>
<tr>
<td>Number of histories</td>
<td>1,000–250,000 primaries (depending on proton energy)</td>
</tr>
<tr>
<td>Simulation time</td>
<td>&lt;10 h to finish a run with a dose of 1 Gy within the nucleus [10 threads mode on Intel(R) Xeon(R) CPU L5640/X5660/E5450@2.27–3.00GHz]</td>
</tr>
<tr>
<td>Statistical uncertainty</td>
<td>DNA damage yield uncertainty &lt;2%</td>
</tr>
<tr>
<td>Geometry</td>
<td>Spherical fibroblast nucleus in G1/G0 phase, 9.3 μm in diameter, water, density = 1 g/cm³ (except backbone density = 1.4 g/cm³)</td>
</tr>
<tr>
<td>Physics and chemistry model</td>
<td>TsEmDNAPhysics, TsEmDNAChemistry (1 ns chemical stage)</td>
</tr>
<tr>
<td>Induction of direct SB</td>
<td>&gt;17.5 eV energy deposition in the backbone and neighboring hydration shell</td>
</tr>
<tr>
<td>Induction of indirect SB</td>
<td>Assuming only HOH can cause indirect damage with a probability of 0.4</td>
</tr>
<tr>
<td>Induction of DSB</td>
<td>2 SB on opposite strands of DNA, located within 10 base pairs</td>
</tr>
</tbody>
</table>
FIG. A1. Energy deposition contribution in the DNA backbone and hydration shell from secondary electrons in the physics stage. Simulations were performed using the TsEmDNAPhysics model for different primary proton energies.

### TABLE A2

Damage Yield (per Gy per Gbp) after Proton (0.5–500 MeV) Irradiation Obtained with TOPAS-nBio

<table>
<thead>
<tr>
<th>Energy (MeV)</th>
<th>LET (keV/μm)</th>
<th>Total SB</th>
<th>Direct SB</th>
<th>Indirect SB</th>
<th>Total SSB</th>
<th>Direct SSB</th>
<th>Indirect SSB</th>
<th>Total DSB</th>
<th>Direct DSB</th>
<th>Indirect DSB</th>
<th>Hybrid DSB</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.5</td>
<td>60.09</td>
<td>134.07</td>
<td>47.90</td>
<td>86.17</td>
<td>108.91</td>
<td>35.29</td>
<td>73.62</td>
<td>21.21</td>
<td>6.20</td>
<td>5.46</td>
<td>9.55</td>
</tr>
<tr>
<td>0.6</td>
<td>47.62</td>
<td>149.68</td>
<td>50.73</td>
<td>98.95</td>
<td>123.76</td>
<td>38.41</td>
<td>85.35</td>
<td>19.60</td>
<td>5.00</td>
<td>5.52</td>
<td>9.07</td>
</tr>
<tr>
<td>0.8</td>
<td>35.77</td>
<td>170.06</td>
<td>53.74</td>
<td>116.33</td>
<td>143.46</td>
<td>41.75</td>
<td>101.71</td>
<td>17.42</td>
<td>3.99</td>
<td>5.47</td>
<td>7.96</td>
</tr>
<tr>
<td>1</td>
<td>29.18</td>
<td>176.34</td>
<td>53.39</td>
<td>122.95</td>
<td>150.61</td>
<td>42.08</td>
<td>108.53</td>
<td>15.98</td>
<td>3.50</td>
<td>5.25</td>
<td>7.23</td>
</tr>
<tr>
<td>1.5</td>
<td>20.95</td>
<td>191.15</td>
<td>56.02</td>
<td>135.13</td>
<td>166.31</td>
<td>45.26</td>
<td>121.04</td>
<td>14.37</td>
<td>3.00</td>
<td>4.82</td>
<td>6.55</td>
</tr>
<tr>
<td>2</td>
<td>16.72</td>
<td>195.57</td>
<td>55.82</td>
<td>139.75</td>
<td>171.87</td>
<td>45.72</td>
<td>126.14</td>
<td>13.39</td>
<td>2.74</td>
<td>4.66</td>
<td>5.99</td>
</tr>
<tr>
<td>5</td>
<td>8.20</td>
<td>204.01</td>
<td>54.62</td>
<td>149.39</td>
<td>183.31</td>
<td>46.00</td>
<td>137.31</td>
<td>11.14</td>
<td>2.22</td>
<td>4.01</td>
<td>4.91</td>
</tr>
<tr>
<td>10</td>
<td>4.64</td>
<td>207.33</td>
<td>55.18</td>
<td>152.15</td>
<td>188.28</td>
<td>47.05</td>
<td>141.22</td>
<td>10.25</td>
<td>2.21</td>
<td>3.60</td>
<td>4.44</td>
</tr>
<tr>
<td>20</td>
<td>2.53</td>
<td>205.36</td>
<td>54.96</td>
<td>150.40</td>
<td>187.23</td>
<td>47.23</td>
<td>140.00</td>
<td>9.76</td>
<td>2.02</td>
<td>3.41</td>
<td>4.33</td>
</tr>
<tr>
<td>50</td>
<td>1.13</td>
<td>192.27</td>
<td>54.84</td>
<td>137.44</td>
<td>175.69</td>
<td>47.48</td>
<td>128.21</td>
<td>8.80</td>
<td>1.95</td>
<td>2.86</td>
<td>3.98</td>
</tr>
<tr>
<td>200</td>
<td>0.34</td>
<td>155.02</td>
<td>57.71</td>
<td>97.31</td>
<td>141.73</td>
<td>51.20</td>
<td>90.53</td>
<td>7.07</td>
<td>1.84</td>
<td>1.95</td>
<td>3.28</td>
</tr>
<tr>
<td>500</td>
<td>0.20</td>
<td>143.96</td>
<td>58.31</td>
<td>85.65</td>
<td>131.72</td>
<td>51.92</td>
<td>79.80</td>
<td>6.52</td>
<td>1.89</td>
<td>1.59</td>
<td>3.04</td>
</tr>
</tbody>
</table>