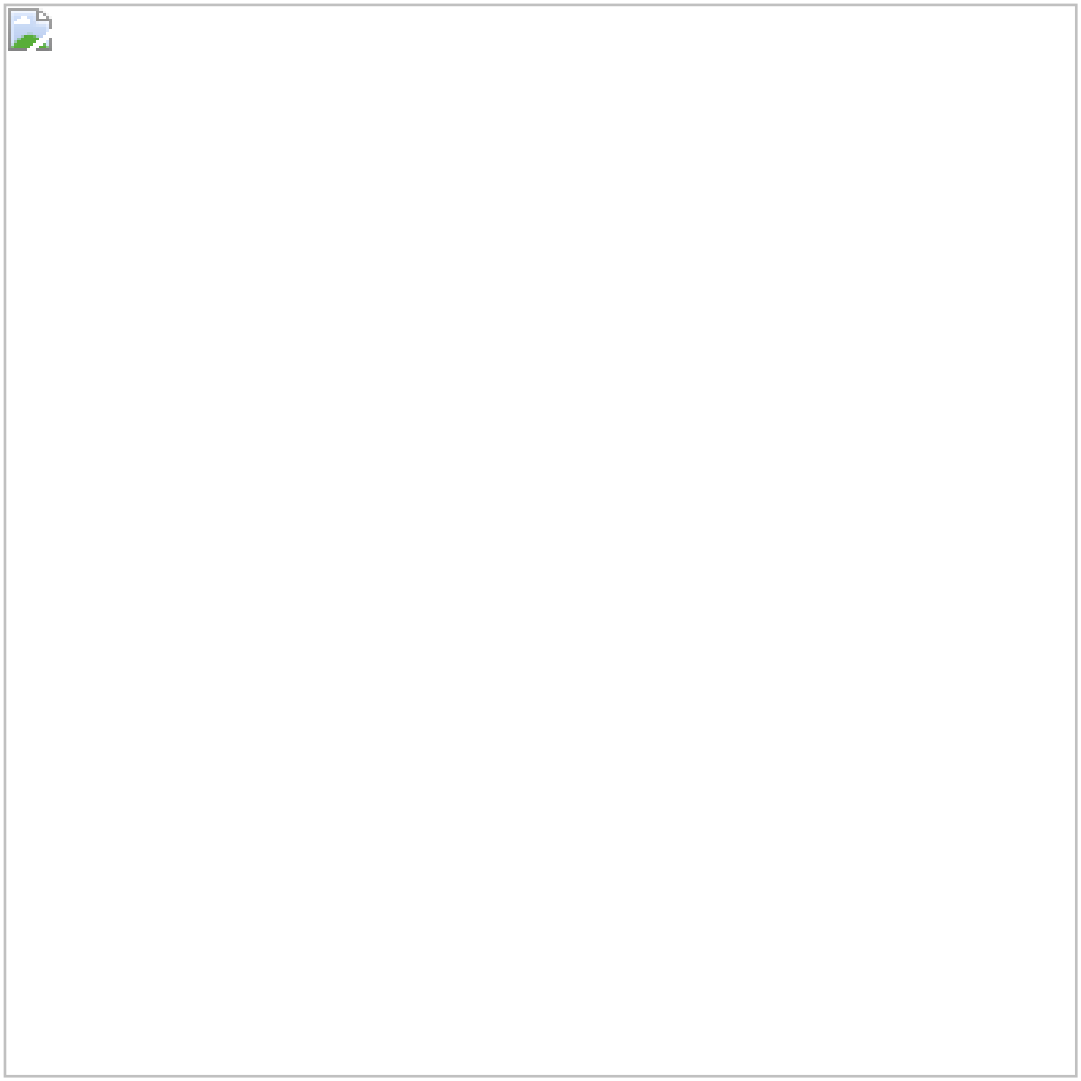

TOPAS-nBio: An Extension to the TOPAS Simulation Toolkit for Cellular and Sub-cellular Radiobiology

Authors: Schuemann, J., McNamara, A. L., Ramos-Méndez, J., Perl, J., Held, K. D., et al.

Source: Radiation Research, 191(2) : 125-138

Published By: Radiation Research Society

URL: https://doi.org/10.1667/RR15226.1

RADIATION RESEARCH **191**, 125–138 (2019)
0033-7587/19 $15.00

DOI: 10.1667/RR15226.1

TOPAS-nBio: An Extension to the TOPAS Simulation Toolkit for Cellular and Sub-cellular Radiobiology

J. Schuemann,[a,1] A. L. McNamara,[a] J. Ramos-Méndez,[b] J. Perl,[c] K. D. Held,[a] H. Paganetti,[a] S. Incerti[d,e] and B. Faddegon[b]

[a] *Department of Radiation Oncology, Massachusetts General Hospital and Harvard Medical School, Boston, Massachusetts,* [b] *Department of Radiation Oncology, University of California San Francisco, San Francisco, California;* [c] *SLAC National Accelerator Laboratory, Menlo Park, California;* [d] *CNRS, IN2P3, CENBG, UMR 5797, F-33170 Gradignan, France; and* [e] *University of Bordeaux, CENBG, UMR 5797, F-33170 Gradignan, France*

Schuemann, J., McNamara, A. L., Ramos-Méndez, J., Perl, J., Held, K. D., Paganetti, H., Incerti, S. and Faddegon, B. TOPAS-nBio: An Extension to the TOPAS Simulation Toolkit for Cellular and Sub-cellular Radiobiology. *Radiat. Res.* 191, 125–138 (2019).

The TOPAS Monte Carlo (MC) system is used in radiation therapy and medical imaging research, having played a significant role in making Monte Carlo simulations widely available for proton therapy related research. While TOPAS provides detailed simulations of patient scale properties, the fundamental unit of the biological response to radiation is a cell. Thus, our goal was to develop TOPAS-nBio, an extension of TOPAS dedicated to advance understanding of radiobiological effects at the (sub-)cellular, (i.e., the cellular and sub-cellular) scale. TOPAS-nBio was designed as a set of open source classes that extends TOPAS to model radiobiological experiments. TOPAS-nBio is based on and extends Geant4-DNA, which extends the Geant4 toolkit, the basis of TOPAS, to include very low-energy interactions of particles down to vibrational energies, explicitly simulates every particle interaction (i.e., without using condensed histories) and propagates radiolysis products. To further facilitate the use of TOPAS-nBio, a graphical user interface was developed. TOPAS-nBio offers full track-structure Monte Carlo simulations, integration of chemical reactions within the first millisecond, an extensive catalogue of specialized cell geometries as well as sub-cellular structures such as DNA and mitochondria, and interfaces to mechanistic models of DNA repair kinetics. We compared TOPAS-nBio simulations to measured and published data of energy deposition patterns and chemical reaction rates (G values). Our simulations agreed well within the experimental uncertainties. Additionally, we expanded the chemical reactions and species provided in Geant4-DNA and developed a new method based on independent reaction times (IRT), including a total of 72 reactions classified into 6 types between neutral and charged species. Chemical stage simulations using IRT were a factor of 145 faster than with step-by-step tracking. Finally, we applied the geometric/chemical modeling to obtain initial yields of double-strand breaks (DSBs) in DNA fibers for proton irradiations of 3 and 50 MeV and compared the effect of including chemical reactions on the number and complexity of DSB induction. Over half of the DSBs were found to include chemical reactions with approximately 5% of DSBs caused only by chemical reactions. In conclusion, the TOPAS-nBio extension to the TOPAS MC application offers access to accurate and detailed multiscale simulations, from a macroscopic description of the radiation field to microscopic description of biological outcome for selected cells. TOPAS-nBio offers detailed physics and chemistry simulations of radiobiological experiments on cells simulating the initially induced damage and links to models of DNA repair kinetics.

INTRODUCTION

A Brief Introduction to TOPAS

Improvements to radiotherapy and imaging can be achieved by understanding how subatomic particles travel through apparatus and tissue. The most precise calculations of such phenomena follow the Monte Carlo (MC) method. The TOol for PArticle Simulation (TOPAS) software project, launched in 2009, has helped to greatly improve the use of Monte Carlo simulations for cancer research and treatment (*1*). Requiring no programming knowledge by its users, TOPAS provides a flexible framework to design simulations for radiation therapy. It enables both clinical applications (e.g., high-precision patient dose calculation) and research (e.g., four-dimensional time-of-flight simulations for detector developments), while its design promotes inter-institutional collaboration (*2–5*).

One of the main reasons for the success of TOPAS is the parameter control system at its core. Parameters specified in one or more text files define the properties of the simulation. For instance, when used for proton dose calculations, they define the extent of the simulated geometry along each cartesian coordinate, the number of voxels in each direction

[1] Address for correspondence: Massachusetts General Hospital and Harvard Medical School, Radiation Oncology, Rm 3206, 125 Nashua St., Boston, MA 02114; email: jschuemann@mgh.harvard.edu.

of the patient computed tomography (CT) scan, the beam angle and the scanning pattern of the beam. The hierarchical nature of these text files allows for sharing and adaptation across radiotherapy centers that use similar equipment (treatment machines, detectors, etc.). The parameter control system syntax is described in detail on the ''Welcome to the TOPAS documentation!'' (https://topas.readthedocs.io/en/latest/) website and by Perl *et al*. (*1*).

TOPAS has been expanded from its initial focus on medical physics to also cover radiation biology, offering scoring options that include linear energy transfer (LET), multiple relative biological effectiveness (RBE) models for proton therapy and organ-based outcome models [tumor control probability (TCP) and normal tissue complication probability (NTCP)] (*6, 7*). TOPAS is now part of the U.S. National Cancer Institute's (NCI) Informatics Technology for Cancer Research (ITCR) Initiative.

An extension system was developed for TOPAS that allows customization for users who need additional features specific to their own application, such as custom scorers or geometries. The extension manager allows users to add new features by writing a short section of simple, very specific C++ classes (a building block of the code) based on templates and C++ helper functions provided within TOPAS. This C++ code can then be integrated into the main TOPAS executable by running a single CMake (https://cmake.org/) command.

Documentation of the latest version of the TOPAS system and the available parameters can be found at ''Welcome to the TOPAS documentation!'' (https://topas.readthedocs.io/en/latest/).

Motivation for TOPAS-nBio

TOPAS has been successfully applied to research in radiation therapy physics and macroscopic organ or cellular biology. However, more fundamental research is needed to understand the underlying mechanisms of radiation action, describe effects of oxygenation, intracellular signaling, drug-induced radiation sensitization or resistance and many other effects (*8–10*). Variations at the (sub-)cellular (cellular and sub-cellular) level, both for tumors and surrounding normal tissues, need to be considered. Such research is ideally supported by detailed *in silico* simulations at the sub-cellular level (*11*).

The goal of this work was to lay the foundation to a deeper understanding of the biological effects of radiation in order to facilitate new research at the boundary between physics, chemistry and biology. By providing detailed physics and chemistry simulations in combination with detailed representations of biological systems, such as cells and their nuclei, we aim to promote a mechanistic description connecting sub-cellular energy deposition phenomena to observable biological outcomes. We have thus developed TOPAS-nBio, an extension to TOPAS specifically aimed at the simulation of radiobiological experiments by modelling detailed biological effects at the nanometer scale. By taking advantage of the simplicity and reliability engineered into TOPAS while providing nanometer-scale Monte Carlo simulations, we have made complex code accessible to researchers who may consider using Monte Carlo simulations to improve the physical, chemical and biological description of their experimental design or data analysis.

METHODS

TOPAS-nBio was, in large part, developed as a library of extensions to the main TOPAS system (www.topasmc.org and https://topas.readthedocs.org). The extension files are to be released, open source, under the Berkeley Software Distribution (BSD) 3-clause or similar. Some of the new features required modification of the TOPAS core to provide additional functionality, made available in TOPAS version 3.2 and above. TOPAS version 3.2 is based on Geant4 version 10.4.p2 (*12–14*).

TOPAS-nBio uses the TOPAS parameter system to control the simulation setup, separating parameters into categories such as geometries (Ge), scoring (Sc), source (So) and time features (Tf). Parameters related to biological damage repair kinetics, specific to TOPAS-nBio, are grouped into the scoring category Sc. An additional category was created to control the chemistry (Ch). A sample parameter file for TOPAS-nBio is shown in Fig. 1.

TOPAS-nBio Extensions

The TOPAS-nBio extensions provide options for sub-cellular geometries, scoring, physics and chemistry. While most users will just need to adjust the values of the parameters used by the new classes, the classes can be modified with minimal coding requirements to adjust each aspect of a user's simulation. Due to the modular nature of the extensions, users can select to only install the features they need for their simulations by downloading the necessary files and adding them to their TOPAS executable.

TOPAS-nBio facilitates and extends the use and configuration of the physical and chemical processes provided by Geant4-DNA (*15–18*). The physical processes of Geant4-DNA, originally intended for radiation transport in liquid water, have recently included cross sections for DNA constituents (*19*), due to the availability of elastic and inelastic cross sections for these materials, which are also available in TOPAS-nBio.

The Geant4 toolkit provides users with basic 3D geometric shapes (solids), which include volumes such as boxes, ellipsoids, cylinders and spheres. All geometries in TOPAS-nBio are made up of either a single Geant4 solid, computer-aided-designed (CAD) solids or a combination of two or more of these solids. In some cases, unions of these solids are used (e.g., to create the double-helix DNA backbone from the union of spheres). The Geant4-DNA user community and others working on track-structure codes also work on new DNA or cell geometries and new DNA repair models. We actively collaborate with some of these groups to include the latest developments in our simulation framework. Several of the features presented here, e.g., interfaces to the DNAFabric code (*20–22*) and DNA repair models (*23–26*), were developed as part of such collaborations.

Physics and Regions

For Monte Carlo track-structure simulations, the transport of charged particles and their interactions are performed in a step-by-step fashion by using the physical processes, their valid energy ranges and associated models provided by Geant4-DNA, described elsewhere (*15, 17, 19, 27–30*). An overview of the performance using different sets of the most recent physics models available in Geant4-DNA is

FIG. 1. Example parameter files for TOPAS-nBio. Parameters are specified by parameter type (pink), category and name (orange), value type (green) and value. A hashtag (#) denotes the start of each comment line. Panel A: Setting up the physics list (Geant4-DNA), a cell with nucleus and DNA and scoring a tuple on the DNA for DSB estimation. Panel B: Setting up a simulation using standard ''step-by-step'' chemical reactions and displaying tracks of selected species. Panel C: Setting up a TOPAS-nBio simulation using the newly implemented independent reaction time (IRT) method.

given by Incerti *et al*. (*15*). Physical processes are available through Geant4 constructors. A constructor is a C++ class which implements a physics list, composed of a specific set of physical processes and associated models, configured according to the energy limits of the models and the particle types associated with the processes. Geant4-DNA currently offers seven physics list constructor options: the default constructor, plus options 1 through 6 that offer different scattering models. Geant4-DNA provides cross sections for electrons, photons, protons and a few selected ions (H through O, Si and Fe). Photon cross sections are based on the Livermore models in Geant4. Currently, Geant4-DNA physics cross sections are available for water and DNA-related materials, i.e., tetrahydrofuran (THF), trimethyl-phosphate (TMP), and pyrimidine (PY) and purine (PU), which serve as precursors for the DNA (and RNA) (*19*). The only other material that has dedicated low-energy cross sections in Geant4-DNA is silicon (*31, 32*) using the ''MicroElec'' package (*33, 34*). Gold cross sections are expected to be available in Geant4-DNA soon (*35, 36*).

TOPAS-nBio gives access to the physics list defined in constructors through parameter settings, while providing flexibility to control the model type involved in each process. Using this approach, the energy cut for applying electron capture or electron solvation is automatically readjusted according to the lower energy limit of the physical models. This makes it possible to combine the elastic models from the CPA100 implementation (*28*) available in constructor G4EmDNAPhysics_opt6 with the inelastic models from the Emfietzoglou-based implementation (*30*) available in G4EmDNAPhysics_opt4.

Regions were originally developed in Geant4 to allow the use of different physics settings (e.g., production cuts of secondary particles) in different parts of the simulated World, i.e., a volume encompassing the entire simulated geometry. This has been exploited in Geant4-DNA to use different physical models in different geometry components (*37*), such as limiting tracking of detailed particle interactions to specific components, in order to speed up the simulation. This feature has been implemented in TOPAS-nBio. An example of parameter settings used to simulate two regions with different physics settings, together with the resulting simulation of a gold nanoparticle (GNP) in water, is shown in Fig. 2.

Simulation Settings and Scenarios

The main difference when setting up simulations using TOPAS-nBio, compared to TOPAS, is the setting of specialized physics lists and chemistry lists to define the basic interaction properties for the simulations. In addition, some settings are optimized by default for all TOPAS-nBio simulations to improve the overall performance. For example, the Geant4 geometric tolerance is adjusted depending on the size of the World or can be adjusted by the user. The tolerance defines an envelope around a surface (similar to an uncertainty band) at which Geant4 considers a point to be on the surface. This is a critical parameter at nanometer scales.

The following are typical, distinct simulation scenarios:

1. Simulations of physics interactions at the nanometer scale using track-structure physics settings. Physics cross sections for simulations using the full track structure are currently limited to water (G4_WATER) and DNA-related materials. The type of the source particles and the energy limits of the physical models can be found on the Geant4-DNA website (http://geant4-dna.org/), updated for the current Geant4 version. To invoke the correct chemistry processes, users should include the settings listed in Fig. 1B.
2. Track-structure simulations that include tracking of chemical species. Simulations of chemical reactions are currently limited to pure water. To invoke the correct chemistry processes, users should include the settings described by Ramos-Méndez *et al*. (*38*).
3. Track-structure simulations that include biological response models. Several published models to predict DNA damage and repair have been included. TOPAS-nBio supports the output of DNA damage in the Standard for DNA Damage (SDD) format [see (*39*) and www.standard-for-dna-damage.readthedocs.org], which can be used as input to several of the available models (*23–26*). In addition, some models have been directly included in TOPAS-nBio to model the biological stage of cell repair.
4. Track-structure simulations containing mixed scales/materials. TOPAS-nBio offers to define regions of interest where Geant4-

A

```
#--- Physics ------------------------------------------------------------------
# Global physics list
sv:Ph/Default/Modules = 1 "g4em-livermore" # Set World default physics list

# Geant4-DNA region:
sv:Ph/Default/ActiveG4DNAInRegionsNamed = 1 "Geant4DNARegion"

# These cuts will enhance the production of secondaries of Livermore
# production threshold in range, default 0.7 mm (by Geant4):
d:Ph/Default/ForRegion/DefaultRegionForTheWorld/CutForElectron = 1 nm

# production threshold in energy, default set by physics list:
d:Ph/Default/SetProductionCutLowerEdge  = 10 eV
d:Ph/Default/SetProductionCutHighEdge   = 500.0 MeV
d:Ph/Default/EMRangeMin = 10 eV     # minimum for EM tables, default 100. eV
d:Ph/Default/EMRangeMax = 500. MeV  # maximum for EM tables, default 500. MeV
i:Ph/Default/EMBins     = 77        # number of bins for EM tables
i:Ph/Default/EMBinsPerDecade = 7    # number of bins per decade for EM tables

b:Ph/Default/Fluorescence          = "True" # Turn on Fluorescence
b:Ph/Default/Auger                 = "True" # Turn on Auger Processes
b:Ph/Default/AugerCascade          = "True" # Turn on Auger Cascade
b:Ph/Default/DeexcitationIgnoreCut = "True" # Simulate full Auger Cascade
b:Ph/Default/PIXE                  = "True" # Turn on PIXE Processes

#--- Geometry -----------------------------------------------------------------
# Define region to use Geant4-DNA
s:Ge/Box/Parent   = "World"
s:Ge/Box/Material = "G4_WATER"
s:Ge/Box/Type     = "TsBox"
d:Ge/Box/HLX      = 0.25 um
d:Ge/Box/HLY      = 0.25 um
d:Ge/Box/HLZ      = 0.25 um
s:Ge/Box/AssignToRegionNamed =  "Geant4DNARegion" # Use Geant4-DNA

# GNP
s:Ge/GNP/Type     = "TsSphere"
s:Ge/GNP/Material = "G4_Au"
s:Ge/GNP/Parent   = "Box"
d:Ge/GNP/RMin     = 0 nm
d:Ge/GNP/RMax     = 50 nm
s:Ge/GNP/Color    = "red"
s:Ge/GNP/AssignToRegionNamed = "DefaultRegionForTheWorld" # Use Livermore
```

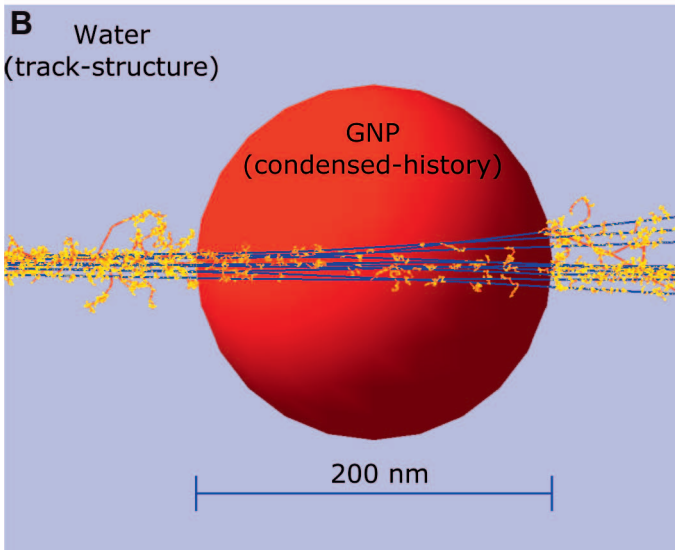

FIG. 2. The use of different models in different regions. Panel A: Parameter file setting physics lists for different regions to combine macroscopic and nanometer scale simulations with several parameters that should be considered. Panel B: Protons (1 MeV) traversing a gold nanoparticle (GNP) surrounded by liquid water. In water, the transport is handled by Monte Carlo track-structure simulations, whereas in the GNP the transport is handled by condensed-history Monte Carlo (Geant4 Livermore EM physics processes). Proton tracks are shown with blue lines and electron tracks are shown with red lines connected with yellow points that represent inelastic and elastic interactions.

DNA processes are activated while using the standard Geant4 electromagnetic physics lists in all other regions (*37*). Figure 2 shows an example parameter file. This is mainly aimed at two scenarios: 1. Simulating larger (>10 μm) volumes but requiring detailed track-structure simulations in smaller sub-volumes, e.g., when simulating irradiation of a flask containing cells in solution and only using track-structure simulations in the nucleus of selected cells; 2. Track-structure simulations in a small volume, but including materials for which there are no cross sections defined yet in Geant4-DNA, e.g., simulating a cell (G4_WATER), which has taken up gold or silver nanoparticles.

Time Dependence

The TOPAS time feature system (*4*) is used to handle time dependencies within the TOPAS-nBio extension. All time features available in the standard TOPAS are available; that is, beam parameters can vary over time, geometries can move over time, scorers can be switched on or off for certain times, and so on. For example, sources can be defined to have different dose rates, influencing the time between potential cellular damage, which can be important information influencing repair kinetics in the biological effect models linked to TOPAS-nBio, e.g., by including the damage time in the SDD format (see sub-section, ''The standard for DNA damage scorer'').

Code Maintenance

TOPAS-nBio was developed by a core team. In addition, alpha users (e.g., University of Manchester) contribute code to the project. Code reliability is fundamental in the development of TOPAS. TOPAS-nBio uses the same system of rigorous development procedures as TOPAS. Each time code is committed to the GitHub repository, a series of instant tests are initiated, including one that checks for a successful build, several ''sanity checks'' that guarantee the committed code does not fundamentally break the code and a simple regression test, ensuring consistency at several key aspects. In addition, before a new version release, each committed code is checked to conform with internal coding guidelines and a more extensive regression test is initiated. This also occurs every time TOPAS is linked to a new Geant4 release. TOPAS-nBio itself is an extension to TOPAS, thus most TOPAS modules are independent of TOPAS-nBio. In addition, similar to TOPAS releases, TOPAS-nBio releases will check each module for coding consistency and initiate a set of sanity checks and regression tests before code is added to the repository.

RESULTS

The main developmental achievements of TOPAS-nBio are access to and control of physics settings, addressed in the Methods section, and the implementation of (sub-)cellular geometries, speed-up options, chemistry and a graphical user interface (GUI). Results of comparison to published studies to test and validate our simulation settings are also reported. In addition, the power of TOPAS-nBio for performing complex simulations is demonstrated by simulating the difference between induced damage on a chromatid with and without tracking of chemical species.

Geometries

TOPAS-nBio offers the user a catalogue of geometries ranging from the micrometer scale (e.g., cells) down to the nanometer scale (e.g., DNA strands). Each geometry has its own set of parameters that can be controlled by the user in the parameter file. For example, the cell geometry has parameters for the size of the cell and also the option to include organelles such as a nucleus. Advanced users also have the option of editing or writing their own geometry extension classes to define new unique geometries. A

```
# Fibre
s:Ge/Fibre/Parent                = "World"
s:Ge/Fibre/Type                  = "TsFibre"
sc:Ge/Fibre/Material             = "G4_WATER"
dc:Ge/Fibre/FibreHalfLength      = 14 nm

# Histone
sc:Ge/Fibre/Histone/Material     = "G4_WATER"
sc:Ge/Fibre/Histone/DrawingStyle = "Solid"
sc:Ge/Fibre/Histone/Color        = "Blue"
# Base 1
sc:Ge/Fibre/base1/Material       = "G4_WATER"
s:Ge/Fibre/base1/DrawingStyle    = "Solid"
s:Ge/Fibre/base1/Color           = "Grass"
# Base 2
s:Ge/Fibre/base2/Material        = "G4_WATER"
s:Ge/Fibre/base2/DrawingStyle    = "Solid"
s:Ge/Fibre/base2/Color           = "Red"

# Backbone 1
sc:Ge/Fibre/back1/Material       = "G4_WATER"
sc:Ge/Fibre/back1/DrawingStyle   = "Solid"
sc:Ge/Fibre/back1/Color          = "Magenta"
# Backbone 2
s:Ge/Fibre/back2/Material        = "G4_WATER"
s:Ge/Fibre/back2/DrawingStyle    = "Solid"
s:Ge/Fibre/back2/Color           = "Yellow"
```

FIG. 3. Parameter file to set up a chromatin fiber defined by a class called TsFiber.

detailed description of the TOPAS-nBio geometries can be found in (*40*).

Possible targets in radiobiological studies may include the whole cell and nucleus or a specific molecule. A collection of geometries is included to cover a range of anticipated simulations, including several cell morphologies, organelles, nuclei models, various DNA models and proteins. An example parameter file setting up a chromatin fiber using the TsFiber geometry, described in detail by Henthorn *et al.* (*26*), is shown in Fig. 3 and a representation of the geometry is shown in Fig. 4C.

Available cell geometries include simple cells such as spherical or ellipsoid cells, as well as highly specialized cells such as neurons (Fig. 4A). Neuron or glial cell geometries are modeled via the TOPAS-nBio interface to the NeuroMorpho database, www.neuromorpho.org (*41, 42*). Since radiation-induced damage to DNA is an important measurement in radiobiology studies, several published DNA models are available, including three full nuclear DNA models incorporating the full hierarchal folding scheme of the DNA in the nucleus (Fig. 4B).

Variance Reduction with Flagged Uniform Particle Splitting

A method to reduce simulation time, in addition to optimization of regions as discussed in the sub-section, ''*Physics and regions*,'' is variance reduction. Population control methods such as particle splitting and cross-section enhancement are examples of variance reduction techniques that have been successfully applied to condensed-history simulations with an impressive reduction in computation time [e.g., (*43, 44*) and papers therein]. In such cases, the quantity of interest is averaged over multiple histories where the track structure of each history is not considered. For

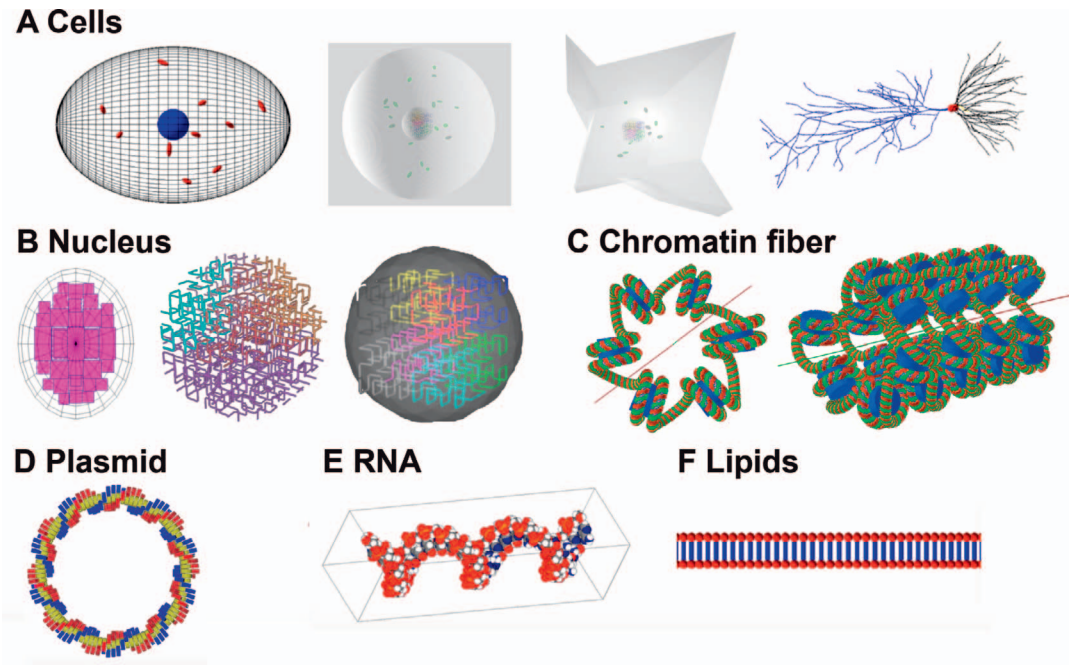

FIG. 4. Geometries available in TOPAS-nBio. Panel A: An ellipsoid cell shown with a nucleus (blue) and mitochondria (red), a spherical and a fibroblast cell with nucleus and mitochondria. Also shown is a hippocampal neuron with a soma (red) and dendrites (black and blue). Panel B: Three full nucleus models, one based on the Geant4-DNA example (left side) and two different fractal models (center and right side). Panel C: A chromatin fiber consisting of nucleosomes each composed of histone proteins (blue) wrapped by two turns of a double helix DNA (green and red). Panel D: A circular plasmid consisting of 100 basepairs. Panel E: RNA strand recreated using the TOPAS-nBio interface to the protein database. Panel F: A lipid (membrane) layer.

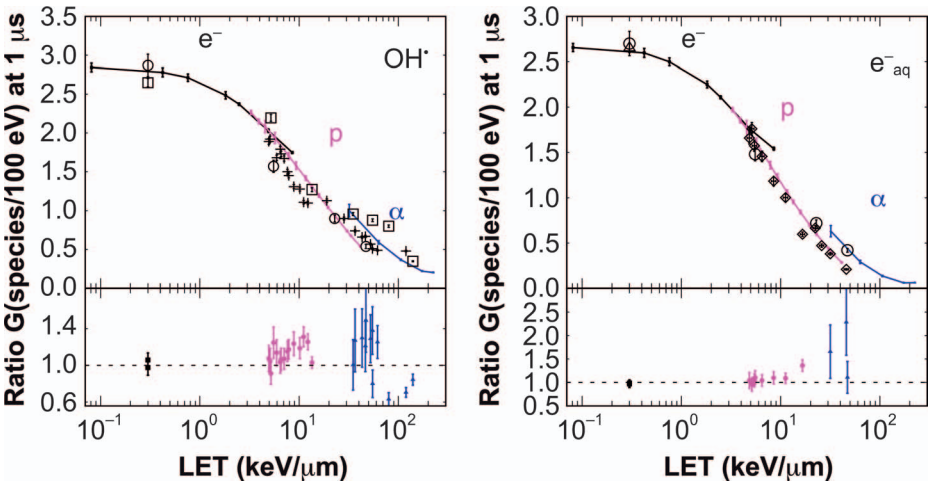

FIG. 5. TOPAS-nBio calculated G values as a function of LET_{100eV} for mono-energetic electrons (e^-), protons (p) and alpha (α) particles. Point-to-point differences with experimental data are shown at the bottom of each panel with error bars (1 standard deviation) including both experimental and simulation errors. Experimental data: (□) (*49*), (○) (*50*), (◇) (*51*), (△) (*52*) and (+) (*53*). (Reproduced with permission. Ramos-Mendez J, Perl J, Schuemann J, McNamara A, Paganetti H, Faddegon B. Monte Carlo simulation of chemistry following radiolysis with TOPAS-nBio. Phys Med Biol. 2018; 63:105014.)

track-structure simulations, however, the stochastic nature of each single history is paramount and the artificial generation of new tracks within the same history potentially biases the results. For example, the use of particle splitting will result in an overestimation of the cluster sizes of ionizations in small volumes and the frequency of DSBs. To exploit the large benefits of variance reduction without bias, we implemented a flagged uniform particle split in TOPAS-nBio, specifically designed for track-structure simulations. This technique performs uniform splitting (*45*) to secondary electrons produced in ionization events at strategically located regions within the geometry and assigns a unique flag number, which is inherited by their progeny (*46*). The flag permits reclassification of each split event as if they were produced by independent histories. This method reduces the variance by improving the statistics of secondary electrons, while keeping the time increase small compared to the generation of additional particles, by only producing them in strategically selected regions. As a result, improvements of the computational efficiency up to a factor of approximately 65 can be achieved without loss of accuracy (*46*).

Chemistry

Step-by-step method. Geant4-DNA is capable of simulating the physical, pre-chemical and chemical stages of the interaction of radiation with biological tissue in a single application (*47, 48*). Configuration of all the simulation parameters can be cumbersome due to the large number of variables involved: dissociation channels for water and their probabilities, reaction rates and diffusion coefficients for each and every reaction, along with configuration of the simulation setup for validation against measured data. TOPAS-nBio provides functionality that facilitates the configuration of simulations that include chemistry (*38*). TOPAS-nBio further allows expansion of the Geant4-DNA default reaction database and type of chemical species. In addition, TOPAS-nBio provides a set of specialized scorers for the spatiotemporal information of chemical species and their yield through a G-value scorer. The chemical parameters were updated from those used by Geant4-DNA and verified through a comparison with experimental data, obtaining satisfactory agreement, as shown in Fig. 5 and (*38*).

Independent reaction time. The pre-chemical stage from Geant4-DNA has been linked to the independent reaction time (IRT) approximation, which has been made available for access in TOPAS-nBio extensions via the TOPAS scoring manager. IRT (*54, 55*) is a computationally efficient alternative to the Brownian dynamics simulation for the calculation of chemical yields resulting from water radiolysis. IRT is a stochastic technique consisting of the sampling of reaction times and chemical reactions of pairs of species, independent from the surrounding neighbors. The method has been shown to be equivalent to full step-by-step Brownian dynamics (*56*). For a given pair of reactive species initially separated by a known distance, the aim is to determine at what time the pair reacts together. For that, the solution to the diffusion equation given by the Green function is inverted to retrieve the reaction time. The form of the Green function depends on the type of reaction (totally or partially diffusion controlled) and the charge between species. The position of reactive products was estimated using the position approach described by Clifford *et al*. (*54*) due to its ease in coding. The iterative sampling process is performed until a time cut specified by the user (e.g., the time to achieve the steady state of the system) is

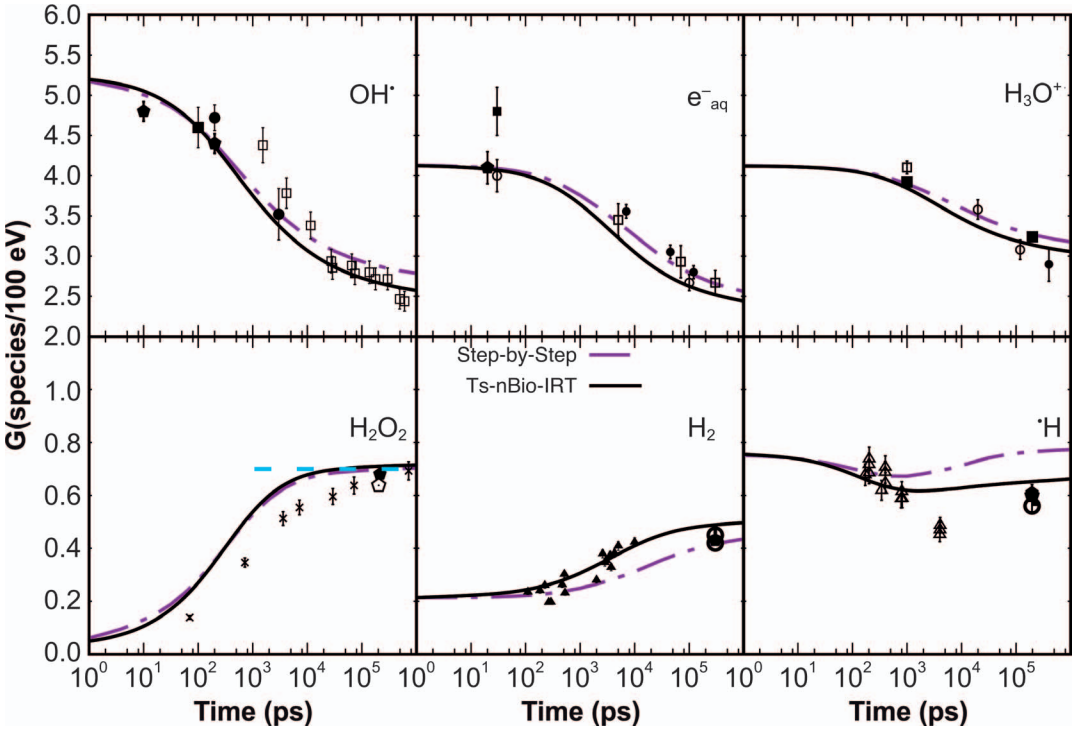

FIG. 6. G values as a function of time for Geant4-DNA/TOPAS-nBio step-by-step simulation (dash-dot line) and IRT (solid line). Experimental data from different sources are represented by symbols. For $^{\bullet}$OH: (□) = ^{60}Co γ rays (*58*), (■) = ~2 MeV electrons (*59*), (•) = 20–22 MeV electrons (*60*), the latter with data scaled by a factor of 0.8 [see (*61*)] and plus sign (+) = 7 MeV electrons (*62*). For e^-_{aq}: (□) = ~35 MeV electrons (*63*), (■) = ~45 MeV electrons (*64*), (○) = ~40 MeV electrons (*65, 66*), (•) = ~2.9 MeV electrons (*67*) and plus sign (+) = 20 MeV electrons (*68*). For H_3O^+: (□) = 5 MeV electrons (*69*), (■) = ^{60}Co and 8 MeV electrons (*70*), (○) = 3.5 MeV electrons (*71*) and (•) = 15 MeV electrons (*72*). For H_2O_2, "×" indicates ^{60}Co γ rays (*58*). For H_2: (△) = ^{60}Co γ-rays (*73*). For H$^{\bullet}$: (△) = ^{60}Co γ rays (*74*). [Reprinted (and expanded) with permission. Ramos-Mendez J, Perl J, Schuemann J, McNamara A, Paganetti H, Faddegon B. Monte Carlo simulation of chemistry following radiolysis with TOPAS-nBio. Phys Med Biol. 2018; 63:105014.]

reached or all possible combinations between species are sampled.

Independent reaction time is computationally more efficient than propagating reactants using Brownian dynamics. However, as the number of species in the track increases, the efficiency decreases due to the iterative process. To mitigate this, Green *et al*. (*55*) proposed the sampling of reaction times only between chemical species separated by a distance of no more than d_{max}, the distance traveled by the most diffusive species before reacting (H_3O^+ in this work). This is done in TOPAS-nBio. In addition, the particle track is binned spatially into a sparse three-dimensional matrix to further reduce computation time. For a given chemical species located at position $\vec{r}$, only those species contained in rectangular parallelepiped voxels fully or partially contained in a virtual sphere of radius d_{max} and centered at $\vec{r}$ are used for the sampling. The voxel size is set to $d_{max}/2$ but it can be defined by the user.

The number of species and reactions currently implemented in the IRT of TOPAS-nBio are those used by Frongillo *et al*. (*57*) and recently by Plante and Devroye (*56*). This implementation includes a total of 72 reactions classified into six types covering first-order reactions, partially and totally diffusion controlled, between neutral and charged species.

Figure 6 shows G values as a function of time for 300 MeV (~0.3 keV/μm) proton track segments of 100 μm in length calculated with both step-by-step Brownian dynamics using TOPAS-nBio with Geant4-DNA chemistry and with IRT from TOPAS-nBio. Statistical uncertainties in both calculations were below 0.5% (1 standard deviation). A revised physics list from Ramos-Méndez *et al*. (*38*) was used, where the relevant Geant4-DNA physics models included the elastic scattering from the CPA100 implementation and the Born model for ionization and excitation. Chemistry parameters with the electron thermalization distance were also from (*38*). As shown, the IRT and step-by-step results are in comparable agreement with the experiment. Both simulation methods used the same input data at the 1-ps starting point (within statistical uncertainties). Note that the different chemical species are tested for contact reactions at zero time by Geant4-DNA before IRT is performed. There is an indication that IRT has improved accuracy of $^{\bullet}$OH yields at 1 μs. This is attributed to the more complete scheme of reactions and the use of partially diffusion-controlled reactions in the implemented IRT. The

```
s:Sc/DBSCAN/Quantity     = "dbscan"
s:Sc/DBSCAN/Component    = "Nucleus"
s:Sc/DBSCAN/OutputFile   = "DBSCAN"
s:Sc/DBSCAN/OutputType   = "Root"
# DBSCAN configuration parameters
# Distance betwen SSB to form a DSB ~10 bp
d:Sc/DBSCAN/MinimumDistanceForDSB        = 3.2 nm
# Hits are scored with 16% probability
u:Sc/DBSCAN/SampleHitsWithProbability    = 0.16
# Hits are sampled from a ramp distribution between
# these two values: 5.0 eV -> 0.0; 37.5 eV -> 1.0
d:Sc/DBSCAN/LowerEnergyForSamplingSSB    = 5.0 eV
d:Sc/DBSCAN/UpperEnergyForSamplingSSB    = 37.5 eV
# Minimum number of SSB to form a DSB
i:Sc/DBSCAN/MinimumNumberofSSBtoFormDSB = 2
```

```
#Scoring using the SDD scorer
s:Sc/dnaDamage/Quantity    = "SDD"
s:Sc/dnaDamage/Component   = "DNA"

#Select whether output should be
# damage patterns or sorted to DSB & SSB
s:Sc/dnaDamage/Format              = "Full" #or partial
d:Sc/dnaDamage/ThresholdEnergy     = 175 eV
i:Sc/dnaDamage/DamageOutput        = 0 #0 for BPs and 1 for nm
i:Sc/dnaDamage/BasepairSeperation = 10
#if DamageOutput is set to nm use:
#d:Sc/dnaDamage/BasepairSeperation = 10 nm
b:Sc/dnaDamage/ScoreBaseLesion     = "true"

s:Sc/dnaDamage/IfOutputFileAlreadyExists = "Overwrite"
s:Sc/dnaDamage/OutputType  = "ASCII" # "binary"
s:Sc/dnaDamage/OutputFile  = "Damage_SDD"
```

FIG. 7. Sample scoring parameter files. Left side: Scorer for the DBSCAN algorithm to determine DSBs without detailed geometry simulation. Right side: Parameter file to create a SDD scorer.

execution time was dramatically reduced with IRT, by a factor of ~145, taking less than 6 s per proton track on a 2.7 GHz Intel® Xeon® processor.

Scorers

TOPAS provides flexibility in scoring quantities for further analysis. We extended the geometry-coupled scoring of TOPAS to cell and sub-cellular geometries. Note that the physical quantities one might want to score will depend on the scale and should be selected accordingly [e.g., macrodosimetric LET vs. microdosimetric lineal energy (y) distributions].

Track-based scoring. While it is generally sufficient for macroscopic Monte Carlo simulations to provide volume and surface scorers, simulations at the sub-cellular scale often require scoring parameters based on each single particle track. For this purpose, the concept of n-tuples, (i.e., an ordered set of values, each containing n elements), has been incorporated. The contents of each different tuple were defined in a new extension class. We have provided a template for users to define the scored quantities, as well as a few options of n-tuple scoring. The n-tuples can be written out in files using ASCII, binary or root data formats. Two special cases of track-based scorers are described below.

SSB and DSB scorers. Simulations of radiation damage often focus on damage to the nucleus, in particular to the DNA double helix. Breaks in the DNA fall into two main categories: SSBs, i.e., breaks that involve only one strand of the DNA, and DSBs, i.e., breaks on both strands of the double helix within a certain defined distance, typically set to be within 10 basepairs. The complexity of the breaks, i.e., how many backbone sugars and bases are damaged within one damage site, is assumed to greatly influence the probability of repair. Accordingly, most studies define additional sub-categories such as a DSB+, a DSB with additional damage within a predefined surrounding base-pair.

Defining an SSB and DSB scorer for generic representations of DNA is nearly impossible due to the inherent subtleties in the design of each DNA geometry. Such scorers need to know how basepairs are defined and how the geometries are set up within Geant4, e.g., as replicas, parameterizations or independent geometries. In TOPAS-nBio, these scorers are thus linked to a specific geometry. Currently, TOPAS-nBio offers two scorers that can generate geometry-dependent SSB and DSB distributions in a tuple format, TsScoreDSBFibre for the TsFibre geometry, and TsScoreDSBFractal for the TsFractal geometry. For both scorers, the user has to select an energy threshold to generate a break (e.g., 17.5 eV) and the number of basepairs between breaks that still count as a single break (e.g., 10 basepairs). These scorers can be used as templates to design scorers for other DNA geometries by updating the geometry dependence of the scorer for each specific implementation.

Another method to score DSBs without considering the actual geometry is one based on the density-based spatial clustering of applications with noise (DBSCAN) algorithm (*75*) that has been adopted by Geant4-DNA (*76*). DBSCAN is used to estimate the number of DSBs based on the distribution of energy deposition events in a given volume (e.g., nucleus) and a likelihood of clusters to form a DSB based on geometric considerations (e.g., the fraction of the volume covered by DNA). We have adopted the DBSCAN algorithm in TOPAS-nBio. A sample parameter file for this scorer is shown in Fig. 7.

The standard for DNA damage scorer. TOPAS-nBio fully supports a new standard data format for DNA damage (SDD) scoring, which has recently been developed (*39*) both to score damage induction and as an input for repair modeling. The SDD scorer is a more generic version of a DNA damage scorer, such as the DSB/SSB scorers, and captures more information about radiation-induced changes to the DNA. However, similar to the SSB/DSB scorers, such a scorer requires detailed information about the simulated geometry. Accordingly, each TOPAS-nBio SDD-scorer has been linked to a specific DNA geometry class. We provide SDD scorers for the same two DNA geometries as for the DSB scorer (TsFibre and TsFractal). These scorers can again be used as templates to design scorer extensions for other DNA geometries. To record damage in an SDD format, users must select one of the two DNA structures, select the SDD fields to be filled and add the information for the DNA repair kinetic modeling that is required by the SDD format or that can optionally be filled.

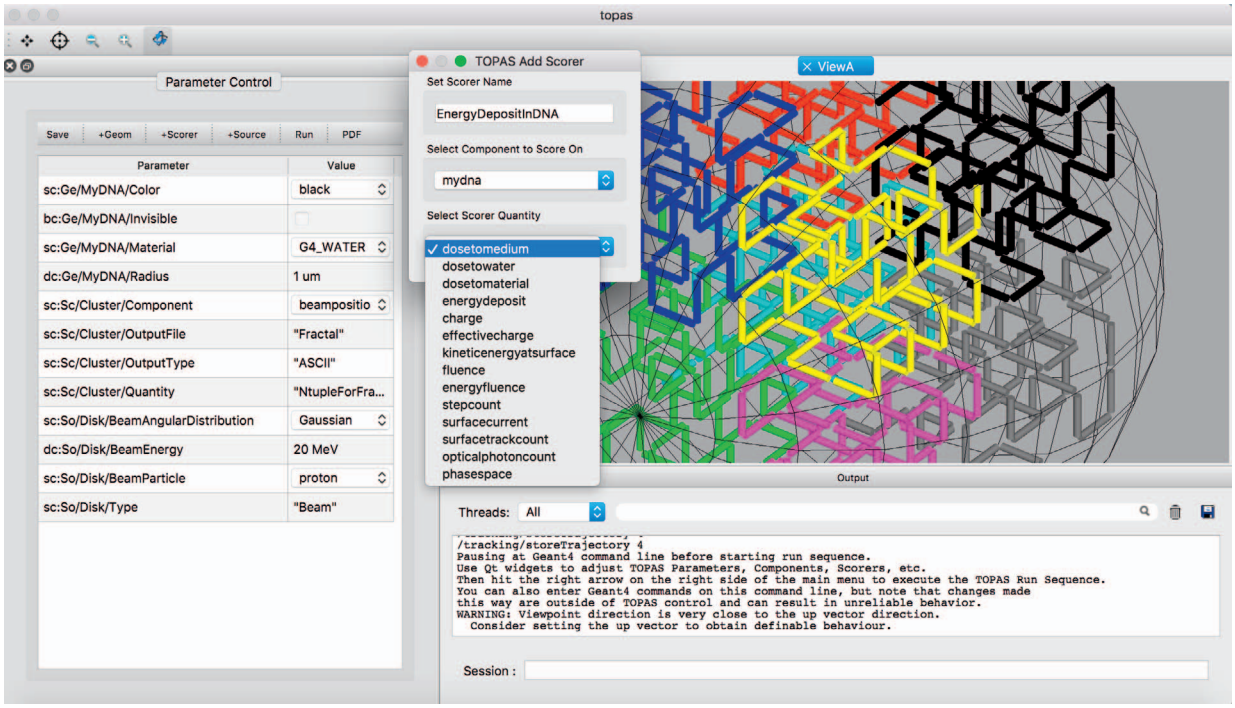

FIG. 8. GUI developed for TOPAS-nBio. Changeable parameters can be set in the GUI, new geometrical components, sources, and scorers (shown) can be added via a drop box interface. The GUI can show the visualization of the simulated geometries and particle tracks once the ''Run'' button is pressed. Images can be exported to PDF.

An example parameter section of an SDD scorer is shown in Fig. 7 for a nucleus containing the fractal walk DNA (Fig. 4B).

Graphical User Interface

The TOPAS text-based parameter system facilitates the setup of Monte Carlo simulations. To further ease the burden of setting up, running and reviewing results of TOPAS-nBio simulations such as virtual radiobiology experiments, we developed a GUI. The goal was to give users access to all the parameters available through the text-based parameter system while at the same time offering a visual interface to TOPAS-nBio (see Fig. 8). The GUI was implemented as part of TOPAS. Thus, while the GUI was designed for TOPAS-nBio users, it will eventually be included in standard TOPAS releases. While users will see the GUI as a whole new front end to TOPAS, the internal design is such that the GUI can be used to adjust values of TOPAS and TOPAS-nBio parameters and then run their simulation from those parameters, so that the already well tested TOPAS parameter control system is still in control. Users can also save the parameter setup with the GUI as a new set of parameter files to later run with or without the GUI. This allows users to run their revised simulations on non-GUI machines such as batch or cloud systems.

The GUI allows users to set up complete simulations. A selection of predefined physics settings is adjusted to be optimal for the length scales of the simulation setup. The user may add geometric objects, such as cells, their nuclei, and their DNA, and scorers to the geometric components, such as a double-strand break (DSB) scorer or scoring DNA damage in the SDD format. In particular, all of the developments described here are available to view and change in the GUI. To allow such multiscale simulation design, the GUI provides a convenient method to zoom between the scales.

Comparisons and Validations of Track-Structure Simulations

Here we summarize our previously published studies validating the simulated track structure, comparing initial DNA damage induction to other codes and experimental data.

Stochastic proton interactions in fluorescent nuclear track detectors. In an effort to validate the simulated track structure, we compared the simulations to experiments performed with fluorescent nuclear track detectors (FNTDs) at the MGH proton beamline (*77*). Multiple measurements were taken at 2–4 mm intervals along pristine and spread-out Bragg peaks; however, due to limited microscope time, only a few data points were evaluated. We concluded that the intensity of the fluorescent signal observed with confocal microscopy should correspond to the number of secondary electrons stopping within a radius corresponding to the optical resolution of the microscope (207 nm). We compared the total intensity of the spot summed from the image data within our region of interest, defined by the microscope resolution and the simulated number of

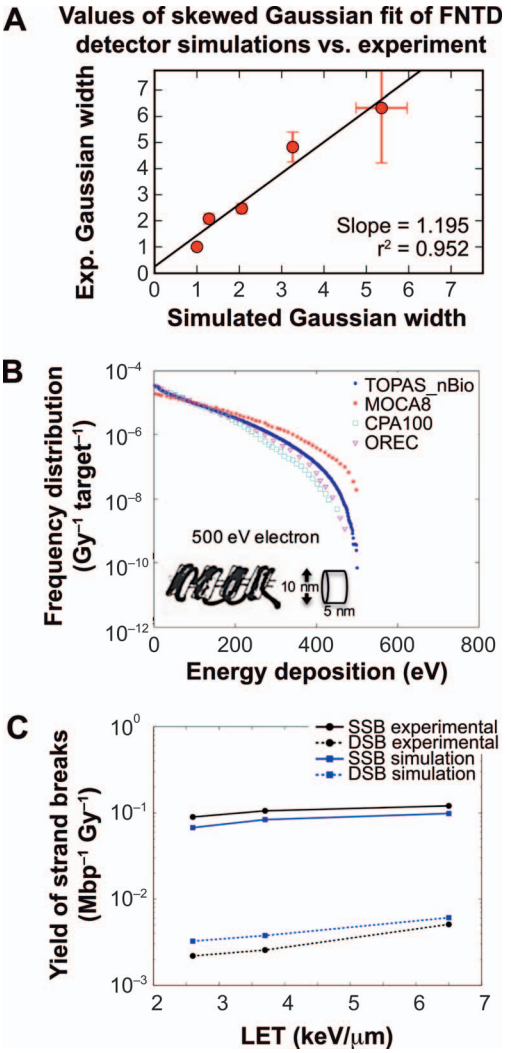

FIG. 9. Validation studies for TOPAS-nBio. Panel A: Comparison of TOPAS-nBio track-structure simulations to measurements using FNTD detectors as described elsewhere (*77*). Shown are correlations of the width of skewed Gaussian fits to the simulated number of electrons stopping within a 207-nm radius of primary proton tracks and experimentally derived integrated brightness of FNTD tracks within a radius of 207 nm. Panel B: Comparison of the frequency of energy depositions in nucleosome-sized cylinders between different codes, TOPAS-nBio, MOCA8, OREC and CPA100 used by Charlton *et al*. (*78, 79*). Panel C: Comparison of TOPAS-nBio SSB and DSB yields in dry plasmids and experimental data (*79, 80*). (Reprinted with permission. McNamara AL, Ramos-Méndez J, Perl J, Held K, Dominguez N, Moreno E, *et al*. Geometrical structures for radiation biology research as implemented in the TOPAS-nBio toolkit. Phys Med Biol. 2018; 63:175018.)

electrons stopping within the same radius of primary proton tracks. The resulting intensity plots at different depths were fitted using a skewed Gaussian. The comparison of the fit results of the Gaussian widths is shown in Fig. 9. We further found strong correlations between the FNTD track intensity and both the track-averaged LET and the frequency mean microdosimetric lineal energy y_F. While this study was limited by detector resolution, it provided a first validation of the Monte Carlo simulated track structure at the subcellular scale.

Simulations in a simple DNA geometry. To validate TOPAS-nBio using historical Monte Carlo track-structure data, a DNA segment using a representation developed by Charlton *et al*. (*81*) was modeled and simulated with both monoenergetic alpha particles and electrons to compare to the simulated results in (*78*). The energy deposited within the DNA segments was modeled using the Geant4-DNA physics processes and both the number and type of strand break within the segment were calculated and compared to that reported in Charlton *et al*. Breaks were defined when the energy deposited within the sub-volume (sugar-phosphate cylinder) was greater than 17.5 eV and then categorized into six different types: SSB, 2SSB, SSB+, DSB, DSB+ and DSB++. Differences between the Charlton *et al*. study and TOPAS-nBio included the generation of more low-energy events as well as more DSBs and fewer DSB++s. However, the overall occurrence of breaks showed similar trends to the Charlton *et al*. study [see (*79*) for details]. The observed differences were attributed to the physics models of the very low-energy events between the two track-structure codes, with TOPAS-nBio predicting more low-energy deposition events compared to the MOCA8 code, which was utilized in the Charlton *et al*. study. Better agreement was found between Geant4-DNA, CPA100 and OREC simulations than with MOCA8, since the cross sections in MOCA8 are based on data for gaseous water, while the other codes are all based on calculations for liquid water (see Fig. 9B). This was consistent with that reported in another comparison study of dose point kernel distribution calculations (*82*).

Plasmid DNA. To validate TOPAS-nBio using experimental data we modeled the irradiation of a dry circular plasmid DNA model with 200 basepairs to compare to the measurements in Vysin *et al*. (*80*). The energy deposited within each component of the DNA plasmid (i.e., the sugar-phosphate backbone) was scored in an n-tuple scorer to determine the total number of SSBs and DSBs as a function of proton LET. Both the experimental and simulation data show an increase in both SSB and DSB yields with increasing LET. Both studies also show that protons are more effective at producing SSBs than DSBs in dry samples. The experimental data, however, predicted higher SSB to DSB ratios for all LET values by approximately a factor of 2 (see Fig. 9C). Differences between the two data sets are likely due to an oversimplified DNA model geometry or other experimental factors not accounted for in the simulation (*79, 80*).

Example of Complex Simulation with TOPAS-nBio

As a proof of concept, and to illustrate application of TOPAS-nBio to an artificial, complex simulation, we simulated protons of 3 and 50 MeV impinging on a chromatin section consisting of seven nucleosomes based

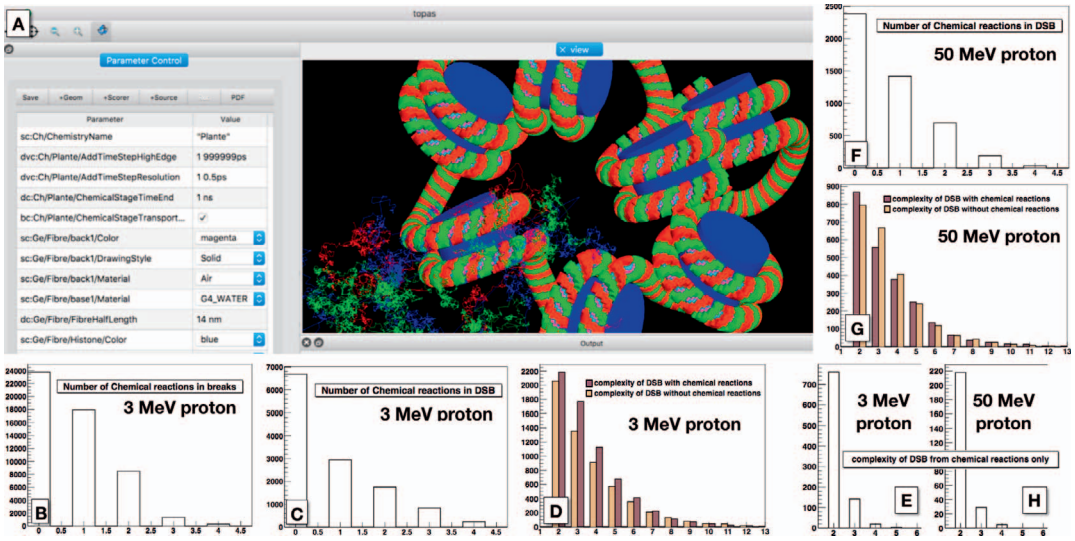

FIG. 10. Simulation of a chromatid fiber irradiated with 3 and 50 MeV protons, including the chemistry of radiolysis. The center shows a representation of the simulated geometry and particle tracks, including the propagation of chemical species. Panels B–E and F–H: The results for 3 and 50 MeV proton irradiations, respectively. The plots demonstrate the effects of including chemical reactions. Here, only interactions with hydroxyl radicals are scored. Approximately one half of the DSBs include a reaction with hydroxyl, which not only influences the overall number but also the complexity of DNA damage (defined as number of lesions within a DSB).

on the TsFibre class, as developed in (*26*), to assess the difference between simulations with and without chemical reactions activated. Damage was scored on the backbones of the DNA geometry (see Fig. 10A). All physics interactions were scored as a damage to the backbone without requiring an energy threshold and a chemical damage was scored when a $^{\bullet}$OH entered a backbone of the DNA geometry, resulting in the immediate removal of the $^{\bullet}$OH from the simulation. Other radicals (e.g., solvated electrons) were tracked but not scored as damage inducing in this simulation unless they produced an $^{\bullet}$OH. The propagation of chemical reactants may be even more important for larger volumes where DNA at more distant sites can be affected. Scavengers limit the distance of chemical propagation in the cell environment. The effect of scavengers was approximated in this demonstration by limiting the transport of chemical reactant propagation to a time of 1 ns. The propagation time can be adjusted through the parameter system to match the anticipated lifetime of chemical species in cell environments with different scavenger concentrations.

Figure 10A shows the results of the simulations on the geometry displayed in that panel. Figure 10B shows the overall number of breaks (SSBs and DSBs) for 3 MeV proton irradiations and the number of breaks that involved $^{\bullet}$OH interactions. Over one half of the DNA damage included at least one $^{\bullet}$OH interaction, with some damage sites involving up to four $^{\bullet}$OH interactions. Damage sites here are defined as a group of DNA lesions (physical energy depositions or $^{\bullet}$OH interaction) with no more than 10 basepairs between neighboring lesions. Accordingly, damage sites can spread over significantly higher numbers of basepairs than 10; in our simulations, the largest DSB observed spread across 41 basepairs. Figure 10C (and F) shows the number of $^{\bullet}$OH interactions involved in damage categorized as DSBs for 3 MeV (50 MeV) protons. It can be seen that the fraction of $^{\bullet}$OH damages in DSBs is slightly less than for all breaks combined. DSBs are defined by having at least one damage pair on opposing strands within 10 basepairs.

Figure 10D and E (G and H) show the complexity of the DSBs, i.e., how many backbones were involved in the damage site for 3 MeV (50 MeV) protons. Multiple physical or chemical damage within the same backbone were counted as single instance. Figure 10D and G show damage including chemical reactions in dark brown and the complexity only from ionization events, i.e., damage sites not involving any chemical reactions in light beige color, and Fig. 10E and H show DSB complexity for damage sites caused entirely by chemical reactions. This latter damage has lower complexity (maximum 5) but would be absent in physics-only simulations. Furthermore, damage involving chemical reactions may be more difficult to repair, since $^{\bullet}$OH attachments can alter the DNA molecules and structure and thus potentially the repair mechanism. Thus, while the damage patterns may be similar, repair probabilities between DSBs induced by physics events alone or a combination of physical and chemical events may differ. Multiple other factors such as chromatin structure (euchromatin vs. heterochromatin) also influence damage induction and repair (*20, 83*). In the presented simulations, only a small section of the TsFibre irradiated with a monoenergetic and monodirectional source was simulated. The propagation of chemical reactants may be even more important for larger

volumes where DNA at more distant sites can be affected. The inclusion of scavengers or, as in this case, the limitation of chemical reactant propagation to a time of 1 ns, limits the distance of chemical propagation to more realistic cell environments. Overall, these simulations show that it is important to study differences in damage patterns and severity when chemistry is considered in order to better understand the processes within a cell.

CONCLUSION

With TOPAS-nBio, we have developed a Monte Carlo framework to model nanometer scale physics and chemical reactions for radiobiology experiments with a graphical user interface that collates progress in the field of track-structure simulations for radiation biology, making these developments accessible to all interested researchers. We summarized our previously published studies validating the simulation framework with improvements to simulation efficiency, validated our implementation of efficient simulation with the independent reaction time approximation, and highlighted the importance of detailed simulations of complex geometries when comparing DSB induction with and without accounting for chemistry.

Thanks to renewed funding by the NCI as part of their program for Informatics Technology for Cancer Research (https://itcr.cancer.gov), TOPAS has recently become free to all nonprofit researchers world-wide with the intention of becoming open source by June 2023. TOPAS-nBio will also be released free and open source in 2019. Licenses will be of the BSD/MIT sort that encourage maximum reuse. We hope that the release of both tools without restrictions for non-profit users will generate further interest from the medical physics and radiation chemistry and radiation biology research communities. We envision that new developments in the field by the Geant4-DNA user community and other researchers working on track-structure codes, new DNA and cell geometries and new DNA repair models, will be included in the TOPAS-nBio code to foster collaborations across the research communities.

The SDD scorer included in the current release is designed as an interface between the physics and chemistry simulations and the biological effect modeling. In addition, biological models can also be directly linked to TOPAS-nBio. The open source format of TOPAS-nBio allows model developers to change any parameter within their simulation framework, or design entirely new geometries or damage induction scorers. This flexibility, in combination with the availability of a large number of precompiled geometries and scorers, provides an ideal framework to design and test new mechanistic models of DNA repair kinetics. TOPAS-nBio further offers new means to test model dependencies on various parameters, for example, by changing geometries or damage definitions such as the damage induction energy threshold.

The developments of TOPAS-nBio presented in this article are early steps to provide researchers with basic tools to further understanding and simulate realistic damage induction from radiation, (sub-)cellular scale responses and eventually tissue and organ level response. In addition, the framework of TOPAS-nBio allows researchers to contribute their own developments as independent extensions to the repository. Such external contributions to the ongoing developments of TOPAS-nBio are highly encouraged and will likely play a significant role to help TOPAS-nBio move the field of mechanistic cell response modeling forward, with the goal to close the gap between simulation of physics events and the biologically observed outcome.

ACKNOWLEDGMENTS

We thank Dudley Goodhead for his advice and serving as our external advisor during the entire TOPAS-nBio project period. We thank John Warmenhoven for developing a TOPAS-nBio version of his repair model, Stephen McMahon for discussions and modeling, Carmen Villagrasa for discussions relating to Geant4-DNA and sharing the DNAFabric model, David Hall who developed the tuple scoring and Siddhanth Bhatt for helping with some figures. This work was supported by the National Institutes of Health (NIH)/National Cancer Institute (NCI), grant no. R01 CA187003.

Received: September 5, 2018; accepted: December 6, 2018; published online: January 4, 2018

REFERENCES

1. Perl J, Shin J, Schuemann J, Faddegon B, Paganetti H. TOPAS: an innovative proton Monte Carlo platform for research and clinical applications. Med Phys 2012; 39:6818–37.
2. Testa M, Schuemann J, Lu HM, Shin J, Faddegon B, Perl J, et al. Experimental validation of the TOPAS Monte Carlo system for passive scattering proton therapy. Med Phys 2013; 40:121719.
3. Schuemann J, Paganetti H, Shin J, Faddegon B, Perl J. Efficient voxel navigation for proton therapy dose calculation in TOPAS and Geant4. Phys Med Biol 2012; 57:3281–93.
4. Shin J, Perl J, Schuemann J, Paganetti H, Faddegon BA. A modular method to handle multiple time-dependent quantities in Monte Carlo simulations. Phys Med Biol 2012; 57:3295–308.
5. Ramos-Méndez J, Perl J, Faddegon B, Schuemann J, Paganetti H. Geometrical splitting technique to improve the computational efficiency in Monte Carlo calculations for proton therapy. Med Phys 2013; 40:041718.
6. Polster L, Schuemann J, Rinaldi I, Burigo L, McNamara AL, Stewart RD, et al. Extension of TOPAS for the simulation of proton radiation effects considering molecular and cellular endpoints. Phys Med Biol 2015; 60:5053–70.
7. Ramos-Mendez J, Perl J, Schuemann J, Shin J, Paganetti H, Faddegon B. A framework for implementation of organ effect models in TOPAS with benchmarks extended to proton therapy. Phys Med Biol 2015; 60:5037–52.
8. Kohandel M, Sivaloganathan S, Oza A. Mathematical modeling of ovarian cancer treatments: Sequencing of surgery and chemotherapy. J Theor Biol 2006; 242:62–8.
9. Appelt AL, Vogelius IR. A method to adjust radiation dose-response relationships for clinical risk factors. Radiother Oncol 2012; 102:352–4.
10. Guerrero M, Carlson DJ. A radiobiological model of reoxygenation and fractionation effects. Med Phys 2017; 44:2002–10.
11. Friedland W, Schmitt E, Kundrat P, Dingfelder M, Baiocco G,

Barbieri S, et al. Comprehensive track-structure based evaluation of DNA damage by light ions from radiotherapy-relevant energies down to stopping. Sci Rep 2017; 7:45161.

12. Agostinelli S, Allison J, Amako K, Apostolakis J, Araujo H, Arce P, et al. GEANT4 – a simulation toolkit. Nucl Instrum Methods Phys Res 2003; A506:250–303.
13. Allison J, Amako K, Apostolakis J, Araujo H, Arce Dubois P, Asai M, et al. Geant4 developments and applications. IEEE Trans Nucl Sci 2006; 53:270–8.
14. Allison J, Amako K, Apostolakis J, Arce P, Asai M, Aso T, et al. Recent developments in Geant 4. Nucl Instrum Methods Phys Res A 2016; 835:186–225.
15. Incerti S, Kyriakou I, Bernal MA, Bordage MC, Francis Z, Guatelli 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:e722–39.
16. Bernal MA, Bordage MC, Brown JMC, Davidkova M, Delage E, Bitar El Z, et al. Track structure modeling in liquid water: A review of the Geant4-DNA very low energy extension of the Geant4 Monte Carlo simulation toolkit. Phys Med 2015; 31:861–74.
17. Incerti S, Ivanchenko A, Karamitros M, Mantero A, Moretto P, Tran HN, et al. Comparison of GEANT4 very low energy cross section models with experimental data in water. Med Phys 2010; 37:4692–708.
18. Incerti S, Baldacchino G, Bernal M, Capra R, Champion C, Francis Z, et al. The Geant4-DNA Project. Int J Model Simul Sci Comput 2010; 01:157–78.
19. Bug MU, Yong Baek W, Rabus H, Villagrasa C, Meylan S, Rosenfeld AB. An electron-impact cross section data set (10 eV–1 keV) of DNA constituents based on consistent experimental data: A requisite for Monte Carlo simulations. Radiat Phys Chem 2017; 130:459–79.
20. Meylan S, Vimont U, Incerti S, Clairand I, Villagrasa C. Geant4-DNA simulations using complex DNA geometries generated by the DnaFabric tool. Comput Phys Commun 2016; 204:159–69.
21. Meylan S, Incerti S, Karamitros M, Tang N, Bueno M, Clairand I, et al. Simulation of early DNA damage after the irradiation of a fibroblast cell nucleus using Geant4-DNA. Sci Rep 2017; 7:11923.
22. Bueno M, Schulte R, Meylan S, Villagrasa C. Influence of the geometrical detail in the description of DNA and the scoring method of ionization clustering on nanodosimetric parameters of track structure: a Monte Carlo study using Geant4-DNA. Phys Med Biol 2015; 60:8583–99.
23. McMahon SJ, Schuemann J, Paganetti H, Prise KM. Mechanistic modelling of DNA repair and cellular survival following radiation-induced DNA damage. Sci Rep 2016; 6:33290.
24. McMahon SJ, McNamara AL, Schuemann J, Paganetti H, Prise KM. A general mechanistic model enables predictions of the biological effectiveness of different qualities of radiation. Sci Rep 2017; 7:688.
25. Henthorn NT, Warmenhoven JW, Sotiropoulos M, Mackay RI, Kirkby NF, Kirkby KJ, et al. In silico non-homologous end joining following ion induced DNA double strand breaks predicts that repair fidelity depends on break density. Sci Rep 2018; 8:2654.
26. Henthorn NT, Warmenhoven JW, Sotiropoulos M, Mackay RI, Kirkby KJ, Merchant MJ. Nanodosimetric simulation of direct ion-induced DNA damage using different chromatin geometry models. Radiat Res 2017; 188:770–83.
27. Francis Z, Incerti S, Karamitros M, Tran HN, Villagrasa C. Stopping power and ranges of electrons, protons and alpha particles in liquid water using the Geant4-DNA package. Nucl Instrum Methods Phys Res 2011; 269:2307–11.
28. Bordage MC, Bordes J, Edel S, Terrissol M, Franceries X, Bardies M, et al. Implementation of new physics models for low energy electrons in liquid water in Geant4-DNA. Phys Med 2016; 32:1833–40.
29. Francis Z, Incerti S, Ivanchenko V, Champion C, Karamitros M, Bernal MA, et al. Monte Carlo simulation of energy-deposit clustering for ions of the same LET in liquid water. Phys Med Biol 2012; 57:209–24.
30. Kyriakou I, Incerti S, Francis Z. Technical note: Improvements in geant4 energy-loss model and the effect on low-energy electron transport in liquid water. Med Phys 2015; 42:3870–6.
31. Raine M, Gaillardin M, Paillet P. Geant4 physics processes for silicon microdosimetry simulation: Improvements and extension of the energy-range validity up to 10 GeV/nucleon. Nucl Instrum Methods Phys Res 2014; 325:97–100.
32. Valentin A, Raine M, Sauvestre JE. Inelastic cross-sections of low-energy electrons in silicon for the simulation of heavy ion tracks with the GEANT4-DNA toolkit. IEEE Nucl Sci Symp Conf Rec 2011; 80–5. (https://bit.ly/2QmHK7f)
33. Valentin A, Raine M, Sauvestre JE, Gaillardin M, Paillet P. Geant4 physics processes for microdosimetry simulation: Very low energy electromagnetic models for electrons in silicon. Nucl Instrum Methods Phys Res 2012; 288:66–73.
34. Valentin A, Raine M, Gaillardin M, Paillet P. Geant4 physics processes for microdosimetry simulation: Very low energy electromagnetic models for protons and heavy ions in silicon. Nucl Instrum Methods Phys Res 2012; 287:124–9.
35. Sakata D, Incerti S, Bordage MC, Lampe N, Okada S, Emfietzoglou D, et al. An implementation of discrete electron transport models for gold in the Geant4 simulation toolkit. J Appl Phys 2016;120:244901.
36. Sakata D, Kyriakou I, Okada S, Tran HN, Lampe N, Guatelli S, et al. Geant4-DNA track-structure simulations for gold nanoparticles: The importance of electron discrete models in nanometer volumes. Med Phys 2018; 45:2230–42.
37. Ivanchenko V, Apostolakis J, Bagulya A, Abdelouahed HB, Black R, Bogdanov A, et al. Recent improvements in Geant4 electromagnetic physics models and interfaces. Proceedings, Joint International Conference on Supercomputing in Nuclear Applications and Monte Carlo: Tokyo, Japan, October 17–21, 2010. Tokai, Japan: Japan Atomic Energy Agency; 2010. p. 898–903. (https://bit.ly/2PqOagb)
38. Ramos-Mendez J, Perl J, Schuemann J, McNamara A, Paganetti H, Faddegon B. Monte Carlo simulation of chemistry following radiolysis with TOPAS-nBio. Phys Med Biol. 2018; 63:105014.
39. Schuemann J, McNamara AL, Warmenhoven JW, Henthorn NT, Kirkby K, Merchant MJ, et al. A new standard DNA damage (SDD) data format. Radiat Res 2019; 191:76–92.
40. McNamara AL, Ramos-Méndez J, Perl J, Held K, Dominguez N, Moreno E, et al. Geometrical structures for radiation biology research as implemented in the TOPAS-nBio toolkit. Phys Med Biol. 2018; 63:175018.
41. Ascoli GA. Mobilizing the base of neuroscience data: the case of neuronal morphologies. Nat Rev Neurosci 2006; 7:318–24.
42. Ascoli GA, Donohue DE, Halavi M. NeuroMorpho.Org: a central resource for neuronal morphologies. J Neurosci 2007; 27:9247–51.
43. Fragoso M, Kawrakow I, Faddegon BA, Solberg TD, Chetty IJ. Fast, accurate photon beam accelerator modeling using BEAMnrc: a systematic investigation of efficiency enhancing methods and cross-section data. Med Phys 2009; 36:5451–66.
44. Kawrakow I, Rogers DWO, Walters BRB. Large efficiency improvements in BEAMnrc using directional bremsstrahlung splitting. Med Phys 2004; 31:2883–98.
45. Rogers DW, Faddegon BA, Ding GX, Ma CM, We J, Mackie TR. BEAM: a Monte Carlo code to simulate radiotherapy treatment units. Med Phys. 1995; 22:503–24.
46. Ramos-Mendez J, Schuemann J, Incerti S, Paganetti H, Schulte R, Faddegon B. Flagged uniform particle splitting for variance reduction in proton and carbon ion track-structure simulations. Phys Med Biol 2017; 62:5908–25.
47. Karamitros M, Luan S, Bernal MA, Allison J, Baldacchino G,

Davidkova M, et al. Diffusion-controlled reactions modeling in Geant4-DNA. J Comput Phys 2014; 274:841–82.

48. Karamitros M, Mantero A, Incerti S, al E. Modeling radiation chemistry in the Geant4 toolkit. Prog Nucl Sci Tec 2011; 2:503–8.

49. Burns WG, Sims HE. Effect of radiation type in water radiolysis. J Chem Soc Faraday Trans 1 1981; 77:2803–13.

50. Appleby A, Schwarz HA. Radical and molecular yields in water irradiated by gamma-rays and heavy ions. J Phys Chem 1969; 73:1937–41.

51. Sauer MC, Schmidt KH, Hart EJ, Naleway CA, Jonah CD. LET dependence of transient yields in the pulse radiolysis of aqueous systems with deuterons and alpha particles. Radiat Res 1977; 70:91–106.

52. Elliot AJ, Chenier MP, Ouellette DC. Temperature dependence of g values for H2O and D2O irradiated with low linear energy transfer radiation. J Chem Soc Faraday Trans 1993; 89:1193–7.

53. Anderson AR, Hart EJ. Molecular product and free radical yields in the decomposition of water by protons, deuterons, and helium ions. Radiat Res 1961; 14:689–704.

54. Clifford P, Green NJB, Oldfield MJ, Pilling MJ, Pimblott SM. Stochastic models of multi-species kinetics in radiation-induced spurs. J Chem Soc Faraday Trans 1 1986; 82:2673–89.

55. Green NJB, Pilling MJ, Pimblott SM, Clifford P. Stochastic modeling of fast kinetics in a radiation track. J Phys Chem 1990; 94:251–8.

56. Plante I, Devroye L. Considerations for the independent reaction times and step-by-step methods for radiation chemistry simulations. Radiat Phys Chem 2017; 139:157–72.

57. Frongillo Y, Goulet T, Fraser M-J, Cobut V, Patau P, Jay-Gerin J-P. Monte Carlo simulation of fast electron and proton tracks in liquid water – II. nonhomogeneous chemistry. Radiat Phys Chem 1998; 51:245–54.

58. LaVerne JA. OH radicals and oxidizing products in the gamma radiolysis of water. Radiat Res 2000; 153:196–200.

59. Jay-Gerin J-P, Ferradini C. A new estimate of the radical yield at early times in the radiolysis of liquid water. Chem Phys Lett 2000; 317:388–91.

60. Jonah CD, Miller JR. Yield and Decay of the Hydroxyl Radical from 200ps to 3ns. J Phys Chem 1977; 81:1974–6.

61. Kreipl MS, Friedland W, Paretzke HG. Time- and space-resolved Monte Carlo study of water radiolysis for photon, electron and ion irradiation. Radiat Environ Biophys 2009; 48:11–20.

62. Omar El AK, Schmidhammer U, Jeunesse P, Larbre J-P, Lin M, Muroya Y, et al. Time-dependent radiolytic yield of $OH^{\bullet}$ radical studied by picosecond pulse radiolysis. J Phys Chem A 2011; 115:12212–6.

63. Shiraishi H, Katsumura Y, Hiroishi D, Ishigure K, Washio M. Pulse-radiolysis study on the yield of hydrated electron at elevated temperatures. J Phys Chem 1988; 92:3011–7.

64. Sumiyoshi T, Katayama M. The yield of hydrated electrons at 30 picoseconds. Chem Lett 1982; 11:1887–90.

65. Wolff RK, Bronskill MJ, Aldrich JE, Hunt JW. Picosecond pulse radiolysis. IV. Yield of the solvated electron at 30 picoseconds. J Phys Chem 1973; 77:1350–5.

66. Hunt JW, Wolff RK, Bronskill MJ, Jonah CD, Hart EJ, Matheson MS. Radiolytic yields of hydrated electrons at 30 to 1000 picoseconds after energy absorption. J Phys Chem 1973; 77:425–6.

67. Buxton GV. Nanosecond pulse radiolysis of aqueous solutions containing proton and hydroxyl radical scavengers. Proc Math Phys Eng Sci 1972; 328:9–21.

68. Muroya Y, Lin M, Wu G, Iijima H, Yoshii K, Ueda T, et al. A re-evaluation of the initial yield of the hydrated electron in the picosecond time range. Radiat Phys Chem 2005; 72:169–72.

69. Pikaev AK, Kabakchi SA, Zansokhova AA. Yields and reactions of hydrogen ions on radiolysis of water and aqueous solutions. Faraday Discuss Chem Soc 1977; 63:112–23.

70. Cercek B, Kongshaug M. Hydrogen ion yields in the radiolysis of neutral aqueous solutions. J Phys Chem 1969; 73:2056–8.

71. Anderson RF, Vojnovic B, Michael BD. The radiation-chemical yields of H3O+ and OH– as determined by nanosecond conductimetric measurements. Radiat Phys Chem (1977) 1985; 26:301–3.

72. Schmidt KH, Ander SM. Formation and recombination of the hydronium ion (H3O+) and hydroxide in irradiated water. J Phys Chem 1969; 73:2846–52.

73. Draganic ZD, Draganic IG. Formation of primary reducing yields (Geaq– and GH2) in the radiolysis of aqueous solutions of some positive ions. Int J Radiat Phys Chem 1975; 7:381–6.

74. Draganic ZD, Draganic IG. Formation of primary hydrogen atom yield (GH) in the gamma radiolysis of water. J Phys Chem 1972; 76:2733–7.

75. Ester M, Kriegel HP, Sander J, Xu X. A density-based algorithm for discovering clusters in large spatial databases with noise. KDD-96 Proceedings. Palo Alto, CA: Association for the Advancement of Artificial Intelligence; 1996. p. 226–31. (https://bit.ly/1rtAMVx)

76. Francis Z, Villagrasa C, Clairand I. Simulation of DNA damage clustering after proton irradiation using an adapted DBSCAN algorithm. Comput Methods Programs Biomed 2011; 101:265–70.

77. Underwood TSA, Sung W, McFadden CH, McMahon SJ, Hall DC, McNamara AL, et al. Comparing stochastic proton interactions simulated using TOPAS-nBio to experimental data from fluorescent nuclear track detectors. Phys Med Biol 2017; 62:3237–49.

78. Charlton DE, Nikjoo H, Humm JL. Calculation of initial yields of single- and double-strand breaks in cell nuclei from electrons, protons and alpha particles. Int J Radiat Biol 1989; 56:1–19.

79. McNamara A, Geng C, Turner R, Mendez JR, Perl J, Held K, et al. Validation of the radiobiology toolkit TOPAS-nBio in simple DNA geometries. Phys Med 2017; 33:207–15.

80. Vysin L, Pachnerova Brabcova K, Stepan V, Moretto-Capelle P, Bugler B, Legube G, et al. Proton-induced direct and indirect damage of plasmid DNA. Radiat Environ Biophys 2015; 54:343–52.

81. Charlton DE, Humm JL. A method of calculating initial DNA strand breakage following the decay of incorporated 125I. Int J Radiat Biol Relat Stud Phys Chem Med 1988; 53:353–65.

82. Champion C, Incerti S, Perrot Y, Delorme R, Bordage MC, Bardies M, et al. Dose point kernels in liquid water: an intra-comparison between GEANT4-DNA and a variety of Monte Carlo codes. Appl Radiat Isot 2014; 83 Pt B:137–41.

83. Goodarzi AA, Jeggo P, Lobrich M. The influence of heterochromatin on DNA double strand break repair: Getting the strong, silent type to relax. DNA Repair 2010; 9:1273–82.