NUMERICAL MODELS FOR RADIATION-INDUCED DNA DAMAGE

Piotr PATER

Doctor of Philosophy

Department of Physics

McGill University

Montreal,Quebec

December 12, 2016

A thesis submitted to McGill University in partial fulfillment of the requirements of the degree of Doctor of Philosophy

DEDICATION

Dla Julusia, kochający ciebie mocno, Tatuś

ACKNOWLEDGEMENTS

First and foremost, I would like to thank my supervisors Issam El Naqa and Jan Seuntjens for guidance, encouragement and support throughout my graduate studies. I would also like to acknowledge the great help I've received from all my co-authors: Dr. Mario A. Bernal, Dr. Gloria Bäckstöm, Dr. Fernanda Villegas, Dr. Anders Ahnesjö, and Dr. Shirin A. Enger. I also have to thank the members of my PhD committee, Dr. Norma Ybarra, B, and C for their helpful suggestions in general. I would like to throw a special shout out to Dr. Sébastien Incerti, who helped me from the beginning getting a better grip on the Monte Carlo code GEANT4. Over the years, I've had the opportunity of meeting with many different persons that had a real impact on my research: Dr. Léon Sanche, Dr. Ianik Plante, Dr. Dudley Goodhead, Dr. Hans Rabus, Dr. Werner Friedland, Dr. Robert Stewart, Dr. Mohammad Rezaee, and Dr. Micheal Dingfelder. I'd like to thank them all for the time they devoted to my questions and ideas. In addition, I also thank my colleagues and friends, Pavlos, Peter, Robert, James, Daniel, Martin, Marc-André, Susannah, Yana, Allison, Jessica, and all other Medical Physics Unit staff and students, who all helped me in different ways during these past five years. I am forever grateful for all the help I've received from both my parents Dorota and Ryszard. I love you both. Finally, from the bottom of my hearth, I would like to thank my loving wife Marie-Pier who supported me continuously through this long and challenging endeavour. *Je t'aime chérie.*

ABSTRACT

This thesis introduces and explores various numerical models for calculating DNA damage produced by photons, electrons, protons, and other light ion beams. Clustered DNA damage in the form of single-strand and double-strand breaks are a hallmark of radiation and a known initiating agent of loss of cellular reproductive ability. A rigorous analysis of the Monte Carlo track structure calculation methodology is presented with emphasis put on standardization of the definitions, values, and algorithms relating the various adjustable parameters in the model. Commonly overlooked problems associated with the strand break threshold energy, the clustering algorithm, the impact of the dose distribution, and the lack of charged particle equilibrium are presented. Estimates of nucleosome strand breakage clustering are obtained to highlight the higher damage complexity of lower kinetic energy electron beams. The algorithms presented offer innovative and fast DNA damage induction searches inside geometrical models of the complete human DNA genome of a biological cell nucleus. Strand break yields induced by proton and light ion beams are described and quantified with respect to the linear energy transfer (LET). It was found that protons create statistically more double-strand breaks than carbon ions of the same LET. In these simulations, computation time places a fundamental limit on the level of detail simulated; however, the data presented shows that the radiation field affecting the nuclear DNA needs to be described with precision in order to reproduce clinical irradiation setups. Firstly, an introduction to the track overlay algorithm built upon a pre-generated

library of electron tracks obtained from a precise description of the photon field experienced around a cell nucleus is presented. Subsequently, an analysis was performed to quantify the creation of double-strand breaks by clinical photon fields, including cone-beam computed-tomography, electronic brachytherapy, and several radio-isotopes. Lastly, the effectiveness at inducing double-strand breaks are computed using the relative biological effectiveness metric. This is complemented by an investigation of the trends of this quantity with respect to particle type and energy.

RÉSUMÉ

Cette thèse explore différents modèles numériques pour le calcul des dommages à l'ADN produits par des faisceaux de photons, d'électrons, de protons et d'autres ions. Les cassures simples et doubles des brins d'ADN sont caractéristiques de l'effet des rayonnements ionisants sur l'appareil cellulaire. Elles sont souvent à l'origine de la perte de fonction des cellules irradiées. Les simulations de ce type de dommage à l'ADN sont possibles avec des techniques de calcul Monte Carlo de la trajectoire exacte des particules passants près de l'ADN. Cette thèse s'intéresse particulièrement à la normalisation des définitions, des paramètres réglables et des algorithmes utilisés lors de ces simulations. Souvent négligée, l'énergie nécessaire pour rompre un brin d'ADN est déterminante dans ce type de simulations. Similairement, l'algorithme de groupage des dommages à l'ADN, la distribution de la dose dans la cellule, et l'obtention d'un équilibre électronique sont discutés. Des estimations du nombre de ruptures de brins dans les nucléosomes sont obtenues afin de mettre en évidence la complexité des dommages à l'ADN créés par les électrons de plus faible énergie cinétique. Les algorithmes proposés permettent une recherche rapide et innovantes des dommages à l'ADN à l'intérieur de modèles géométriques de la totalité du génome humain localisé dans un noyau de cellule. Les ruptures de brins d'ADN dues à l'action de protons et d'autres ions sont caractérisées par rapport au transfert linéique d'énergie (TLE). Il a été constaté que les protons créent statistiquement plus de double bris de brins que les ions de carbone d'égal TLE. Dans ce type de simulations, le

temps de calcul impose une limite fondamentale au niveau de détail accessible. Cependant, les données présentées dans cette thèse montrent que le champ de rayonnement affectant le noyau a une grande influence sur les résultats et doit être décrit avec précision. Cette thèse présente tout d'abord une introduction à l'algorithme de superposition des trajectoires de particules se basant sur une bibliothèque de trajectoires générées à l'avance. Subséquemment, des descriptions précises des champs de photons et d'électrons qui agissent autour du noyau sont présentées. Les cassures double brins sont calculées pour des champs de photons cliniques, comprenant des tomographies à faisceau coniques, des sources de rayons-x miniatures et des radio-isotopes. Enfin, l'efficacité biologique à induire des cassures double brins est obtenue et comparée.

Table of Contents

List of Tables

List of Figures

CHAPTER 1
Introduction

"To explain all nature is too difficult a task for any one man or even for any one age. It is much better to do a little with certainty and leave the rest for others that come after than to explain all things by conjecture without making sure of any thing."

— Sir Isaac Newton

1.1 Thesis outline

This thesis is divided into 7 chapters. This chapter gives a general introduction to radiation therapy for the treatment of cancer, with special attention to the motivations for predicting the biological consequences of irradiation. Chapter 2 reviews selected methods and results of numerical modeling of radiation-induced DNA damage from the literature. Chapter 3 proposes an in-depth discussion of useful concepts and quantities in use in radiobiological modeling. These chapters serve as an introduction for the next three chapters, consisting of two published articles in *Medical Physics* and one manuscript currently under revision in Physics in Medicine and Biology. Chapter 4 presents our initial findings concerning the influence of various user-defined parameters and algorithms in numerical models of biological damage. Chapter 5 improves on the work of Chapter 4 and focuses on the implementation of a numerical model applied to protons and light ions.

Chapter 6 adapts our previous models to photon beams in different clinical situations. Chapter 7 summarizes the general conclusions from this work and discusses potential future research perspectives.

1.2 Contribution of authors

This thesis is fully written by Piotr Pater. The core work of this thesis was published in 2 research articles and 1 review article, in addition to a manuscript that is currently under review. As per university policy, here are the details of the contributions of each author.

The first manuscript (Chapter 4), "On the consistency of Monte Carlo Track Structure DNA Damage Simulations", P Pater, J Seuntjens, I El Naqa, M Bernal, *Medical Physics*, **41** 121708 (2014),[1] presents published work consisting of the characterization of the adjustable parameters necessary in simulations of DNA damage by electron beams. PP constructed the model, conducted the simulations and wrote the first draft. MB helped with the DNA model parameters and DNA damage induction algorithms. MB, JS, and IEN provided general guidance, manuscript correction and approval.

The second manuscript (Chapter 5), "Proton and Light Ion RBE for the Induction of Direct DNA Double Strand Breaks", P Pater, G Bäckstöm, F Villegas, A Ahnesjö, S A Enger, J Seuntjens, and I El Naqa, *Medical Physics*, **45**, 5, pp 2131–2140, (2016),[2] presents published work considering the simulation of DNA damage induction by proton and other light ion beams. PP created the model and the algorithms, conducted the simulations, analyzed and discussed the data, and wrote the first draft. GB provided initial simulation data from another code. FV and

AA assisted in the formulation of the methodology. GB, AA, FV, SAE, JS, and IEN were responsible for general guidance, correction and approval of the manuscript.

The third manuscript (Chapter 6), "Event-by-Event Electron Spectra in Cells for RBE Calculations", P Pater, G Famulari, I El Naqa, and J Seuntjens, was submitted on May 27th, 2016 and is currently under review in *Physics in Medicine and Biology*. It presents a more exact simulation methodology for DNA damage induction by clinical photon beams, including event-by-event electron track simulation in cm-sized volumes. PP constructed the modeling technique, conducted the simulations, analyzed and discussed the data, and wrote the first draft. GF provided track data for some brachytherapy sources and corrected the manuscript. IEN and JS provided general guidance, corrections and approval of the manuscript.

In addition, several key observations of this thesis were published in a review article, "Monte Carlo Role in Radiobiological Modeling", I El Naqa, P Pater, and J Seuntjens, *Physics in Medicine and Biology*, **57**, 11, R75–R97, (2012).[3] Piotr Pater participated in the description of the numerical codes used in the field and produced illustrations demonstrating their capabilities. However, this review article is not reproduced in this thesis.

1.3 Background

Cancer is a systemic disease that can affect any organ or tissue in the human body. According to the 2015 statistics of the Canadian Cancer Society, roughly 2 in 5 Canadians will develop cancer in their lifetimes and 1 in 4 will die of the disease.[4] Treatment options are dictated by several factors, such as age, stage, cell type, location, etc. Typical treatments include a combination of surgery, radiation therapy

(RT) and/or chemotherapy. In particular, RT is a modality administered to more than half of cancer patients.[5] Several RT modalities were developed over the years, aiming at controlling tumor growth while limiting side-effects on healthy tissue. Treatments with external x-rays include the use of orthovoltage tubes, cobalt-60 (^{60}Co) irradiators and high energy linear accelerators. Treatments with proton and ion beams necessitate large facilities for the acceleration of these particles to high energies. In the case of brachytherapy, radioactive seeds are placed near the tumor, either surgically or through various body orifices. In targeted radionuclide therapy (TRT), the radiation is delivered with the help of a pharmaceutical with special propensity to accumulate in tumor cells. In addition, x-ray images produced by computed tomography (CT), mammography, and positron emission tomography (PET) are routinely used in the diagnosis of cancer or for alignment of the tumor with radiation. In fact, ionizing radiation (IR) has been used in medical applications since shortly after the discovery of x-rays by Roentgen in 1895.

In 1900, Kienbock demonstrated that the dose of x-rays, (i.e., the quantity of absorbed radiation[I]) is the biologically effective agent.[6,7] Dose deposition in biological tissue starts a series of physical, chemical, and biological effects that can ultimately lead to the death of the irradiated cells or to mutations. In fact, IR is at the same time a cure and a cause of cancer. IR is known to break the DNA molecule responsible for the genetic coding of every cell. If the damage is too complex to be repaired, the cell ceases to carry out its functions (i.e., it dies). If

[I] a more thorough definition is given in Section 3.1.

enough tumor cells are killed, the tumor shrinks and can be eradicated. Typically, healthy tissues have a higher resistance to radiation, because of more developed repair mechanisms; however, excessive dose will lead to unwanted side-effects. It is the responsibility of a clinical medical physicist to ensure that the patient receives the prescribed treatment with minimal damage to healthy tissue. This is achieved by careful choice of the RT modality, collimation, gantry angles, precise calculation of dosage, and quality assurance of the medical equipment used for the treatment (for further details see Ref. 8). Additional improvements to the separation between tumoricidal doses and acceptable tolerance of normal tissues may be achieved by adaptation of the fractionation schedule, concomitant use of radio-chemistry, utilization of radio-sensitive pharmaceuticals accumulating in the tumor, or by gene radiotherapy.[9] The rest of this section will cover introductory notions to this research field, necessary for the the formulation of the hypothesis and aims in Section 1.4.

1.3.1 Radiation therapy (RT)

IR is any type of radiation with energy high enough to ionize atoms or molecules such as DNA, in which case, a series of biological effects may occur. Formally, electromagnetic radiation with energy higher than $\approx$ 10–30 eV, sub-atomic particles such as electrons, positrons, alpha particles, protons, or neutrons, and high speed ions and atoms are all examples of IR used in RT.

The current work-horse of RT is the clinical linear accelerator (LINAC) used in the majority of treatments worldwide. LINACs collide a microwave-accelerated pencil beam of electrons with a heavy metal target to produce high-energy (6–20

MV)[II] x-rays through the bremsstrahlung effect. The x-ray beam is shaped and directed to the patients tumor(s) while they are laying on a treatment bed. The beam exits a gantry that shapes the beam using a series of collimators and that can be rotated around the patient to deliver radiation from any angle. Medical LINACs were invented in the 1950–60 in replacement of ^{60}Co irradiators, that are limited in energy. ^{60}Co is a radioactive isotope of cobalt produced in nuclear reactors and it decays by emitting two gamma rays of 1.17 and 1.33 MeV.[III] The isotope is placed in a lead-shielded vessel with an adjustable collimator, and can be mounted on a gantry for treatment from all angles. This technology was in fact invented by a Canadian medical physicist, Harold E. Johns. Prior to the invention of ^{60}Co irradiators, RT was conducted using x-ray tubes or orthovoltage machines, limited to x-rays of $\approx 500\,\mathrm{kV}$.

The energy of the photon beam dictates its penetration depth in matter. LINACs are more adapted for treating deep-seated tumors, whereas orthovoltage machines are still in use for superficial and subdermal treatments. Photons deposit their energy through the photoelectric effect, Compton effect, or pair-production (for photon energies higher than 1.22 MeV), in each case generating energetic secondary electrons and positrons (in the case of pair-production). These energetic

[II] Since the resulting beam is poly-energetic, it is customary to refer to the photon beam energy by using the acceleration potential in volts instead of the energy in electronvolts.

[III] To be precise, ^{60}Co decays by beta decay into ^{60}Ni, which then decays by gamma emission.

secondary charged particles ionize and excite the atoms and molecules on their paths by depositing minute amounts of energy locally and producing their own secondary electrons, in a snowball effect, until all the remaining energy has been dispersed. In this sense, photons are indirectly ionizing, as most of the energy is deposited by the secondary electrons they put in motion. The distinction between directly and indirectly ionizing particles was recommended by an international Commission on radiation Units and measurements (ICRU) report;[10] however, the more adequate terminology, *indirectly transferring*, should be instead adopted, as photons may create ionizations through photoelectric and Compton effects, yet most of their energy is transferred by the secondary electrons created.

The photon modalities described previously are categorized as external beam RT (EBRT). In some clinical situations, it is beneficial for the patient to be treated with radiation sources placed in proximity to the treated region. This type of treatment is called brachytherapy. Radioisotopes the size of a grain of rice with photon emission energy, intensity, and half-lives carefully considered are permanently implanted surgically, inserted in body cavities or in surgically-implanted catheters, or placed on the skin of the patients. Typical radioactive isotopes include iodine-125, ytterbium-169, iridium-192, and cobalt-60. The main advantage of brachytherapy compared to EBRT is the higher relative dose to the tumor compared to surrounding healthy tissue; however, brachytherapy is limited in its ability to conform the dose to the tumor site. Recently, several electronic brachytherapy (EBT) sources consisting of miniature x-ray tubes with adjustable energy spectra generally below 50 kVp were commercialized. These sources are especially interesting during

intraoperative RT (IORT) for their ease of use as they are only emitting radiation when electrified.

LINACs can also be used to generate clinical electron beams by scattering the accelerated electron pencil beam instead of producing bremsstrahlung x-rays. Electron beams are less penetrative than their photon counterpart, in part due to their higher scattering probability, deflecting them away from their incident direction. They are reserved for superficial or shallow treatments within the first few centimeters of tissue.

In contrast, protons and other light ion beams (notably $^{12}C^{6+}$) are massive charged particles experiencing minimal scatter when traversing matter. They deposit a large amount of energy near the end of their range, resulting in a sharp peak in deposited dose. This so-called Bragg peak is the main advantage of ion beams over photons. A secondary advantage of ions resides in their ability to cause higher biological damage for the same dose level. This is a consequence of the differences in the patterns of energy deposition between these two radiation types. Whereas photons create energetic secondary electrons that tend to scatter, ions decelerate in a straight line, generating many low-energy electrons that stay relatively close to the primary particle. These differences in energy deposition patterns are characterized by the quantity linear energy transfer (LET), defined formally in Section 3.2, or by measurements or simulations of ionization clusters size distribution (ICSD) and their relevant statistical moments. In this context, ICSDs represent the frequency of occurrence of clusters of a given number of ionizations within a gas volume of a few nms irradiated by the field of interest.

In conjunction to its various therapeutic uses, IR also has numerous applications in visualization and diagnosis of cancer. X-ray imaging is frequently used in breast cancer screening in a technique referred to as mammography. This low-dose imaging methodology consists of passing ≈ 30 kV x-rays through compressed breast tissue to form an image of the breast composition. This image is then read by trained doctors looking for microcalcifications that may lead to clues on malignant breast conditions. X-ray imaging in the form of cone-beam computed tomography (CBCT) is also used during LINAC RT. In this case, a kV x-ray source is mounted on the gantry and rotates around the patient, while projection images of the interior of the patient body are acquired on an image detector mounted opposite to the source. A three dimensional image is then reconstructed and helps in determining tumor position, movement and shape just prior to RT treatment.

1.3.2 Radiation biology

Typical RT treatments on LINACs are given in 30 daily sessions over 6–7 weeks in order to allow for repair of the irradiated healthy tissue. During each of these sessions, the patient lies on the treatment bed for approximately 15 minutes while being irradiated by the rotating photon beam from the LINAC, which is collimated to the tumor site. Radiation ionizes and excites the atoms and molecules in its path through various physical interactions that are dictated by the radiation type, energy, and the material composition. Radiation may interact with human cells, ranging from a few μm to a few tens of μm in diameter. Their shapes are diverse and dependant on their type, but each (normal) cell has a nucleus where the DNA molecule resides. It is well known that damage to DNA, which

contains all the genetic information, is the main contributing factor causing harm to organs and tissues exposed to radiation. DNA is thus considered the primary target of radiation. While the vast majority of DNA resides in the cell nucleus, a small fraction is located in the mitochondria; therefore, most simulation work to date focused only on the cell nucleus. Recently, some work investigated the role of mitochondria in radiation-induced cell killing.[11] Within the cell nucleus, the DNA is arranged in organisational levels from the double-stranded helix to the chromosomes (for reference, see a model of the DNA molecule in Figure 5–3). The genetic information is coded in nucleotides in combinations of the four DNA nucleobases (adenine, guanine, thymine, and cytosine), connected by chemical bonds to the sugar-phosphate backbone composed of a deoxyribose molecule and a phosphate group. Each nucleotide has a complimentary nucleotide on the opposite strand, forming a base pair (bp) smaller than 3 nm. Figure 1–1 presents the dimensions of the DNA helix. Segments of 100–200 bps are wound around histone proteins and form nucleosomes. Then, nucleosomes are arranged into higher order structures called chromatin fibers with the help of scaffolding proteins. The addition of further scaffolding proteins to the chromatin allows further packaging of DNA into chromosomes. Except for germ cells, a copy of each chromosome exists inside the nucleus of each cell. Therefore, in typical normal human cells, more than 6 billion bps are fitted within a volume with diameter of a few microns and agglomerated in a number of organisational layers. When the radiation ionizes and excites the DNA molecule or its surroundings, it can induce various sorts of DNA damage.

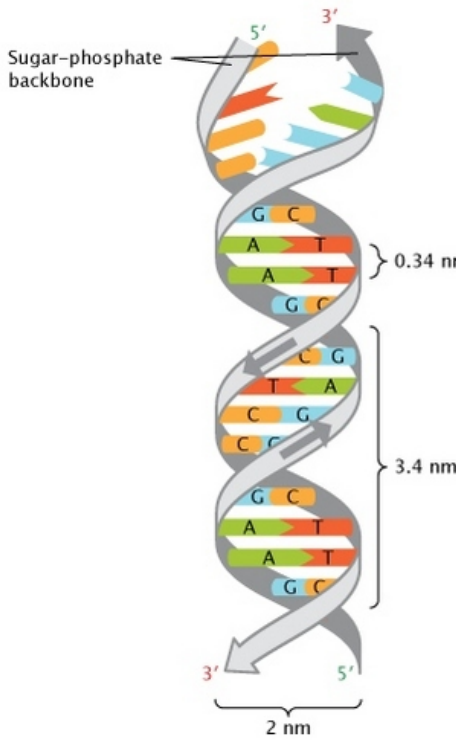

Figure 1–1: DNA double helix dimensions. Reproduced with permission from 12.

Direct radiation damage refers to molecular bond breakage by ionizations and excitations occurring directly in the atoms forming the DNA molecule. These lesions are formed within the first few *picoseconds* after the passage of radiation. However, the probability of the radiation interacting with DNA is very small, since DNA represents less than 5 % of the volume of a cell nucleus. Radiation-induced DNA damage can also arise via an indirect pathway. In this case, radiation produces reactive chemical species in the H_2O molecule and other biological material in the cellular environment. Most of these species originate from the dissociation of H_2O which accounts for 70 % of the cellular mass. For example, the OH free radical ($OH^{\bullet}$) can diffuse up to 6 nm away in a cellular milieu, a distance corresponding to roughly 2–3 times the diameter of the DNA molecule. These reactive species can recombine together or react with the DNA molecule and lead to indirect DNA damage. Typically, equilibrium in the chemical species is attained one *microsecond* after the passage of radiation, after which point no more indirect DNA damage occurs. In addition, endogenous cellular processes not related to radiation can also lead to DNA damage.

Radiation may break the strand, alter bases, destruct the sugar, and lead to the formation of cross-links and dimers. Strand breakage can occur either at the bond between the base and the sugar or in the phosphodiester bond. Nucleobases can be destroyed or chemically modified by radiation. Additionally, $OH^{\bullet}$ can add themselves to the nucleobase. Some of these lesions are produced by specific agents such as UV radiation for pyrimidine dimers. Generally we can categorize the elemental lesions, irrespective of their direct or indirect nature, as either a

DNA strand break (SB) or DNA base damage (BD). DNA SBs occur when the sugar-phosphate backbone (SPB) is sectioned, whereas DNA BD arises when the interactions affect the nucleobase. Generally, BDs are less toxic and require less complex cell machinery to repair than SBs. However, the hallmark of radiation is the formation of clustered DNA damage, where both SBs and BD can be present within a few bps (a few tens of nms), forming DNA damage clusters. In contrast to chemical agents such as peroxide, radiation-induced clusters of damage are much more difficult to repair. A single-strand break (SSB) cluster arises when one or more SBs occur on the same DNA strand within a few nms. A double-strand break (DSB) cluster occurs if each strand is broken at least once within a short distance. Additional BD can complement the SBs in both SSB and DSB clusters. Finally, clusters composed uniquely of BD (i.e., without SBs) are categorized as base damage clusters. Irradiation of a typical human cell to a dose of 1 Gy is expected to produce about 3000 BD clusters, 1000 SSB clusters and 40–100 DSB clusters. The number of lesions composing each cluster is a measure of the cluster's size or complexity. These definitions uniquely define each type of DNA damage, but other schemes were proposed.[13] These are reviewed more thoroughly in Chapter 4.

Cells have repair mechanisms to counteract radiation injuries depending on the type of lesions and cluster complexity.[14] Normal tissue cells are notably more effective at repairing DNA damage than tumor cells. In certain cases, the damage is too complex for repair, or the repair is incomplete, and the cell loses its ability to

divide (i.e., the cell dies). In a small proportion of cases, misrepairs may also lead to mutations and potentially to the induction of cancer.

The relative biological effectiveness (RBE) is a measure that relates the ability of a given radiation type to induce a given biological effect compared to a reference radiation, often ^{60}Co. It is mathematically expressed as $RBE = D_{ref}/D_q$, with D_q and D_{ref} being the dose necessary to reach a given biological endpoint with the given radiation quality q and the reference radiation quality, respectively. The biological effect can be expressed as a percentage of cell survival, induction of DNA DSB clusters, intestinal crypt regeneration in mice, foci formation, chromosome aberrations, or induction of micronuclei.[15] The RBE is an additional factor to take into account when prescribing a dose of radiation. For instance, the average RBE of proton beams with respect to MV photons is generally in line with a value of 1.1.[15] Relative biological effectiveness for the induction of DNA double-strand breaks (RBE_{DSB}) can be obtained with demanding irradiation experiments of cells or live animals. On the contrary, this work explores methodologies that would allow simulating RBE, notably RBE_{DSB}, using numerical models.

1.3.3 Numerical models

The most common and direct method to simulate radiation-induced DNA damage is by tracking radiation inside a cell nucleus, and accounting for interactions with DNA. Since it is known that DSB clusters are responsible for a large proportion of lethal cellular damage,[16,17] estimating the RBE_{DSB} is of great importance in modeling studies. These simulations require an event-by-event description of the radiation slowing down in water, used as a surrogate for biological

tissue, and the complete understanding of the molecular damage initially created by energy deposition (ED) sites (i.e., the individual ionizations and excitations of atoms/molecules) and knowledge of their spacial distribution along the radiation track.[18] This information can be obtained with deterministic one-dimensional models of ion transport based on the solution of the Boltzmann transport equation,[19,20] with amorphous track codes based on the Katz track-structure model and theory of RBE,[21] with condended history Monte Carlo (MC) codes providing averages over discrete particle histories,[22,23] or with Monte Carlo track structure (MCTS) codes that can provide the full distribution of ED sites in 3D space.[24,25] Out of all these methods, MCTS offers the higher resolution and precision in the simulations especially when coupled with geometrical models of the DNA molecule. While initial work by Berger *et al.*[26] on the penetration of charged particles laid the path for condensed-history MC, it is not until the 1980s that MCTS codes became available.[27] When coupled to a 3D model of the DNA molecule with all its intricacies, the simulated ED sites become surrogates for DNA damage, traditionally scored as such only when a threshold energy is deposited within the DNA sugar-phosphate backbone or nucleotide volumes of the geometrical model.

The joint MCTS-DNA model simulations are carried out in a water environment with specific subregions representing the DNA atoms or molecules of relevance for DNA damage. Alternatively, an algorithm based code making use of published primary DNA damage data, the Monte Carlo Damage Simulation (MCDS) code,[28–31] provides fast and relatively accurate estimates, while replacing the exact DNA modeling by a randomized approach. Lastly, radiation-induced

DNA damage can be estimated using combinatorial mathematics relying on ICSDs measured experimentally within the low-pressure gas cavity of an ion counting nanodosimeter or simulated with MCTS simulations.[32]

1.4 Hypothesis, objectives, and motivations

Conventional RT courses consist of ≈ 30 daily mega-voltage x-ray fractions of dose over a one month period, a regiment established on the biophysics of tissue response to radiation damage, largely tested and verified in the clinic for the past half-century. Theoretical and technical advances allow for novel and drastic treatment options, at the expense of a loss in this historical medical confidence. The advantages provided by proton and ion beams, targeted radionuclide therapy, nanoparticles, micro-beams, or hypo-fractionation are typically predicted from basic understanding of theory and experiments and yet to be proven in randomised clinical trials. Efforts in the field of dosimetry have resulted in precision to within a few percent in the physical dose, which is an order of magnitude better than most uncertainties involving biological experiments or clinical outcome studies. The next logical step towards accurate RT outcome prediction is to apply physics-thinking to biology and attempt, at least partially, to reduce uncertainties in our understanding of biological phenomena. The multiscale response of tissue to radiation is fundamentally reliant on the initial patterns of energy deposition of the radiation. Therefore, physics-propelled simulation tools, based notably on MC techniques and dosimetry experiments, were designed to guide biologists and clinicians by providing, first principles, mechanistic predictions of radiation-induced DNA damage.

The hypothesis of this thesis is that numerical simulations of initial DNA damage following irradiation can predict the RBE and can complement experiments for studying novel irradiation sources and techniques, predicting treatment outcomes through multiscale modeling, and in absolute dosimetry of hadron beams. Conducting experimental validation is often very costly, uncertain, impractical, long, and difficult to adequately control for all variables. On the other hand, numerical simulations driven by adjustable parameters estimated from selected experimentally measured SSBs and DSBs yields, can complement these experiments. Nevertheless, a thorough standardization in the methodologies used in the simulations is needed. Currently simulations also have a large uncontrolled variability due to a lack of standardization in the terminology used to simulate DNA damage, conceptual errors in user-defined source and phantom choices, and fundamental differences in the cross-sections and models used for physical interaction simulations. Therefore, the specific objectives of this thesis are:

- To improve on existing and develop new, potentially faster and statistically more accurate, methods for the simulation of DNA damage, particularly DNA SSBs and DSBs, following irradiation with a variety of techniques used in RT.
- To characterize the sensitivity on the final DNA damage results of several key simulation parameters and assumptions relating to the simulated radiation sources.

- To showcase and explain several conceptual errors and inaccuracies in previously published work in order to achieve a better standardization of the field of DNA damage simulations.

To attain these objectives, Chapter 4 presents a methodology based on MCTS simulations overlaid on a geometrical DNA model for simulating the direct DNA damage following electron irradiation. Chapter 5 builds upon and adapts this methodology for proton and other light ion beams. The methodology considers algorithms that decrease simulation time by reusing pre-generated data, hence allowing the gathering of results with an increased statistical significance. Importantly, this work allows the study of the influence of the irradiation angle, of the SB energy threshold, definitions, and clustering schemes, and of the source geometry and dosimetric effect within the cell nucleus. On the other hand, in Chapter 6 the RBE_{DSB} of several clinical photon sources below 1 MeV is obtained by adapting the techniques of Chapters 4 and 5 and replacing the geometrical DNA model with a probabilistic approach. The fields considered include CBCT, EBT, and selected radio-isotopes. Throughout these three chapters, it was found that previously published DNA damage yields were falsely estimated, either due to wrong assumptions or conceptual errors, and that there was a lack of standardization in the simulation methodologies among different groups.

Simulations of the initial DNA damage has direct applications in at least three areas of study, namely comparisons of the RBE of various treatment techniques, multiscale modeling of radiation response, and biologically relevant absolute dosimetry.

Novel treatment techniques

Over the past few years, the clinic has witnessed an increased use of particle beam treatments. In addition, several completely new types of treatments have also become available with unconventional radiation fields. Dosimetry of these fields is a challenge, but so is the knowledge of their biological contribution when compared to historical treatments using MV photons. Below approximately 50 kVp, photon sources become as much as 10 % more effective at cell killing and DSB induction than their MV counterparts. Miniature x-ray sources used in EBT or IORT are becoming widely available and used instead of traditional MV photon or electron treatments for selected treatment sites. Increased use of mammography screening and on-board imaging during RT treatments may lead to increases in secondary cancer incidence. Treatments involving radio-pharmaceuticals and/or high atomic number nanoparticles lead to the generation of very low energy photons or electrons in the vicinity of the DNA, as well as alter the radiation chemistry and radical production. Proton and other light ion beams offer advantageous dosimetric properties, but necessitate much care, as the biological effectiveness depends on the energy, atomic number, dose-rate, dose level, depth of treatment, and cellular composition. In this work, the ability to induce DSB clusters by electrons, proton and other light ions, and clinical photon beams is studied in Chapters 4, 5, and 6, respectively.

Multiscale modeling of the radiation response

IR acts on cells through successive physical, chemical, and biological processes. These chain reactions span several orders of magnitude on both the time

and space scales, and are responsible for cell mutations and cell death. Predicting these pathways and their final outcomes is crucial, particularly in radiation oncology for tailoring patient-specific or disease-specific treatments. Radiation interacts almost instantaneously with living samples; however, its effects span many time and space scales. Through ionizations, excitations and other physical processes, molecular bonds are broken, very reactive species such as $OH^{\bullet}$ are created and diffuse and react throughout the sample, proteins are activated to repair the DNA damage that occurred, cellular apoptosis might be initiated and different cellular signals are sent to surrounding cells. The medical response is quantified by assessing the tumor control *versus* normal tissue damage. Furthermore, radiation is known to be the cause of late normal tissue complications and even secondary cancers, which may occur tens of years following the irradiation. This complex chain of events is initiated by a surprisingly small total energy deposited inside the cell, barely enough to raise its temperature by about 0.5 mK. Fundamentally, the patterns or clustering of the individual ED sites are responsible for the DNA damage. Physical phenomena such as atomic and molecular ionizations and excitations are therefore the initiators of the chain reactions. Modeling of RBE across multiple physical, chemical, and biological scales can help interpret laboratory experiments and can provide insight into methodologies for safe and effective therapy. However, this challenging modeling task needs to rely on complementary experimental and theoretical findings. Experiments drive accurate theoretical models and simulations allow an insight into experimentally non-measurable quantities. For instance, measurements cannot provide the exact location of the

lesions in the DNA, whereas this type of information is frequently simulated. Treatment outcome modelling often follows a top-down approach, where past clinical experience drives research in genetics or biology. Novel predictive models using large data sets including genetic markers and/or medical images are being developed in the context of personalized medicine, and could allow doctors to make informed decisions based on probabilities of success. Often, these models lack detailed, rational or first-principle reasoning underlying the predicted outcomes; clinical decisions are thus faith-dependant on the statistical model. On the contrary, bottom-up modeling starts from the physical initial radiation insult and builds each step in order to construct a unified model. This approach has the advantage of offering a clear understanding of the reasons leading to a specific result at the expense of some simplicity. However, the data and models able to translate these findings into clinical outcomes are still preliminary at best. The present thesis attempted to shape order in methodologies used in simulating the initial radiation-induced DNA damage, while focusing on the physical scale.

Biologically relevant absolute dosimetry

Absolute dosimetry is interested in the knowledge of the dose delivered in absolute terms by a given RT modality. It is typically done in controlled reference conditions. Metrological standard labs calibrate the dosimeters for further use in the clinics in order to keep the dose definition as constant as possible throughout different clinics, often within about 2%. This is a crucial point as dose delivered in RT treatments needs to stay within ±5% of the prescribed dose in order to keep the

same effectiveness. However, dose itself is often not enough to reflect the effectiveness of a given treatment, especially in the case of particle therapy. Therefore, accurate calibration in terms of absorbed dose does not guarantee that the same treatments are delivered in different clinics. On the other hand, nanodosimeters can measure physical beam quantities relatable to initial DNA damage patterns with proper combinatorial models.[32] It is also believed that biologically relevant damage patterns can be used as tools for the definition of a standard reference dosimetry calibration procedure of proton and light ion therapeutic beams.[33] In fact, in a joint European research project, Biologically weighted Quantities in Radiation Therapy (BioQuaRT),[33] efforts are made to construct tools needed to create a multiscale framework that would allow the dosimetric standardization of those beams, and potentially even lead to a change in the current prescription paradigm based on physical dose. Linking experimentally measured and MC simulated ED sites and their associated radiation damage patterns would be a step in this direction. The models presented in this thesis would allow linking experimental measurements of ICSDs, such as those obtained by the nanodosimeters presented by the BioQuaRT project, to relevant initial biological radiation damage patterns in DNA such as SSBs and DSBs. This could provide a straightforward comparison of experimentally measured ionization clusters and MC simulated tracks, allowing a description of beam radiation quality and, potentially, biological effectiveness.

CHAPTER 2
Simulating DNA damage

Several groups attempted to describe and predict the induction of DNA damage by radiation. This chapter presents the methods involved in conducting such simulations using MCTS simulations. Section 2.1 consists of a brief introduction to MC methods and more specifically to the GEANT4 toolkit and Geant4-DNA extension.[23,24] It is followed by a description of a large project to develop models for the risk of cancer and circulatory diseases by the National Research Center for Environment and Health (GSF) in Germany, that lead to the creation of the PARticle TRACking (PARTRAC) MCTS code.[34] This code considers the event-by-event action of radiation on a geometrical representations of DNA, consisting of an arrangement of spheres representing each of the constituting atoms. Next, Section 2.3 portrays the methodology and key results obtained with the MCDS algorithm for the generation of clusters of DNA lesions using a probabilistic model.[28–31] In this approach, neither the track structure nor the DNA model are explicitly simulated; however, surprisingly accurate results are produced with this very fast computational technique. Finally, a combinatorial DNA damage model revolving on ICSD experimentally measured or simulated with MCTS simulations is briefly explored in Section 2.4.[32]

2.1 Monte Carlo (MC) methods for particle tracking

The MC technique makes use of random variable sampling to solve mathematical equations. It is widely used to solve the radiation transport equations that can be modelled using probabilities of radiation to interact with matter. These probabilities, called cross-sections, are known from quantum mechanical considerations and validated through large series of experiments. There are two general possibilities when conducting MC simulations with charged particles. In general purpose MC simulations, individuals energy depositions of the charged particles are not directly simulated; instead, their global, condensed effect on a portion of the matter is simulated allowing for much faster simulation while preserving macroscopic properties of the fields. These codes are often used in radiation therapy planning in order to calculate doses in macroscopic tumors in a precise and timely manner. On the other hand, event-by-event MC techniques, explicitly simulate each and every interaction of the charged particles, which makes them well adapted to studying the effects of radiation on small structures such as DNA.

2.1.1 Basics of event-by-event MC methods

MCTS codes simulate the discrete events encountered by an incident particle in water. In RT, the incident particles of interest are kV–MV photons and electrons, protons, and different low atomic number ions and atoms. All these particles lose energy as they travel through water by setting in motion a shower of secondary electrons that will, in turn, ionize and excite H_2O molecules until all their energy is dispersed. To adequately follow the event-by-event slowing down

of secondary electrons, detailed knowledge of the ionization, excitation, and elastic scattering interactions of electrons in H_2O is needed down to very low energies ($\approx$ 10 eV). These processes are partially described by their total and doubly differential cross-sections, which provide the mean-free paths to the next interaction, the type of interaction, the energy deposited and the energy transferred to the secondary particles, and the angles of emission of the incident and secondary particles. Experimental data of water cross-sections for low-energy electrons (< 1 keV) is available for the vapor phase and is the basis of many models derived from first principles. However, experimental data for condensed states of H_2O (liquid, ice) is scarce, and cross-sections over the entire span of energy loss and momentum transfer are not adequately known. MCTS codes have to either rely on the vapor phase cross-sections measurements or on a mixture of models originating from vapor and liquid phase measurements.

Recently Thomson and Kawrakow[35] criticized the use of low energy simulations in liquid water below about 100 eV, because at those energies and dimensions, the electrons are delocalized. In fact, it is not possible to know the position and energy of these electrons to within 30% without violating the Heisenberg principle. Therefore the authors interpreted that trajectory methods are not applicable at these energies. Nevertheless, as was later shown by Liljequist and Nikjoo, this reasoning is too simplistic and not correct.[36] In fact, the trajectory of a given low energy electron should be seen as a one contribution to an approximate modelling of the results of many different results of low energy scatterings. In any case, large

theoretical and experimental uncertainties are to be expected when comparing simulated to actual electron trajectories.

2.1.2 GEANT4 and GEANT4-DNA

The GEANT4 general purpose Monte Carlo simulation toolkit[23,37] is continuously being extended with physical, chemical and biological models in order to simulate cellular and subcellular damage induced by ionising radiation. The GEANT4-DNA project[24,38] has successfully incorporated a new set of electromagnetic processes able to track low energy electrons , protons, alpha particles, and several ions. Although, it's a new player in the MCTS field, GEANT4-DNA is open-source and supported by many worldwide collaborators already producing promising results. The latest version of GEANT4 (version 10.2) is freely available on their website. The GEANT4-DNA processes simulate explicitly every interaction without relying on condensed history techniques, as these are inappropriate for biological simulations of interactions with DNA. The GEANT4-DNA processes include elastic scattering, electronic excitation and ionisation of the water molecule, dissociative electron attachment below ≈ 100 eV, and excitation of vibrational modes of the molecule of the order of the milli-electronvolt. The interaction processes used in GEANT4-DNA are based on semi-empirical models and on the plane-wave Born approximation.[39] The creation of detailed DNA and cellular geometries is possible with the GEANT4 *detector construction* class, as demonstrated by Bernal *et al.*[40] Recently, GEANT4-DNA was extended to allow the possibility of tracking the diffusion of radiolytic products and their mutual interactions in liquid

water following the initial physical interactions of charged particles.[41] This approach allows the modeling of chemical radiolysis up to 1 μs after irradiation and thus the production of indirect damage to biological sub-units by oxygen reactive species. In GEANT4-DNA, the electron track structure is simulated by following a step-by-step recipe. First, the distance travelled by the electron (i.e., its free path) is determined by direct MC sampling of the total interaction cross-section, $\phi = \sum \phi_i$, where i represents all the possible interaction processes: ionization , excitation, dissociative electron attachment, vibrational excitation, and elastic scattering. Furthermore, the type of interaction is decided by sampling the relative magnitudes of the individual cross-sections. The final states of the local energy deposited, the energy transferred to the secondary particles (if any), and the angles of emission of the incident and secondary particles are calculated.

Photon interactions

In this work, the Livermore models for the photoelectric effect, Rayleigh scattering, and Compton scattering processes were used to track photons below 1 MeV. The Livermore models exploits the Evaluated Photon Data Library 1997 (EPDL97)[42] to track photons in the range from 250 eV to 100 TeV. The library provides the doubly differential and total cross-sections for sampling the final state after interaction. The pair-production process, in which an electron-positron pair might be created, is impossible below 1 MeV and was not studied.

Elastic scattering

Charged particles may scatter elastically by transferring momentum, but not energy, leading to changes of direction without energy loss. The scattering angle

depends on the masses of the particle present; therefore, elastic scattering is generally neglected for proton and light ions. For electrons, angular differential and total elastic cross-section are calculated theoretically using the partial wave expansion. The GEANT4-DNA project offers two models of the elastic scattering of electrons by liquid H_2O. The screened Rutherford analytical model is fully described by Emfietzoglou *et al*.[43] The final state is given by the Rutherford differential scattering cross-section with an additional screening term to account for the electron cloud. Below 200 eV, a fit to experimental data is used instead.[44] This model was developed for the vapor phase and extended to the liquid phase, which is the norm in the field, because of the scarcity of experimental data available and untested assumptions in the liquid phase. It is expected to predict angular deflections not very different from those actually occurring in the liquid phase. The second model available in GEANT4-DNA for elastic scattering of electrons (Champion model) is based on quantum mechanical calculations in the partial-wave framework and was specifically developed for the liquid phase of water.[45] It is based on a spherical potential including a static contribution (deduced from experimental electron density measurements[46]) and two fine structure correction terms corresponding to the correlation-polarization and the exchange interactions. Overall, the Champion model is valid from 10 eV to 10 keV and shows improved agreement with available experimental data than the screened Rutherford model.[24] It is the default model in GEANT4 and the only one used in this work.

Inelastic scattering

The models for electronic excitations and ionizations in GEANT4-DNA are based on the first Born approximation and are making use of optical oscillator strength measurements.[47] They also incorporates semiempirical corrections for low energies as presented by Emfietzoglou and Nikjoo.[48,49]

Vibrational excitation and dissociative electron attachment

GEANT4-DNA considers two additional processes at very low energies for electrons. Below 100 eV, electrons may lead to molecular vibrational excitations by depositing a few *milli*-eV of energy directly to the water molecules. The cross-section data for this process were taken directly from experimental measurements by Michaud and Sanche in amorphous ice.[50] Additionally, between 4 and 13 eV a model for electron dissociative attachment in liquid water is available.[51] These processes become the dominant energy loss processes below 10 eV.

Validation

GEANT4-DNA was thoroughly validated in the past 10 years by comparisons with various set of simulation and experimental data.[38] Qualitative comparisons of the cross-sections with experimental data in the gas phase were performed[24] Reasonable agreement between ICSD, frequencies of energy deposition, lineal energy measured with GEANT4-DNA and other simulation results was also obtained.[52–54]

2.2 The PARTRAC experience

The biophysical code PARTRAC follows the mechanisms of interaction of radiation with matter from the early stage.[55] It is a well established code that can model the formation of initial DNA damage and its successive evolution. The

computation of DNA damage in PARTRAC comes from the superposition of simulated particle tracks on an atomistic description of DNA. Separate modules exist for the radiolysis of water, the diffusion of radical species, and their interactions with DNA. Additionally, two modules were recently introduced for mechanistic simulations of the non-homologous end joining (NHEJ) DNA repair mechanism and the subsequent calculation of kinetics and yields of radiation-induced chromosome aberrations.[56,57]

The PARTRAC code offers the most advanced simulations in the field and is therefore considered as the gold standard. In this thesis, we referred a number of times to work done with the help of this code. In chapter 4 PARTRAC's SB definitions and clustering algorithms were examined and compared to other schemes. In chapter 5, the RBE_{DSB} obtained for protons was compared to that obtained with the novel methods presented. It is therefore natural to have an in-depth section reviewing the methodology and significant results obtained with this code. However, PARTRAC is not open-source and was never distributed outside of a select group of researchers. The comparisons between the work presented in this thesis and PARTRAC are simply based on published data.

2.2.1 Methods and description

The physical module

This module focuses on the transport of IR and the subsequent creation of ionized and excited atoms or molecules. The EPDL97 provides the relevant atomic cross-sections for tracking of photons in media with arbitrary elemental composition.[42] As in GEANT4, coherent scattering, photo-electric effect, Compton

scattering, and pair-production, as well as the production of Auger electrons and fluorescence photons are considered. Electrons are tracked exclusively in liquid water from 10 eV to 10 MeV using excitation and ionization cross-sections obtained in the first-order plane-wave Born approximation formalism coupled to the dielectric theory, with a semi-empirical correction below 500 eV.[58,59] Electron exchange and elastic scattering are also taken into account. Proton transport between 1 keV to 1 GeV considers ionization, excitation, electron capture, and electron loss, while elastic scattering is ignored. Relevant cross-sections were calculated in the first-order plane-wave Born approximation above 1 MeV. Semi-empirical models were used instead below that energy threshold. For heavier ions, proton ionization and excitation cross-sections are scaled by the particle's effective charge.[60] However, this methodology was recently criticized for underestimating energy loss and mean ionisation cluster sizes.[61]

The prechemical and chemical modules

Ionized and excited water molecules decay into $OH^{\bullet}$, $H^{\bullet}$, H_3O^+, and H_2, following predefined branching ratios dependent on the excited level or the ionized shell. Furthermore, electrons with kinetic energy below 10 eV are converted into hydrated electrons (e^-_{aq}) via attachment of water molecules. Therefore, PARTRAC does not simulate vibrational excitation nor dissociative electron attachment as GEANT4 does. The generated species then undergo diffuse Brownian motion for about 1 micro-second and react with one-another, generating further products (OH^- and H_2O_2). After each diffusion step of $\approx$ 30 pico-seconds in a random

direction, the reaction between partners are assessed on specified reaction radii from observation of radiolysis reaction rates.

The DNA model

On the smallest scale, the nucleotides composed of a desoxyribose, a phosphate group and a nucleobase (adenine, guanine, thymine, or cytosine) are modeled as the union of spheres representing the atoms (H, C, N, O, or P) composing each of these constructs. The radius of each sphere corresponds to the Van der Waals radius of each atom times a factor of 2. This factor represents the hydration layer of the DNA; a certain number of water molecules are always bound to the DNA at all times, and interactions in these water molecules and in the nucleotides themselves are modeled as non-scavengeable damage instead of being processed by the chemical module. Special rates in the chemical module are used to assess interactions between DNA constituents and chemical species. A great feature of this code is that all the structural DNA levels, including the DNA double helix, the nucleosomes, the chromatin fibers, fiber loops, chromatin domains, and chromosomes are considered; however, the models have considerably changed over the years and modified for different applications. In the most recent iteration,[62] the cell nuclei are modelled as spheres or ellipsoids composed of 6070 of spherical chromatin domains of 1 Mbp, linked together by entropic spring potentials (for further details on the PARTRAC DNA models, see Refs. 62–64 and references within).

DNA damage module

SBs in PARTRAC are distinguished by their nature, scavengeable or not. Scavengeable SBs are those created by reactions of chemical species with DNA. Non-scavengeable SBs comprise those created in the direct interactions of radiation with the DNA constituents, including the hydration layer, in which case the damage is coined *quasi-direct*. The creation of DNA SBs at a given nucleotide position is based on the sum of the energy deposited within the desoxyribose and phosphate group. The thresholding function was adjusted over the years, notably in light of DNA induction measurements by low energy photons and electrons (down to about 5 eV)[65,66] The current approach is to assign a probability of SB induction of 0 for total energy deposits below 5 eV, of 1 above 37.5 eV or 40 eV, with a linear interpolation between. This function was actually parametrized in order to obtain a non-scavengeable total strand break (TSB) yield of 350 $Gy^{-1}cell^{-1}$ for ^{60}Co. As for scavengeable SBs, the parametrization yielded a SBs induction probability of 65% for a $OH^{\bullet}$ interacting with the desoxyribose or with the phosphate group. A DSB is induced if 2 SSBs on opposite strands occur within 10 bp. An additional empirical transformation of 1% of SSBs into DSBs is implemented, to account for the potential transfer of radical sites from one broken strand to the opposite strand.[67]

2.2.2 Important results

After cell irradiation, it is possible to extract DNA fragments and measure their size distributions. Experimental distribution of fragment sizes were compared to PARTRAC simulations for ^{60}Co gamma rays and for protons, with overall

very good agreement. In addition, it was found that simulations allow to study the production of fragments ≤ 1000 bp, which are not revealed by conventional experimental methods using gel electrophoresis for instance.[68] Moreover, it was found that at higher LET, the number of short fragments increases.

PARTRAC was used to obtain the SSB and DSB yields in a cell irradiated by proton beams with LET between 1.6 and 70 keV μm^{-1}. The study showed that the SSB yields decrease and the DSB yields rise with LET. The RBE_{DSB} steadily rose up to 2.2 for 0.5 MeV protons. Fragment size analysis showed that experimentally unresolved fragments account for 30% of the simulated RBE_{DSB} and that a random breakage algorithm to determine DSB genomic positions and fragment sizes is inappropriate.

2.2.3 Discussion in context of present work

As previously stated, this code lived through a series of reparametrizations, changes in the DNA geometries, modification in SBs definitions and classifications, and additions of higher-order modules throughout the years. These changes were motivated mostly by new experimental data that became available. In Chapter 4, many references to publications using the PARTRAC code will be made. Particularly, we attempted to compare our methodology for the simulation of DNA damage by photon beams to a reference paper in the field.[25] Unfortunately, all the details of the PARTRAC implementation were never fully published. For instance, details such as the counting of two adjacent SBs on the same strand as a SSB cluster of size 1, the size of the beam and the size of the uniform homogeneous field used to irradiate the cell nuclei, and the exact algorithm for the classification of DSBs

were missing. We obtained these details through extensive communication with the authors. In itself, this might seem like a small problem; however PARTRAC was used as a gold standard by numerous other simulation studies, each of which redefined their own methodology and terminology. Therefore, direct comparisons with PARTRAC were flawed. The field seemed to lack standardization and interest in explaining differences between simulations. Most differences were attributed to the differences in the codes used (cross-sections for instance) or ignored by referring to the large experimental uncertainties. In Chapter 4, we considered several other reasons why differences in simulated data were so large.

In Chapter 5, we devised a methodology to obtain the relative biological effectiveness for the induction of direct DNA double-strand breaks (RBE_{DSB}^{direct}) for protons and light ions and compared it to PARTRAC.[63,69] One limitation of the PARTRAC code and/or literature resides in the beam specification. Often beam dimensions and orientation with respect to the cell nucleus are not specified. Moreover, the simulations are conducted in a single-cell geometry and influences from neighboring cells are not considered. In other words, simulations do not reflect clinical situations where charged particle equilibrium (CPE) is present throughout the cell and part of the dose to the cell comes from scattered radiation. These two limitations are further discussed in Chapter 5 and 6. Also, the next section present a code that is partially based on PARTRAC's simulated DNA damage yields, but can also be paired with general purpose MC simulations for the determination of the exact spectra influencing a given cell.

2.3 The MCDS experience

The MCDS code is an algorithm for the simulation of DNA damage under the effect of photons, electron, alpha particles and other atoms and ions up to 1 GeV.[28,29,31] MCDS accounts for the energy loss by electrons (and the secondary electrons they create) within the cell nucleus. The methodology relies on interpolation of damage yields from computationally expensive MCTS simulations. The code outputs DNA SSB, DSB and BD cluster yields per cell, as well as cluster complexity. It was previously used to obtain RBE_{DSB} estimates for several x-rays and radio-isotopes,[70] for kV CBCT,[71] for EBT sources,[72,73] in the context of Auger electron therapy with TRT,[74] for therapeutic proton beams,[75] and even recently for various neutron fields.[31] The MCDS algorithm is implemented in a computer program freely available online. Typical simulations take less than a few minutes, which is 1–2 order of magnitude faster than equivalent MCTS simulations.

The MCDS algorithm has three adjustable parameters independent of the type and energy of the particle (σ_{SB}, f, and N_{min}) and one parameter (n_{seg}) which depends on both the type and energy of the incident particle. The number of SBs ($\sigma_{SB} = 1300$ Gy^{-1} cell^{-1}), the ratio of BDs to SBs ($f = 3$), and the maximal distance between lesions to be considered in the same cluster ($N_{min} = 9$ bps) were all optimized by comparison with fitted data for SSB and DSB yields from selected MCTS simulations.[29] The fourth parameter, n_{seg}, is an *ad hoc* DNA length with,

$$n_{seg} = 149\,200\frac{123\,600x}{x+267}\text{bp Gy}^{-1}\text{ cell}^{-1}, \quad x = \frac{Z_{eff}^2}{\beta^2}, \tag{2.1}$$

where Z^2_{eff} is the effective charge and $\beta = v/c$, where v is the particle speed and c is the speed of light. Equation 2.1 was optimized using experimental and simulation DNA damage induction yields of particles of varying $\frac{Z^2_{\text{eff}}}{\beta^2}$. In MCDS, the particles are not transported. Instead, their effect on DNA is estimated by modulating the n_{seg} parameter representing the length of a DNA segments in which σ_{SB} SBs and $f\sigma_{\text{SB}}$ BDs are randomly distributed. First, SBs and BDs are assigned at random in the DNA segment. This involves (1), the selection of a random bp in $[1, n_{\text{seg}}]$, (2), the selection of a random strand, and (3), the assignment of one SB on the selected bp/strand combination until none are left to be placed. The procedure is then repeated for the $f\sigma_{\text{SB}}$ BDs with the provision that a given bp/strand combination cannot contain more than 1 SB and 1 BD. In the second part of the MCDS algorithm, the DNA lesions are grouped into clusters. A cluster comprises all lesions (bps and BDs) within $N_{\text{min}} = 9$ bp of one another. Some clusters can therefore contain only one lesion. Finally, each cluster is categorized exclusively into a DSB cluster if 2 SBs on opposite strands are found within 10 bp, an SSB cluster if it is not a DSB cluster and it exhibits at least one SB, or otherwise, in a BD cluster. In addition to the category, the cluster size (i.e., the total number of SBs and BDs in the cluster) is also reported.

It is well documented that higher LET particles create more complex lesions. This increased clustering is achieved in MCDS by modulation of n_{seg}. For higher LET particles, n_{seg} is shorter and therefore lesions are clustered closer together,

effectively inducing more complex lesions and more DSBs. It is somewhat surprising that such a simple approach is able to reproduce more complex MCTS simulations.

The MCDS algorithm was initially created for monoenergetic charged particles and the parameters were adjusted based on a few MCTS simulations not covering a vast therapeutic range.[28] It was later reparametrized to cover a larger range of particle types and energies, by integrating data from an extended list of electrons and proton simulations.[29] It was since upgraded to also simulate the effect of polyenergetic spectra of charged particles.[70] This is achieved by inputting a list of energies and frequencies. The frequency needs to be proportional to the dose-effect of each particle and not to the particle fluence. This is formally defined as was shown in Ref. 70. The yield Σ_i of the ith type of cluster (i.e., SSB, DSB or BD clusters) can be computed as

$$\Sigma_i = \frac{\int_0^\infty \mathrm{d}E \Sigma_i(E)\Phi(E)\bar{z}_F(E)}{\int_0^\infty \mathrm{d}E \Phi(E)\bar{z}_F(E)}, \tag{2.2}$$

where $\Phi(E)$ is the energy fluence of charged particles, $\Sigma_i(E)$ is the initial yield of the ith type of cluster for particles of energy E, and $\bar{z}_F(E)$ is the mean specific energy in the cell. The integrals are taken over the whole particle fluence spectrum. This formulation explicitly states that the particle fluence weighted by its dose contribution to the cell is required in order to produce meaningful average DNA damage yields. Photons and neutrons are handled indirectly by inputting the spectra of the secondary particles they put in motion.

In addition, MCDS allows the adjustment of the cellular oxygen concentration ($[O_2]$), which affects the total strand break yields. The anoxic ($[O_2]$ = 0%), normoxic ($[O_2]$ = 20%), and aerobic ($[O_2]$ = 100%) states of oxygenation, and anything in between, are modelled using a probability function (see Eq. (7) in Ref. 30). As the cellular oxygen concentration increases, the number of DNA damage increases. For instance, for low LET radiation below $\approx 0.3\,\mathrm{keV}\,\mu\mathrm{m}^{-1}$, cells irradiated under normal cellular oxygen concentrations sustain 2.9 times more DSBs as those irradiated under full anoxia.[30]

2.3.1 Important results with MCDS

Recently, Stewart *et al.*[31] proposed an approach to integrate the cell level DSBs yields generated with the MCDS code into a general purpose MC program, Monte Carlo N-Particle (MCNP). This integration is warranted by the need to calculate voxel-based RBE in treatment planning systems in large scale systems such as the body while conserving CPE conditions as they occur in many RT treatments. In this work, the authors showed that the RBE_{DSB} for ^{137}Cs is 1.7% higher than that for ^{60}Co. In addition, 60–250 kV x-rays were between 1.1 and 1.25 times more efficient at inducing DNA DSBs than ^{60}Co in normoxic conditions.

Kirkby *et al.*[71] presented a method for coupling MC radiation transport to DNA damage simulations and applied it to the simulation of the RBE_{DSB} for CBCT x-rays (80 and 125 kVp) relative to ^{60}Co gamma rays. The authors focused on determining the spectrum of electrons incident on the cell nucleus when irradiated by photons in a much larger region of tissue. The MC tracking capabilities of the Penetration and ENErgy LOss of Positrons and Electrons (PENELOPE) MC

code and Eq. 2.2 were used to obtain the kinetic energies of electrons incident on a cell and the dose they deposit in the cell. To speed up the calculation, the reciprocity theorem was used and regions with unequal E_{cut} were defined. The reciprocity theorem[76–78] states that for the purposes of determining absorbed dose or fluence in the volume, a small volume irradiated by a large beam, or a pencil beam irradiating a large volume are equivalent; the latter solution being computationally less demanding. In addition, the electron tracking cutoff E_{cut} was lowered to 50 eV around the scoring volume, but was set at 10 keV further away. The methodology led to an RBE_{DSB}= 1.1 for 80 and 125 kVp photon beams, with no significant change with depth, filtration, or cellular oxygen concentrations.

In addition, the MCDS code was used to investigate the RBE_{DSB} at various depths in a phantom irradiated by clinical proton beams.[75] The authors notably found that the biological range of the proton beam is extended by 1.9 mm beyond the physical Bragg peak due to secondary particles. In a study looking at spectral differences in and out of the field of a 10 MV clinical photon beam, MCDS simulations showed that the MCDS varied by less than 1%.

2.3.2 Discussion

MCDS is an easy to use, very fast, and accurate code that generate DNA-level damage to nucleotides. The accuracy of the code in predicting DNA damage absolute yields and trends as a function of energy, particle type, oxygen concentration and dose was previously investigated. It is often used as a benchmark for other simulation codes or experimental measurements. The trends of DNA damage as a function of LET, particle type, energy, and cellular oxygen concentration are

accurately reproduced for many particles. The DNA damage yields are expressed per Gy and per cell. As previously explained, MCDS does not track particles inside a cell. Instead, the effects of the particle incident on a cell (and all the potential secondary particles it emits in the cell) are estimated based on adjustable parameters fitted to MCTS simulations, notably the PARTRAC code. Inherently, MCDS accounts for the slowing down of particles inside the cell. Only the incident spectra and type of particles are of importance.

Except in specific cases such as irradiation of small thin cell cultures in vaccum or for the purposes of theoretical investigations, single-particle type monoenergetic beams do not exist. MCDS requires the spectra of all radiation types and energy incident on a given cell. For instance, in the case of a 160 MeV proton beam, the fluence of primary protons, secondary protons, alpha particles, deuteron ions and other less frequent products are needed for an accurate simulation.[75] These spectra are often obtained with general purpose MC codes.

A central part of the MCDS algorithm, and probably the least understood one, is the relationship between the yield of a given damage type and the mean specific energy, as depicted in Eq. 2.2. Take for example a hypothetical cell irradiated by electrons of 100 eV and 1 MeV in equal proportions. According to MCDS, 100 eV and 1 MeV electrons are responsible for 24.9 DSBs/Gy/Gbp and 8.2 DSBs/Gy/Gbp, respectively. In order to compute the DSB yield for the entire spectrum, it is not enough to simply average these numbers, because a 100 eV electron will deposit less dose in the cell (therefore create less DSBs/Gbp) than a 1 MeV electron. Therefore, the fluence of incident electrons needs to be weighted by

the dose-effect or each electron, i.e., the mean specific energy. In the hypothetical case, the dose-weighted DSB yield would be only 11.4 DSBs/Gy/Gbp, whereas it would be 16.6 DSBs/Gy/Gbp if the mean specific energy was ignored. In most situations, ignoring the mean specific energy artificially increases the calculated DSB yield. Chapter 6 presents some situations from the literature where this inconsistency leads to much higher than anticipated RBE_{DSB} values.

As explained earlier, MCDS assigns randomly a SB (or a BD) to a given bp/strand combination. Mathematically, clusters with i SBs has a $\tilde{p}(i) = 1 - (\frac{1}{2})^{i-1}$ probability for all SB not to be on the same strand. Therefore, for clusters composed of exactly 2 SB, the cluster has an equal probability of being categorized as an SSB or a DSB cluster. In the case of clusters with 3 (4) SBs, DSB clusters become 3 (7) times more likely than SSB ones. This result is not unique to MCDS. It is also not in agreement with results from MCTS methods. This is further explored in Chapters 4 and 5.

2.4 Combinatorial DNA damage model

Alternatively, radiation-induced DNA damage can be obtained from experimental measurements of ICSDs with an ion counting nanodosimeter or with MCTS simulations using combinatorial mathematics.[32] This modeling approach considers an adjustable parameter representing the probability of an ionization to lead to an SB (p_{sb}). Ionizations clusters are determined either by counting the number of ionizations occurring in the cavity of the nanodosimeter per primary particle traversal or by grouping simulated ED sites using data clustering algorithms, such

as density-based spatial clustering of applications with noise (DBSCAN). Combinatorial mathematics are then used to construct the probabilities of a given cluster of n ionizations to be an SSB or a DSB cluster. A final step is required to transform the probabilities into absolute yields. Several assumptions and approximations are inherent to this formulation. First, it is assumed that the ICSD in a gas system, surrogate to condensed matter, is adequate, because the kinetic energy of the electrons put in motion is not dependant on the phase. Second, p_{sb} is assumed to be independent on the cluster size, which is not adequate for high LET radiation. Effectively, as the ionizations density increases in a given volume, so does the free-radical recombination probability. Furthermore, the combinatorial mathematics for the selection of the type of cluster (SSB or DSB) assume that a SB has an equal probability of occurrence on either DNA strands. This last assumption is challenged in Chapters 4 and 5, where it is shown that due to the DNA conformation and to the distances involved between different strands, consecutive sbs on the same strands are more frequent than those on the opposite strand. These assumptions can be revisited and improved upon; however, even in their current form, combinatorial models based on nanodosimetric measurements are able to predict the LET dependence of clustered DNA damage *in vivo*.[32]

CHAPTER 3
Further Concepts in Radiobiological Modeling

This chapter presents additional background and context on several key concepts that are touched upon in the core of this work in Chapters 4–6. The reader may skip sections he is familiar with without loosing continuity.

3.1 Specific energy

The quantity absorbed dose D is defined as the the average of the stochastic quantity energy imparted $\bar{\epsilon}$ per mass m at a given point of interest, where ϵ is the energy imparted by one or more events in a site of mass m .[79] The special unit of dose is the Gray (Gy) with $1\,\mathrm{Gy} = 1\,\mathrm{J\,kg^{-1}}$. This physical quantity can be measured directly using calorimetry, a measurement of the minute increase of temperature in a known sample of material following irradiation, or it can be inferred indirectly using ionization chambers that count the number of ion pairs (electric current) created in a cavity under the influence of the radiation field. RT treatments are prescribed in terms of an absolute dose in Gy to be given to a point in the tumor or to the tumor volume. Typical curative treatments deliver between 60–80 Gy in 20–40 daily fractions.

The specific energy z is the quotient of ϵ by m.[79] Repeated measurements of z provide an estimate of the probability distribution of z, $f(z)$. The mean specific

energy $\bar{z}$ is the expectation value of the probability density function,

$$\bar{z} = \int_0^\infty z f(z) dz. \tag{3.1}$$

The mean specific energy $\bar{z}$ is therefore the microdosimetric equivalent of absorbed dose. In the case of single-events, i.e., when the energy is imparted by a single particle track, the frequency-mean specific energy per event,

$$\overline{z_{\mathrm{F}}} = \int_0^\infty z f_1(z) dz, \tag{3.2}$$

is used instead. This quantity represents the average energy deposited per mass in single track traversal of a target. It is a useful measure to evaluate the type of radiation present in a beam. The frequency-mean specific energy per event can be measured using microdosimeters, which are gas proportional counters kept at a very low pressure. Adjustment of the pressure can emulate volumes of several tens of nm to a few cm. As the volume of the cavity decreases, the probability density function of single-events $f_1(z)$ enlarges. The probability distribution of z is important because radiation effects are more related to z than D for small values of m such as biological cells.

As will be shown, the stochastic nature of energy deposition will explain differences of biological response. Therefore, dose can be understood as the average of a series of i irritations of a small volume of matter of mass m, in which a total energy imparted of ϵ_i is deposited for each attempt. In contrast, specific energy represent each individual measurement of ϵ_i/m. For completeness, the *macroscopic*

dose is often used to refer to an average dose over a volume (not a point) such as a CT voxel, an organ, or a tumor.

In this work, we simulated the frequency-mean specific energy per event for several electron energies using the GEANT4-DNA code as described in Chapter 6. We use these simulations to estimate the average number of DSB clusters created in the traversal of a cell by a single particle of each radiation type.

3.2 Linear energy transfer (LET)

LET describes the action of radiation in matter. It equals to the energy transferred ϵ_{tr} by an ionising particle to the material it traverses per unit distance l. Therefore, LET depends on the radiation type and on the material traversed. It is dimensionally equal to the stopping power (SP) of the material; SP is the loss of energy of the radiation field, whereas LET represents the absorption of energy by the medium. In theory, LET is only defined for charged particles. In the case of photons, LET in often used in the sense of the LET of the secondary electrons put in motion by the photon beam. There is a non-trivial relationship between LET and RBE, dependent on the biological material, the biological endpoint, the particle type and energy. However, for mammalian cells, the RBE increases as a function of LET up to a maximum around 100–200 $\text{keV}\,\mu\text{m}^{-1}$. It is believed that this peak in efficacy is reached when the clustering of ED sites reflects the average interstrand distance in the DNA molecule (i.e., $\approx 2\,\text{nm}$). At even higher LET, energy is deposited in between the DNA strands, effectively increasing the dose without any increase of the DNA damage. In addition, RBE is generally independent of LET below $10\,\text{keV}\,\mu\text{m}^{-1}$. One of the main results of Chapter 5 is

that RBE_{DSB} increases with increasing LET for protons of kinetic energy between 10 keV μm^{-1} and 30 keV μm^{-1}, and that it is higher for protons than for carbon ions of the same LET (see Figure 5–6), thereby demonstrating the inadequacy of LET for the specification of RBE_{DSB} for these beams.

3.3 Fluence

The ICRU defines the fluence Φ as the quotient of dN by da, where dN is the number of particles incident on a sphere of cross-sectional area da. This quantity is defined at a point; however any measurements or MC simulations require a non-zero scoring volume. Papiez and Battista[80] showed that the ICRU definition of fluence is equivalent to a second, more general, definition based on path lengths. This definition is especially useful when estimating fluence in MC simulations. The fluence can equally well be defined as the quotient of l by V where l is the average length of track segments contained within any shape of volume V. The fluence is often used to characterize a spectrum of radiation. In this sense, it is a function of the energy of the particles present in the field. In Chapter 6, the track-length fluence is calculated for several simulations and compared to the photon interaction spectrum that only considers photons interacting inside the volume.

3.4 Charged particle equilibrium (CPE)

The CPE refers to a state where an equal number of charged particles of any energy enter and exit a given region. It is colloquially subdivided into lateral CPE in the plane perpendicular to the radiation beam and into longitudinal CPE defined in the beam's direction. It is mostly a theoretical concept, that can be achieved with exactitude only under exceptional or theoretical conditions. For instance, in an

MV photon beam irradiating a phantom, the entrance region is where a build-up of electrons occur, in the sense that there is always more electrons leaving than entering. Deeper in the phantom, an approximate state of longitudinal CPE is reached, the transient CPE. In this region, attenuation is responsible for a constant diminution of electrons as a function of depth. However, for very thin volumes attenuation can be ignored and longitudinal CPE is reached. On the other hand, in order to reach lateral CPE, the beam irradiating the region of interest (ROI) must reach passed the ROI by at least the maximal range of the most energetic electrons potentially produced. In the case of ^{60}Co photons irradiating a biological cell, lateral and longitudinal CPE is attained if the cell is not in the build-up region and if it is irradiated by a beam with a radius larger than about 5 mm. CPE is not always reached in RT treatments and it is not a problem in itself. However, as showed and discussed several times in Chapters 4–6, when comparing DNA damage between different radiation qualities, it is primordial to consider similar CPE conditions. In practice, it is sufficient to reach a state of uniform dose throughout the volume of interest such as the cell or cell nucleus. Numerically, CPE can be simulated by using photon regeneration (see for instance an implementation of this technique in Ref 81)

3.5 Associated volume of a track

CHAPTER 4
On the Consistency of Monte Carlo Track Structure DNA Damage Simulations

Simulations have the potential to replace costly, long, and uncertain experimental measurements, while theoretically providing statistically significant and controlled results; however, they are highly dependant on the adjustable parameters, definitions, and algorithms inputed by the user. Therefore, in this work, an MCTS–DNA framework for the simulation of initial DNA damage by electrons was constructed and used to characterize the influence of user-defined adjustable parameters. Particularly, the dose distribution related to the source shape, size, and position, the definition of the simulated SB, and the algorithms used in the implementation of the SB classification into SSB or DSB clusters were investigated, all of which were found to have significant impact on the estimated DNA damage yields. Additionally, we introduced a metric, the nucleosome damage patterns, which is the biophysical equivalent of ICSDs obtained from nanodosimetry experiments. The goal of this work was to raise awareness and discussions on the need for standardization of DNA damage yield simulations using MCTS codes, while shedding some light on some methodological errors in previously published results. Notably, we showed that preservation of CPE is a prerequisite to comparisons of results; therefore, pencil beams through the center of the DNA model are inappropriate types of sources. This knowledge led to the development of a novel

method for simulating the event-by-event electron spectra incident on a cell based on an overlay of pregenerated tracks in Chapter 6.

Authors: Piotr Pater, Jan Seuntjens, Issam El Naqa, and Mario A. Bernal.

Published in: *Medical Physics*, **41** 121708 (2014).

Abstract

Purpose: MCTS simulations have been recognized as useful tools for radiobiological modeling. However, the authors noticed several issues regarding the consistency of reported data. Therefore, in this work, they analyze the impact of various user defined parameters on simulated direct DNA damage yields. In addition, they draw attention to discrepancies in published literature in DNA SB yields and selected methodologies.

Methods: The MCTS code GEANT4-DNA was used to compare radial dose profiles in a nanometer-scale ROI for photon sources of varying sizes and energies. Then, electron tracks of 0.28 keV–220 keV were superimposed on a geometric DNA model composed of 2.7×10^6 nucleosomes, and SBs were simulated according to four definitions based on energy deposits or energy transfers in DNA strand targets compared to a threshold energy E_{th}. The SB frequencies and complexities in nucleosomes as a function of incident electron energies were obtained. SBs were classified into SSB and

glsdsb clusters based on inter-SB distances and on the number of affected strands.

Results: Comparisons of different nonuniform dose distributions lacking CPE may lead to erroneous conclusions regarding the effect of energy on relative

biological effectiveness. The energy transfer-based SB definitions give similar SB yields as the one based on energy deposit when $E_{\text{th}} \approx 10.79\,\text{eV}$, but deviate significantly for higher E_{th} values. Between 30 and 40 nucleosomes/Gy show at least one SB in the ROI. The number of nucleosomes that present a complex damage pattern of more than 2 SBs and the degree of complexity of the damage in these nucleosomes diminish as the incident electron energy increases. DNA damage classification into SSB and DSB clusters is highly dependent on the definitions and their implementations. The authors show that, for the four studied models, different yields are expected by up to 54% for SSB clusters and by up to 32% for DSB clusters, as a function of the incident electrons energy and of the models being compared.

Conclusions: MCTS simulations allow to compare direct DNA damage types and complexities induced by ionizing radiation. However, simulation results depend to a large degree on user-defined parameters, definitions, and algorithms such as: DNA model, dose distribution, SB definition, and the DNA damage clustering algorithm. These interdependencies should be well controlled during the simulations and explicitly reported when comparing results to experiments or calculations.

4.1 Introduction

Ionizing radiation causes SSB and DSB clusters in DNA either through direct interactions with the DNA itself, such as ionizations of the SPB, or through indirect interactions with chemical species produced by water radiolysis, notably $OH^{\bullet}$. A

damaged DNA molecule can potentially lead to lethal consequences for the cell, to mutations, or it can also be flawlessly repaired.

Radiation therapy makes use of the cell-killing ability of radiation to treat cancer patients. The study of the effects induced by ionizing radiation on DNA, through experiments or simulations, is of great interest in the medical physics community where various ionizing radiation types and energies are used to treat cancer patients with advanced techniques. In previous work,[3] we argued that MCTS simulations, validated by adequate experimental data, and implemented in a bottom-up framework linking microdosimetric quantities such as physical interactions with DNA to biological end-points such as cell survival, could be possible[82–84] and would allow for estimating tumor control probability (TCP) and normal tissue complication probability (NTCP), from first principles.

MCTS codes were developed (see review in Ref. 85) to track particles to subionization energies and DNA damage yields can be obtained by superimposing these electron tracks on to a geometric DNA model,[86,87] which could be atomistic[88,89] or by postprocessing these tracks with a probabilistic model.[28,29,54] Some codes also simulate indirect effects that can amount to an important proportion of the damage.[16,90] For instance, Holley and Chatterjee[91] proposed a general theoretical model to simulate direct and indirect DNA damage, without relying on event-by-event tracking of particles. Common MCTS codes such as PARTRAC,[92] PITS,[93] KURBUC,[85] GEANT4-DNA,[24] and PENELOPE[94] rely on somewhat different empirical, semiempirical, and/or experimental models of interaction cross-sections that influence the simulated DNA damage yields. For instance, Li *et*

al.[95] compared DNA damage simulated using six different electron inelastic cross-sections in liquid water and concluded that significant differences in SSB and DSB yields are expected. It is now well documented that electrons of kinetic energies around 100 eV are the major contributors to direct radiation damage of DNA in a cell.[96] Although inelastic cross-sections for DNA molecules were measured and theoretically modeled,[97–102] many MCTS codes still rely on liquid or gaseous water cross-sections. The measurement and theoretical modeling of cross-sections in liquid water and/or DNA is challenging, as high-energy theories gradually fail as energy decreases.[103,104] Zhang and Tan[105] proposed a calculation method to incorporate knowledge of electron cross-sections in base-pair molecules, and show for instance, that guanine–cytosine (GC) base pairs experience more frequent damage than adenine–thymine (AT) pairs.

It has been shown that a step-by-step tracking of low energy electronss (LEEs) may violate the Heisenberg uncertainty principle, which questions the experimental significance of such simulations.[35] Nevertheless, recent work by Liljequist and Nikjoo provide a different perspective on this paradox through the concept of circumstantial validity.[36] They estimate that electrons of 100 eV in liquid water have a relative error in position or in momentum of 1%–2 %, under certain conditions. In the view of these results, and previous experimental validations of MCTS simulations,[106,107] we presume they can be applied to simulate relative DNA damage as a function of irradiation setup and parameters.

We acknowledge that the choice of the MCTS code, with associated cross-sections, impacts the simulated direct SB yields, however user-controlled parameters also have an important role. For instance, users create ROIs, define types and dimensions of sources, implement different classifications and clustering of DNA damage and set limits on the tracking cutoff energy of electrons (E_{cut}) and the threshold energy for the creation of an SB (E_{th}). In this work, we investigate the impacts of these parameters on simulated direct DNA damage yields. Specifically, we analyze the impact of the following parameters: dose distribution dependence, SB definition, choice of E_{th}, nucleosome damage patterns and SB classification into clusters. We also compare simulated SSB, DSB, and complex SSB (SSB+) after irradiation with electron sources of 0.28 keV, 1.5 keV, 5 keV, 10 keV and 220 keV. This work explores the effects of user-defined parameters and their impacts on simulated DNA damage yields for given scenarios. We mainly focus on examples and parameters governing the direct effect of electron irradiation, however conclusions are also valid for mixed direct–indirect simulations and simulations with other particle types.

4.2 Materials and Methods

4.2.1 Track structure simulations

MCTS simulations were carried out using the GEANT4 (4.9.5) simulation toolkit[23] with the GEANT4-DNA processes.[24] The simulation phantom was composed of three enclosed volumes filled with liquid water: the water slab, the low energy process region (LEPR) and the ROI. Simulations are conducted in a liquid

water medium, as DNA material represents a small fraction of the ROI, and cross-section for DNA targets are not so detailed as those used in GEANT4-DNA for liquid water. In recent work by De Vera *et al.*,[108] it was found that despite they accounted for differences of chemical composition and cross-sections, the specific energy deposited by ions and secondary electrons in the cytoplasm and nucleus are practically equal. Nevertheless, the liquid water density was scaled to 1.06 $g\,cm^{-3}$ to approximate the density of a cell nucleus.[25] The ROI and LEPR are centered in the water slab, which is a cube with sides of 1 cm. The LEPR is a virtual volume surrounding the ROI in which the following GEANT4-DNA models and processes, from the G4EMLOW6.23 data file, were active: Champion elastic,[45] Born ionization and excitation (see details in 24), Melton attachment and Sanche excitation. GEANT4 distinguishes itself from other codes notably because all created particles are tracked to zero range and there is no tracking cutoff. Therefore, an additional process (G4eCapture) was used to stop electrons with kinetic energy lower than E_{cut} and deposit locally their energy. A value of $E_{cut} = 10.79\,eV$, equal to the lowest ionization potential of liquid water, was used in this work. Therefore, subionization electrons were not tracked, and their energy deposited locally. These low energy electrons have a residual range that could allow them to reach DNA strand targets and potentially lead to additional, nonsimulated, SBs through resonant effects.[109,110] Outside the LEPR, electrons are tracked using multiple scattering, Livermore ionization, and bremsstrahlung models to preserve accuracy, while reducing the calculation time. Photons are followed using the Livermore models for photoelectric effect, Compton, and Rayleigh scatterings. Fluorescence

and Auger electron deexcitations are active and particles are produced if their energy is higher than 14 eV. All interactions depositing energy in the ROI are saved to a binary file, which is further analyzed with MATLAB (R2012a, The MathWorks, Natick, MA).

4.2.2 Impact of dose distribution

The objective of this section is to demonstrate that the choice of source size and type changes the dose distribution with consequences on simulated DNA damage. A cylindrical ROI of 15 nm radius and half-height of 525 μm, is irradiated by photon sources of 1.5 keV or 1.25 MeV. We used a similar irradiation setup as presented by Bernal and Liendo to emulate their results.[86] Ten simulations of dose distributions binned radially into 100 equal-volume cylindrical shells in the ROI were obtained, each for a total ROI dose of 100 Gy, and averaged. The mean dose per bin is reported as a function of the bin's maximum radius. For 1.5 keV, photons were generated isotropically and uniformly throughout a cylindrical volume of size (a) equal to the ROI or (b) exceeding the ROI by 0.1 μm isotropically, with the LEPR size equal to the source size. For 1.25 MeV, photons were generated in a beam centered on the ROI central axis. The beam was either (c) a pencil beam with the LEPR size equal to the ROI, or (d) a 1 μm-radius circular parallel beam with the LEPR size exceeding that of the ROI by 1 μm isotropically. In irradiation setup (b) CPE is achieved in the ROI, whereas it was not achieved for all other cases.

4.2.3 DNA models and strand break definition

This section discusses different DNA models in the literature by comparing the size of the DNA strand target and contrasts four definitions of the simulated

SB. DNA models consist of spatial arrangements on several organizational levels of volumes corresponding to nucleobases and to theSPB. In Bernal and Liendo,[86] the SPB volumes were modelled as prisms with circular sector base with a volume of 0.24 nm^3. In Bernal *et al.*,[40] the volume was reduced to 0.13 nm^3 to avoid overlapping, while preserving the overall shape. Friedland *et al.*[34,88] model the SPB as the union of spheres centered at the positions of all constituent atoms (1 Phosphor, 5 Oxygen, 7 Hydrogen, and 5 Carbon) with radii corresponding to the atoms van der Waals radii (respectively, 0.19, 0.14, 0.12, 0.17 nm). We calculated the volume of the union of these spheres and obtained 0.13 nm^3. In addition, 60 % of interactions occurring within the water shell (union of spheres of 0.35 nm radius centered on each atoms center) also contributed to the creation of SBs. We calculated that the total effective volume of the direct DNA strand target in their model is close to 0.36 nm^3. Charlton *et al.*[111] and Nikjoo *et al.*[13] simulated DNA strand targets as half cylindrical shells of volume 1.73 nm^3 inside small cylinders. In their probabilistic model, Francis *et al.*[54] used an adjustable empirical parameter to sample interactions located in DNA strand targets. On the other hand, in the model presented by Semenenko and Stewart,[28,29] DNA strand target volumes are not simulated, and only genomic distances are distributed.

In addition to differences in size and geometry of the DNA strand target and the whole geometric DNA model, the SB definition varies also from one study to another. One definition (Sum E_{dep}) requires that the sum of all energy deposits in a DNA strand target exceeds a threshold energy E_{th} in order to create an SB. To define the SB, various authors used E_{th} of 10 eV to 2000 eV,[87] 10.79 eV[25,52] and 17.5 eV.[111]

A ramp probability function (linearly increasing from 0 to 1 between 5 eV and 37.5 eV[62] or 40 eV[88]) was also proposed. Francis *et al.*[54] used this approach to look at maximal energy deposits in DNA strand targets. Alternative definitions include: (Max E_{dep}) requiring that the maximal energy deposited in a DNA strand target exceeds E_{th} (Sum E_{trans}) requiring that the sum of all energy transfers exceeds E_{th} and (Max E_{trans}) requiring that the maximal energy transfer exceeds E_{th}.[40,86]

The DNA model we used is described in details by Bernal *et al.*[40,86] It consists of 2.7×10^6 independent DNA nucleosomes, arranged in groups of 6 around a 30 nm diameter circle and stacked 500 times to form the 30 nm chromatin fiber. The ROI is filled with 900 copies of the 30 nm chromatin fiber, totalling 524.6×10^6 bps, arranged around a cylindrical shell of 2625 nm half-height, with inner and outer radii of 4984.76 nm and 5015.24 nm, respectively. Each nucleosome consists of a double stranded two-turn structure of 198 bps. Each SPB volume is equal to $0.1344\,\mathrm{nm}^3$ and each nucleosome turn has a diameter of 10.5 nm and a thickness of 2.37 nm. The ROI was irradiated with monoenergetic electrons generated isotropically and uniformly throughout a cylindrical volume with radial and half-height dimensions exceeding those of the ROI by $0.4\,\mu\mathrm{m}$ for 0.28 keV, $0.6\,\mu\mathrm{m}$ for 1.5 keV, $1\,\mu\mathrm{m}$ for 5 keV and $2.5\,\mu\mathrm{m}$ for both 10 keV and 220 keV electrons sources. The LEPR volume was equal to the source volume for each energy. The source size was made large enough to achieve CPE inside the ROI except for the 220 keV source. For this case, a source exceeding the ROI by 0.5 mm would be necessary to achieve CPE. Such a source is larger by an order of 10^5 than the one we used and would be computationally too expensive to simulate. The full set

of inelastic interactions inside the ROI is processed by an in-house algorithm to rapidly determine which interactions happened inside the DNA strand targets. Our algorithm also identifies SBs according to the four strand break definitions (Sum E_{dep}, Max E_{dep}, Sum E_{trans} and Max E_{trans}) and was used to compare the TSB yield for all simulations. In six consecutive and similar steps, our algorithm finds all interactions occurring within six subunits: (1) a 30 nm chromatin fiber, (2) a group of 6 nucleosomes, (3) a nucleosome, (4) a nucleosome turn, (5) a bp and (6) a DNA strand target. In each step, a state is associated with every inelastic interaction at once. For instance, in step (1), interactions will be associated with state '0' if they are outside of all 30 nm chromatin fibers or to a state between '1' and '900' corresponding to the position of the 30 nm chromatin fiber that was hit. Once a state is obtained, we use symmetry considerations to move all interactions within a representative structure for the given subunit. This is repeated for each subunit and the set of six states for each interaction determines if and which DNA strand target will be hit.

4.2.4 Nucleosome damage patterns

In this part, we estimated the frequency and the complexity of damage in nucleosomes. Each nucleosome is a circular double-stranded DNA fragment of 198 bps. These nucleosome damage patterns give an estimate of the complexity of the damage created and are independent of SB clustering algorithms. We used the SBs determined with Sum E_{dep} and E_{th} = 10.79 eV. Our algorithm determines how many nucleosomes per unit dose show at least one, exactly one, two, three, four, or five SBs or more as a function of the incident electron energy.

4.2.5 Differences in strand break clustering algorithms

We implemented three algorithms from the literature to group SBs into clusters in addition to one other algorithm proposed by us. Typically, SB distributions are clustered using inter-SB distances and the knowledge of the strand that was hit. Charlton and Humm[112] definitions (SSB, SSB+, 2 SSB, DSB, DSB+, DSB++) of clustered damage are commonly used, notably the DSB is defined as a pair of SBs on opposite strands located within 10 bps. These definitions are adapted and implemented differently depending on the authors and the DNA models adopted. Fig. 4–1 compares the expected clustered DNA damage as obtained by three different algorithms from the literature (Bernal,[40,86] Friedland[25] and Charlton and Nikjoo[87,111]) and our own proposed clustering algorithm. Bernal's algorithm looks for pairs of SBs in a unidirectional manner, and categorizes each SB in more than one complex type of damage. This results in scoring 2 DSBs if 3 alternating SBs are found within 10 bp. The algorithm also scores a SSB+ when a pair of SBs within 10 bps on the same strand is found. Also, all SBs not related to DSB clusters are counted as SSB clusters, including those only leading to SSB+ clusters. On the other hand, Friedland's algorithm starts by scoring all DSB clusters with a preference given to the closest pair of SBs. Then, all remaining SBs are classified as SSB clusters. The case when two directly adjacent SBs on the same strand are found is exceptionally counted as a SSBs composed of a single SB. Using Charlton's and Nikjoo's definitions, the example damage pattern in Fig. 4–1 is simply classified as a DSB++, a type of damage where both strands are hit twice or more, creating a short unbound DNA fragment, that can potentially lead to a

deletion after incomplete repair. To compare their classification created for linear DNA fragments, we defined the SSB_{NIKJOO} = (SSB)+2∗(2 SSB)+2∗(SSB+)+(DSB+), to count the total number of SSBs clusters, i.e 1 for SSB, 2 for 2SSB, 2 for SSB+ and 1 for DSB+, which is composed of 1 SSB and 1 DSB. We also defined the DSB_{NIKJOO} = (DSB) + (DSB+) + 2 ∗ (DSB + +), to count the total number of DSB clusters, i.e. 1 for DSB, 1 for DSB+ and 2 for DSB++, which could potentially be composed of more than 2 DSBs.

In addition, a modified algorithm is proposed that finds the same DSB yields as the Friedland's algorithm, and the same SSB yields as Bernal's one. Our approach, also shown in Fig. 4–1, starts by looking for the closest pairs of opposite strand SBs and categorizes each as a DSB. When no more DSBs are present, the algorithm looks for pairs of SBs on the same strand and categorizes them as SSB+. These pairs have to be within 10 bps to be counted in these complex damage types. In addition, all SBs unrelated to DSBs are counted as an SSB and the sum of all individual SBs gives the TSB. Using the Sum E_{dep} SB definition (see section 4.2.3), we compare the percent difference of SSB and DSB yields between all these classification rules for our set of data. In addition, we report the TSB, SSB, SSB+ and DSB yields according to our own clustering conditions as average values over 10 batches. The total dose in each simulation was 1000 Gy, except for 10 keV and 220 keV where only 100 Gy were delivered to reduce simulation time.

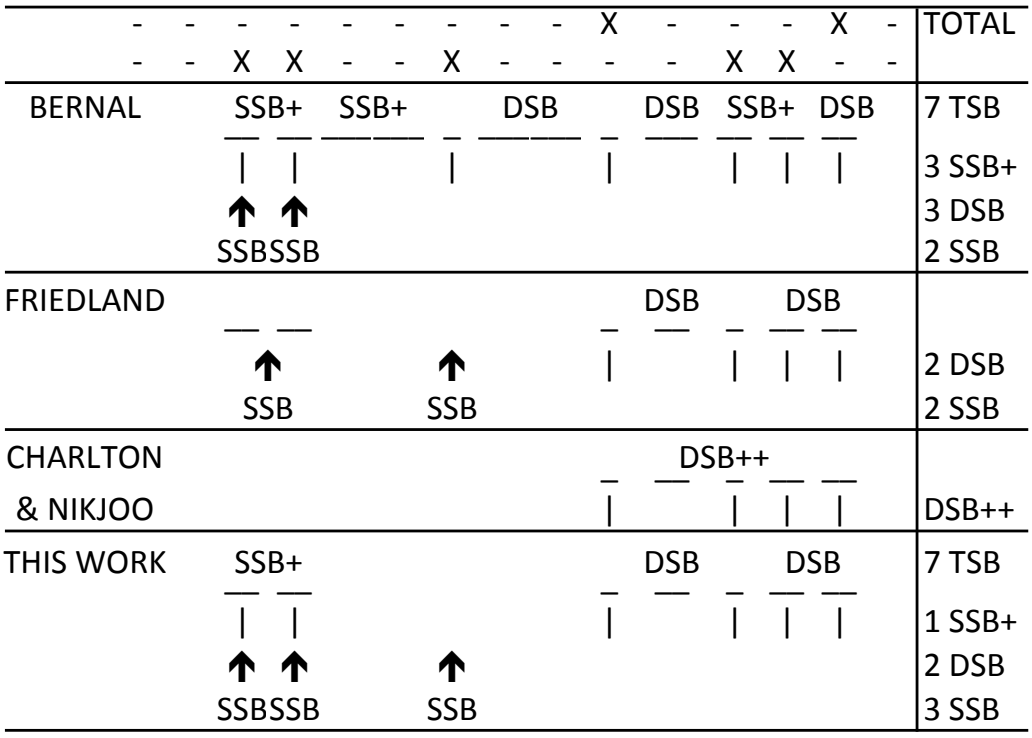

Figure 4–1: Comparison of various SSB, DSB, SSB+ and TSB clustering conditions, including the ones from Bernal and Liendo,[86] Friedland *et al.*,[25] Charlton and Nikjoo[87,111,112] and our own implementation. The top part represents two DNA strands with x representing a strand break and - representing an unaffected DNA strand target. The DSB++ can be understood as a double DSB where each strand is hit at least two times within 10 bps.

4.3 Results

4.3.1 Impact of dose distribution

The mean dose in the cylindrical shells irradiated with photon sources described in section 4.2.2 as a function of the outer radius of the shell is given in

Fig. 4–2. In all cases, the total dose to the ROI was 100 Gy, but the radial dose distributions differ. For 1.5 keV photons, a source that produces CPE in the ROI (●) yields a constant dose of 100 Gy in all shells, within the statistical uncertainty. In contrast, when CPE is not achieved (○), a non-uniform radial dose distribution is obtained, with a peak dose of 160 Gy in the center and a low dose of 38 Gy around the edge of the ROI. For 1.25 MeV photons, none of the presented simulations yields an uniform dose in the ROI. The plane beam with 1 μm radius (■) produces a dose of 130 Gy near the center and of 60 Gy near the ROI boundary. A pencil beam of 1.25 MeV photons (□) produces a dose distribution that is heavily peaked around the center of the ROI. At this energy, a 1 cm-diameter beam should be simulated in order to achieve CPE within the ROI.

4.3.2 Impact of the strand break definition

The TSB yield is dependent on both the definition of the SB and on the value of E_{th}. Fig. 4–3 compares the simulated TSB yield as a function of E_{th} for the Sum E_{dep}, Max E_{dep} and Max E_{trans} SB definitions. The Sum E_{trans} definition is not shown but follows a similar trend as Max E_{trans}. For all cases, the TSB yields decrease monotonically as E_{th} increases. The Max E_{dep} curves follow step-like functions related to the different ionization and excitation potentials of liquid water.[95] The Sum E_{dep} curves decrease continuously and no significant dependence on the incident electron energy is observed. For both sets of curves based on energy deposited, virtually no SBs are created for $E_{\text{th}} \geq 100\,\text{eV}$. Both set of curves based on energy transfer exhibit an energy dependence, i.e., for a fixed E_{th} value, more SBs are produced for higher incident energies. An interesting point is that if

$E_{th} \leq 10.79\,eV$, energy transferred and energy deposited are now equal for all interactions. Finally, definitions based on maximum and total energies differ only when more than one interaction occurs inside a given DNA strand target.

4.3.3 Nucleosome damage patterns

Fig. 4–4 shows that only 30.6 nucleosomes at 0.28 keV and 44.2 nucleosomes at 220 keV presented at least one SB per Gy (×) out of 2.7×10^6 nucleosomes in total. Also, the number of nucleosomes hit exactly once (●) increases with incident electron energy, whereas the number of nucleosomes presenting complex damage patterns (2 (■), 3 (◆), 4 (□) and 5+ (○) SBs) decreases with increasing incident electron energy (i.e., decreasing linear energy transfer). Finally, the fraction of highly complex damage (5+ SBs) over the sum of all complex damage diminishes as the incident energy increases. This last observation entails that both the frequency and the complexity of the damage is higher for incident electrons of lower energy.

4.3.4 Differences in strand break clustering algorithms

Fig. 4–5 presents the percent differences of SSB and DSB yields obtained when using different SB clustering algorithms from the literature compared to our proposed algorithm, using the same initial nucleosome damage patterns, based on the Sum E_{dep} SB definition with $E_{th} = 10.79\,eV$. As expected, our method finds the same SSB yields as in Bernal's and the same DSB yields as in Friedland's. However, our classification detects between 6 % and 16 % more SSBs than Friedland's, and between 8 % and 12 % less DSBs than Bernal's as a function of the incident electron energy. Charlton and Nikjoo's algorithm classifies damage into 5 different

categories, and for purposes of comparison with this work, we defined SSB_{NIKJOO} and DSB_{NIKJOO} in section 4.2.5. Using these definitions, our classification method finds between 12 % and 24 % more DSBs than using DSB_{NIKJOO} and between 18 % and 48 % less SSBs than using SSB_{NIKJOO} as a function of energy.

Table 4–1 presents TSB, SSB, SSB++ and DSB yields obtained with our proposed SB clustering algorithm and Fig. 4–6 compares these values (●, ○) to SB yields published by Bernal and Liendo[113] (■, □) and Friedland *et al.*[25] (◆, ◇).

Table 4–1: Direct DNA strand break yields in $Gy^{-1}\,Gbp^{-1}$, including TSB, SSB, DSB, and SSB+ obtained using our classification presented in Section 4.2.5 for selected incident electron energies between 0.28 keV and 220 keV.

	0.28 keV	1.5 keV	5 keV	10 keV	220 keV
TSB	93.8 ± 1.9	97.7 ± 1.4	99.7 ± 1.4	99.5 ± 3.7	97.4 ± 6.0
SSB	76.6 ± 1.4	85.1 ± 1.7	88.5 ± 1.4	90.0 ± 2.9	89.4 ± 5.3
DSB	8.1 ± 0.5	6.1 ± 0.3	5.4 ± 0.2	4.6 ± 0.8	3.8 ± 1.0
SSB+	9.7 ± 0.4	7.5 ± 0.3	6.4 ± 0.3	5.9 ± 0.9	5.0 ± 0.6

4.4 Discussion

Table 4–2 presents a summary of the investigated parameters and their impact on simulated direct DNA damage from irradiation with electrons sources. Calculations should be seen valid from the relative instead of the absolute point of view. The conclusions presented are valid for irradiations with other particles such as ions or photons, however the absolute numerical differences may differ. Sections 4.4.1–4.4.4 discuss these issues in depth.

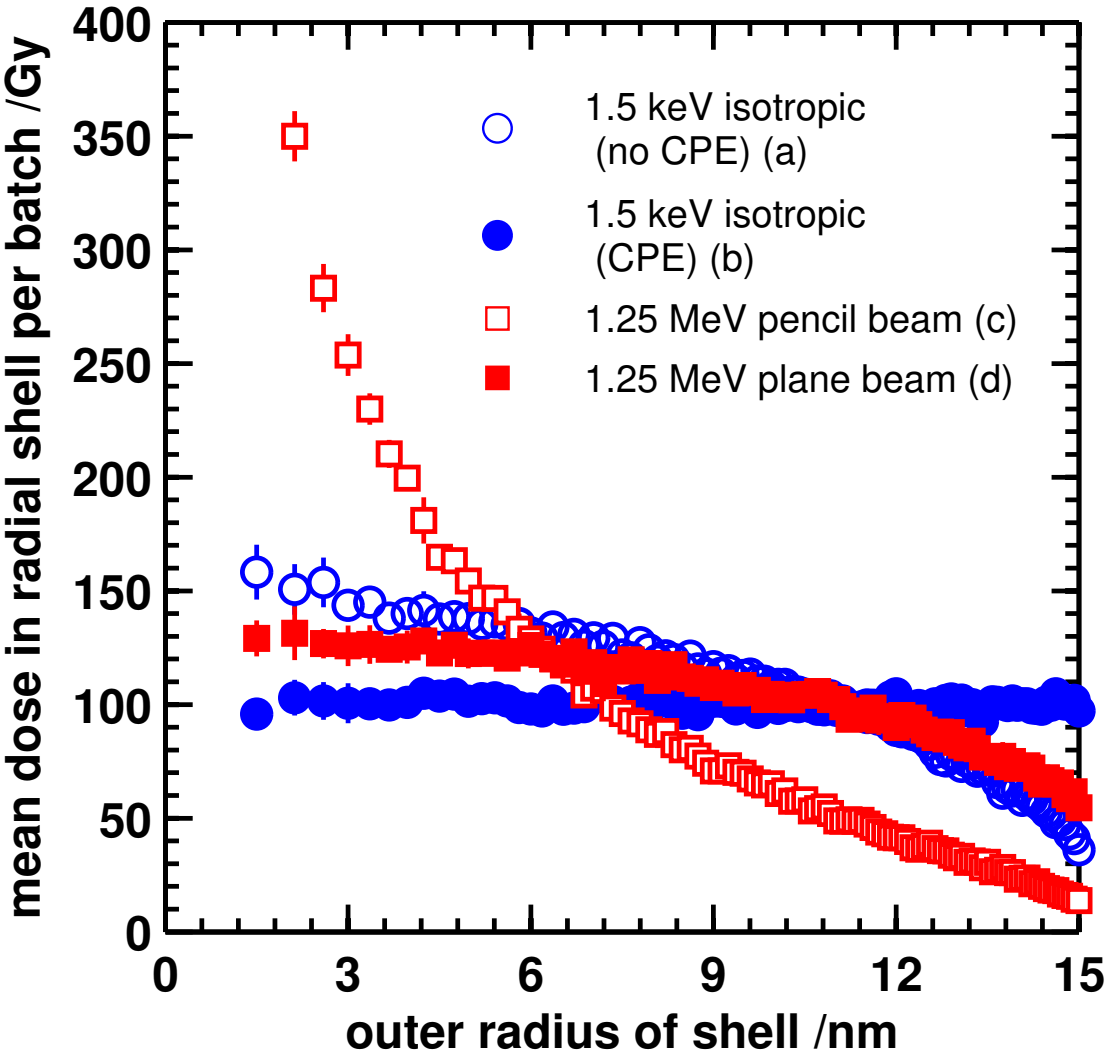

Figure 4–2: Radial dose distribution in the 15 nm-radius cylindrical ROI after the irradiation with isotropic sources of 1.5 keV photons (○, ●) or plane beams sources of 1.25 MeV photons (□, ■) with the following dimensions : (a) ○ source size = LEPR size = ROI size , (b) ● source size = LEPR size = ROI size + 0.6 μm in all directions, (c) □ pencil beam with LEPR size = ROI size (d) ■ plane beam (radius = 1 μm) with LEPR size = ROI size + 1 μm in all directions. CPE in the ROI is attained only for ●, the source (b). All shells have equal volume. Error bars correspond to two times the standard deviation of the mean.

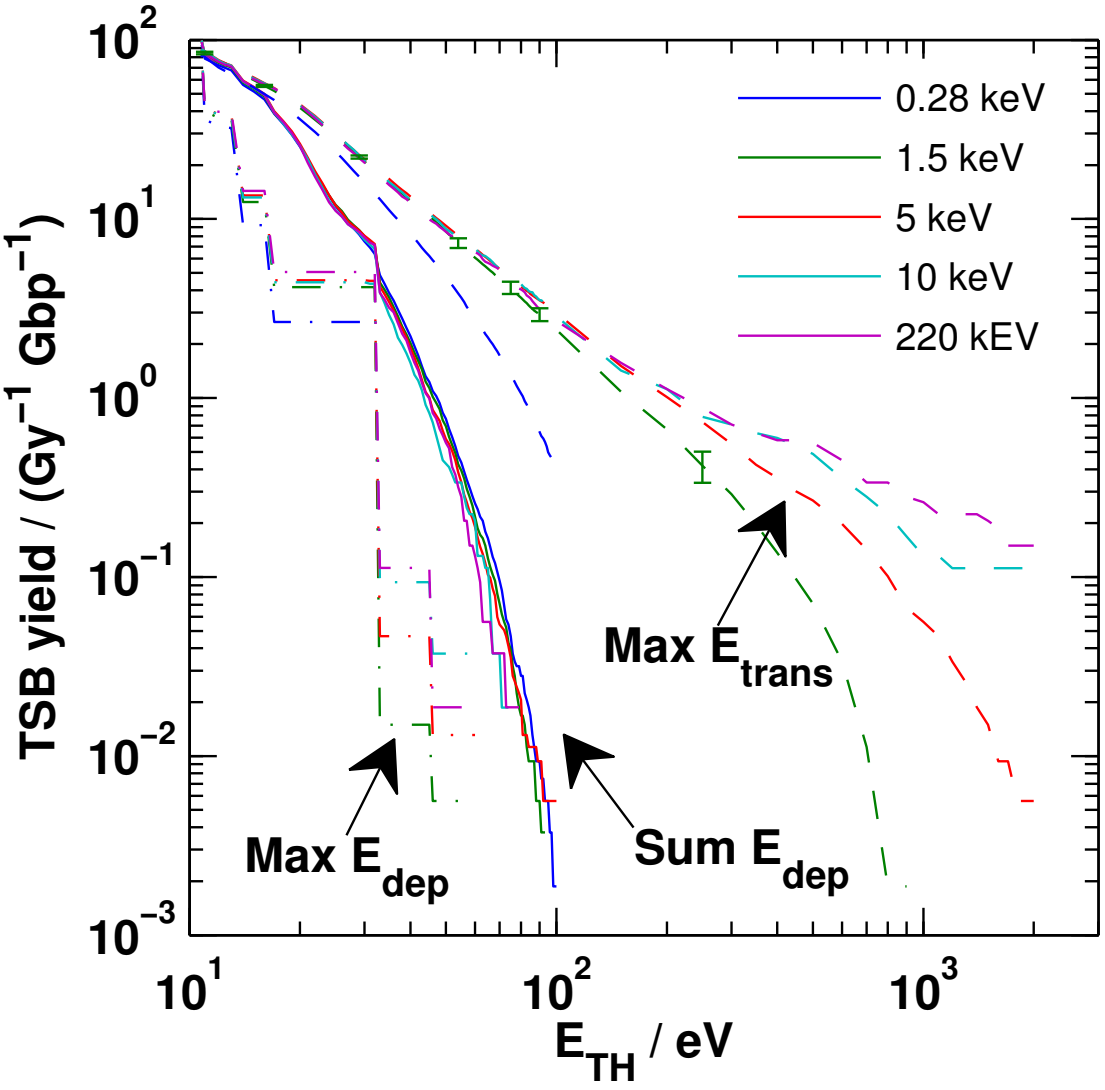

Figure 4–3: TSB as a function of E_{th} for simulations with 0.28 keV, 1.5 keV, 5 keV, 10 keV and 220 keV electron sources. Dashed-dotted curves represent the SB definition Max E_{dep}, solid curves represent Sum E_{dep} and dashed curves represent Max E_{trans} (see Section 4.2.3). Curves based on Sum E_{trans} were omitted for clarity but mainly follows similar trends as Max E_{trans}. Representative error bars are shown on one curve (Max E_{trans} at 1.5 keV) and equal to two times the standard deviation of the mean over 10 simulations.

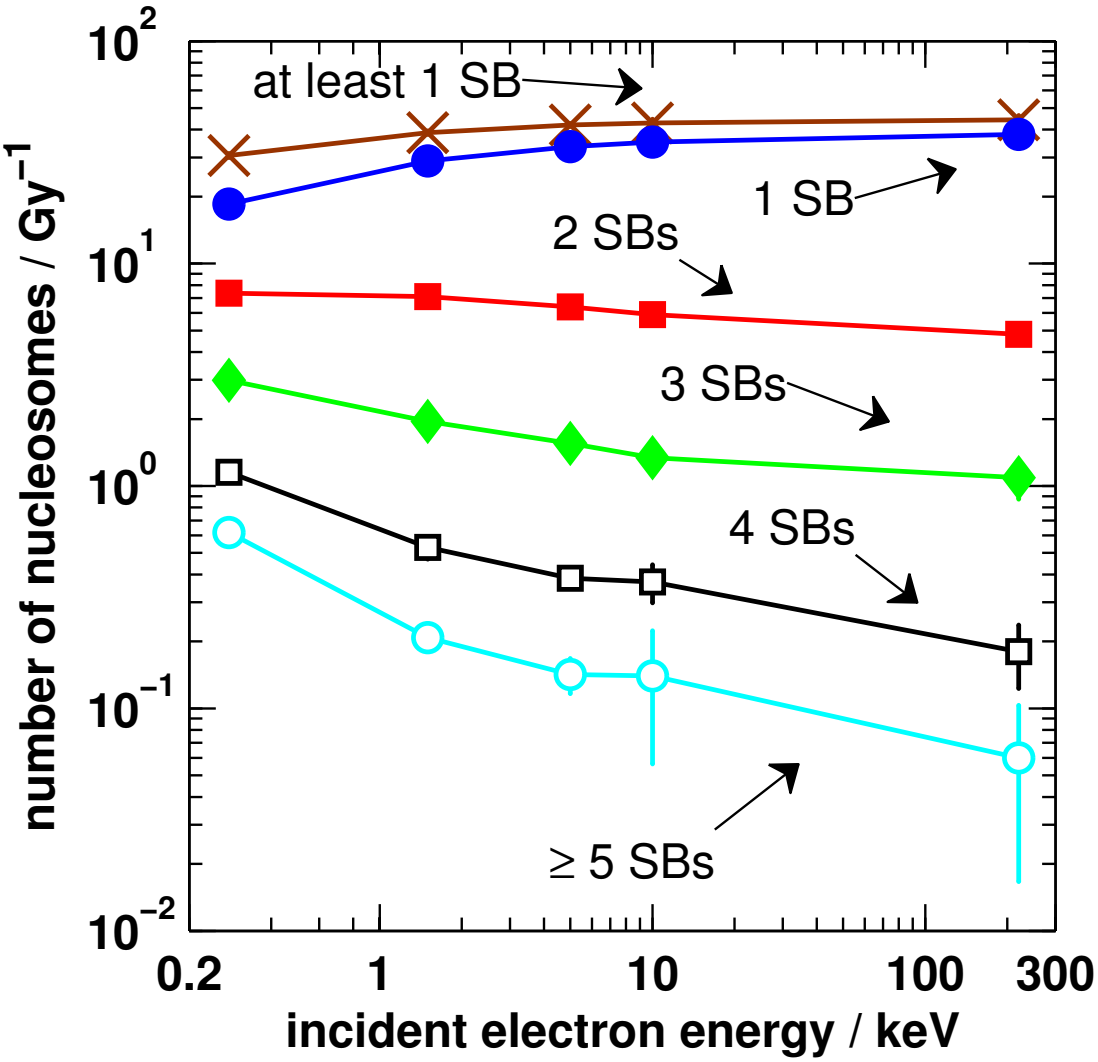

Figure 4–4: Number of nucleosomes/Gy presenting at least 1 (×), exactly 1 (●), 2 (■), 3 (♦), 4 (□) and 5 and more (○) SBs as a function of the incident electron energy. Error bars correspond to two times the standard deviation of the mean for 10 batches.

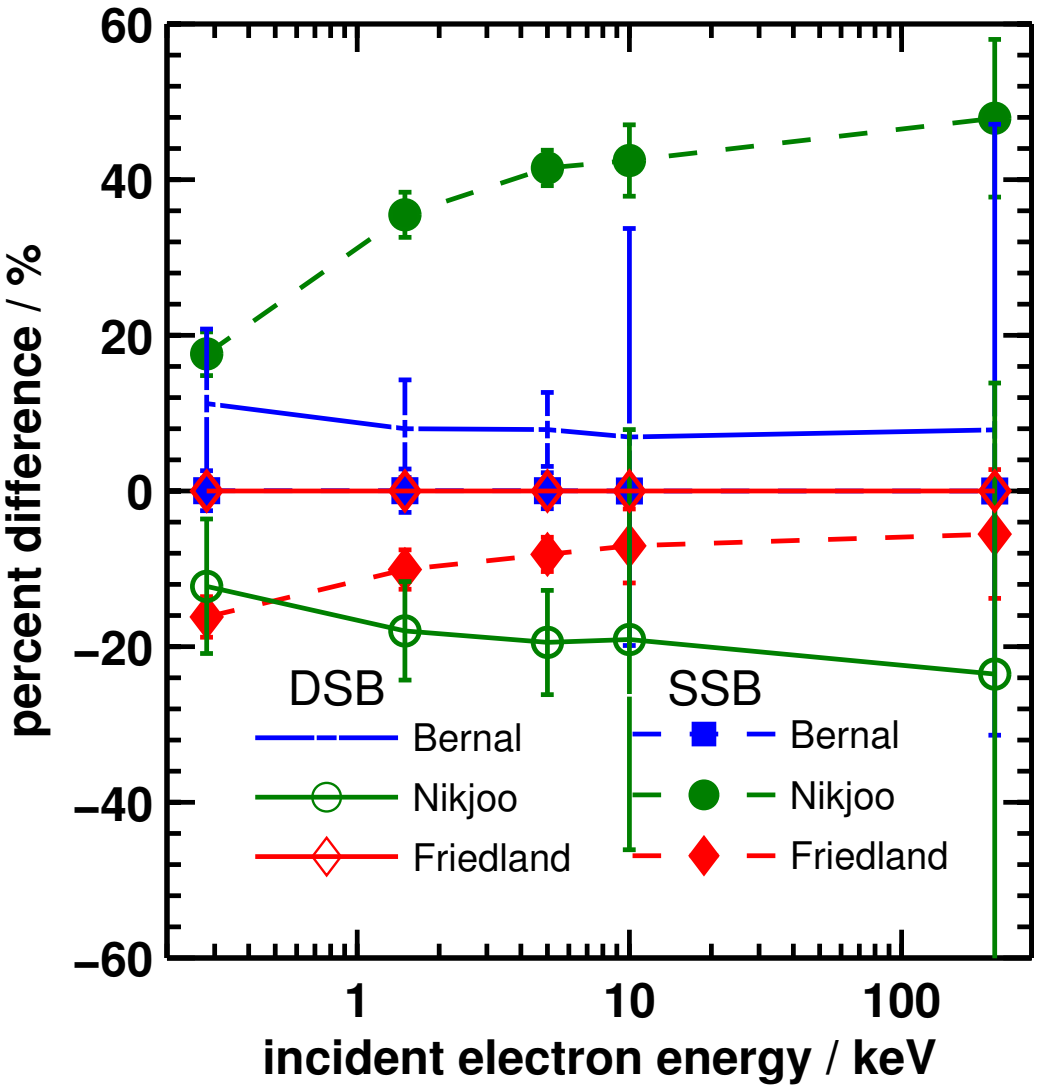

Figure 4–5: Percent difference of SSB and DSB yields obtained with strand break clustering algorithms from Bernal (Refs. 86 and 40), Friedland *et al.* (Ref. 25) and Charlton and Nikjoo (Refs. 111 and 87) as compared to our own algorithm. It is defined as % diff = 100% ∗ (value – this work)/this work. Error bars represent one standard deviation of the sample.

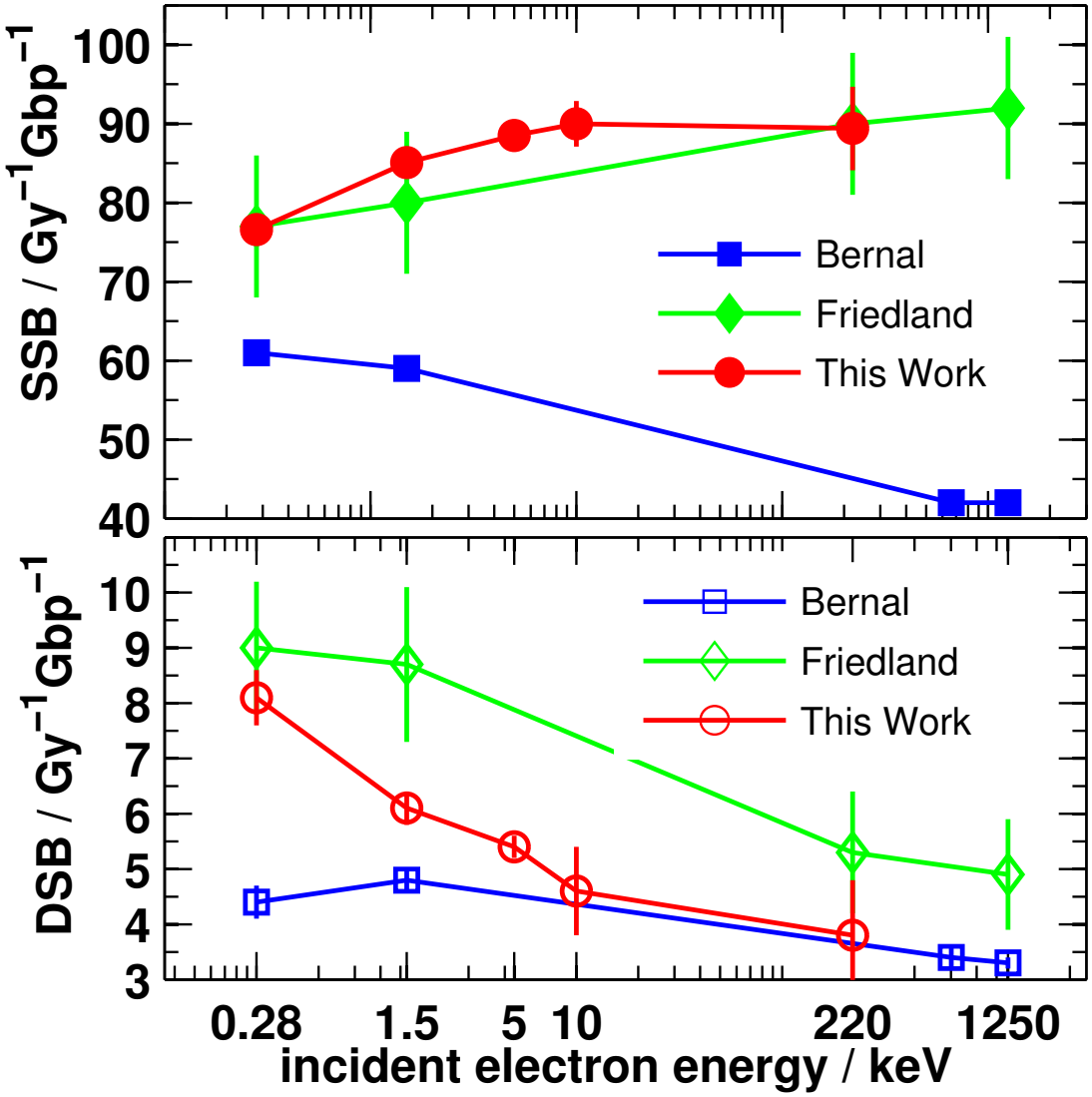

Figure 4–6: Simulated direct SSB (●, ♦, ■) and DSB (○, ◇, □) yields comparisons between this work (●, ○), Bernal and Liendo (Ref. 113) (■, □) and Friedland *et al.* (Ref. 25) (♦, ◇). Error bars for 'This Work' equal one standard deviation of the sample, whereas literature values are shown with one standard deviation (unspecified).

Table 4–2: Impacts of parameter choices in MCTS simulations on DNA strand break yields.

Parameter choices	Impact on DNA strand break yields
Dose distribution in the ROI	Achieving CPE in the ROI, guarantees an uniform dose distribution. This can be achieved by using a source with dimensions that exceeds the ROI by at least the range of the highest energy electrons. Similar, but nonuniform dose distributions can also be compared relatively. However attempting to compare results from different dose distributions in the ROI can lead to erroneous conclusions on the causes of DNA damage differences (see section 4.3.1).
Cutoff energy E_{cut}	This parameter is limited by the MCTS cross-sections. It should be as low as possible to avoid underestimation of the total number of ionizations and thus of SBs. Higher values of E_{cut} can potentially lead to lower incidences of complex damage patterns. In addition, electrons with kinetic energy below E_{cut}, may also contribute to DNA damage through resonant effects that may or may not be simulated (see Section 4.2.1).
ROI dimensions, genome size and dose	Larger ROI requires longer simulations, and larger sources to achieve CPE, but have no influence on yields. Increasing genome size for fixed ROI dimensions or increasing the dose will decrease the statistical uncertainty on the yields.
DNA strand target size and positions	The DNA strand target volume affects the number of simulated SBs as multiple interactions in the same DNA target are more probable for larger target sizes (Ref. 40). Relative distances between neighbouring DNA strand targets can affect clustered yields such as DSBs.
SB definition and E_{th}	SBs defined based on Sum E_{dep} or Max E_{trans} are both used in the literature and produce decreasing TSB yields as a function of E_{th} (see Fig. 4–3). These definitions are equivalent for E_{th} = 10.79 eV. For higher values of E_{th}, energy-transfer based definitions overestimate TSB yields compared to energy deposit- based definitions. In addition, TSB yields defined with energy transfer vary as a function of the incident electron energy.
DNA damage classification	Implementations of DSB and SSB classifications differ among authors. This can result in differences of up to 54 % for SSB yields and of up to 32 % for DSB yields, as a function of the incident electrons energy and of the models compared (see Fig. 4–5). Comparison of DNA damage yields between different authors, or with experimental values should account for this effect.

4.4.1 Impact of dose distribution

In order to obtain full CPE in the ROI, the dimensions of a uniform isotropic source must be larger than the ROI by the maximal secondary electron range, and electrons need to be tracked with track structure processes (i.e., down to subionization or subexcitation energies) throughout the source volume. At incident energies of the order of the kilo-electron-Volt, this is computationally achievable, but as the maximum range of electrons increases, this becomes challenging. For instance, looking at the 1.5 keV simulations of Fig. 4–2, when CPE is achieved, doses in all radial bins are equal, and thus there is an equal interaction density in all bins. When nonuniform interaction densities are superimposed on nonuniform DNA densities in the ROI, this can lead to erroneous conclusions. In the work of Bernal and Liendo,[86] no DNA strand targets were present in the central 4 nm-radius of the ROI. They used a pencil beam for the 1.25 MeV radiation to simulate SBs, which created a strongly peaked radial dose distribution around the center, where dose was deposited, but no SBs created and thus underestimated SSB yields (see Fig. 4–6). In addition, they compared results by calculating the RBE_{DSB} as a function of energy, overlooking the nonuniform interaction densities that contributed to these differences. If a beam is used instead of a volumetric source, the depth-attenuation in the microscopic ROI can be neglected for higher electron energies, but is significant for low energy. This effect was observed by Friedland *et al.*[25] when homogeneous irradiation yielded an increase of about 10 % of the SSB and DSB yields when compared to irradiations using a beam. They used a source of the same size for all simulations, not accounting for lateral CPE

which produced slightly different dose distributions in the ROI at each energy. The spectrum of secondary electrons incident on the ROI is highly dependent on the irradiation setup and blinded comparison based on nominal beam energy can lead to misrepresentation of DNA damage.

4.4.2 Impact of the strand break definition

The SB definition as well as the choice of E_{th} impacts the TSB yield as shown in Fig. 4–3. At all values of $E_{\text{th}} > 10.79\,\text{eV}$, energy transfer-based definitions yield more SBs than energy deposit-based definitions. This is in accordance with the definition of these energies. The actual value of the E_{th} parameter is still unknown and it can be seen as a fitting parameter. The user can chose an E_{th} value that predicts a given experimental or simulated yield of direct strand breaks, as was previously done in other work (see references in section 4.2.3). One could argue that energy transfer-based definitions are a closer approximation of the amount of energy available to create an SB in a DNA molecule. Most simulations are based on interactions in liquid water, and energy depositions are calculated based on the ionization potentials of liquid water, not DNA molecule constituents. However, as E_{th} increases, energy transfer-based definitions exhibit a dependence on the initial spectrum. In other words, such definitions do not conserve the invariance of the TSB yield as a function of the incident energy which is not an expected behavior.[40] It is worth noting that for E_{th} values close to 10.79 eV, TSB yields are less dependant on the definition.

4.4.3 Nucleosome damage patterns

We chose to use the Sum E_{dep} with $E_{th} = 10.79\,eV$ definition of an SB as this corresponds to the lowest ionization potential of liquid water, and thus all simulated ionizations in SPB volumes produce an SB. We did not attempt to reproduce an experimental data set and thus our results can be seen as relatively valid with respect to each other. Although DNA SBs from direct effects of radiation have a low probability of occurring, they are still likely to induce biological effects. As seen on Fig. 4–4, at all energies, less than 8 nucleosomes/Gy exhibit more than two SBs, becoming even less probable as the incident electron energy is increased. In addition, the complexity of the damage is increased with lower incident electron energies. These results predict that complex DNA damage is more frequent and more complex as the incident electron energy decreases. Nucleosome damage patterns can also be seen as strand break frequencies in DNA fragments of around 200 bps. We believe that nucleosome damage patterns, which depend only on the SB definition, on the value of E_{th}, and on the geometrical DNA model could be good candidates for consistency checks between different MCTS codes. They do not suffer from the added complexity of classification of SBs into SSB and DSB clusters, yet they predict similar trends.

4.4.4 Differences in strand break clustering algorithms

Fig. 4–4 gives only the total number of SBs in a nucleosome, but provides no information on the SB distributions in the nucleosome. For example, it is well known that two isolated SBs are easier to repair than two clustered SBs on opposite strands.[114] In addition, experimental results are often presented as SSB and DSB

yields, defined as either a change in the conformation of the molecule or as the creation of fragments of the DNA molecule. Authors mostly agree on the definition of the DSB, regarded as a pair of SBs on both strands within 10 bps, but implementations of this definition differ and can lead to discrepancies as shown by our results (see Fig. 4–5). These differences become more apparent for lower energy electrons, where the damage complexity increases. From Fig. 4–5, differences in DNA damage clustering of as much as 32 % in DSB yields and of up to 54 % for SSB yields are expected as a function of the clustering implementations and electron energy. In published data, these relative differences in the clustering algorithms can be masked by differences in MCTS codes, cross-sections, or other parameter choices. In addition, when simulation data are validated against experiments, these differences could be not significant. Nevertheless, we believe that standardizing the reporting of the clustering methods and possibly the methods themselves, with adequate experimental validations, would allow easier interpretation, intercomparisons and fewer floating-parameters in simulations that may increase the chance of over-fitting the experimental data. Our simulation data shown in Table 4–1 confirms that the yields of complex damage (i.e., DSB and SSB+ clusters) decrease with increasing energy. This is concordant with our nucleosome damage results that showed a decrease in multiple hits in the same nucleosome and a decrease of the complexity of the damage as the incident electron energy is increased. The TSB yields are invariant with energy except for 0.28 keV, where multiple ionizations in the same DNA strand target become more probable, reducing the TSB yield. This invariance has already been studied in detail for ion beams, using the

site-hit probability.[40] For incident electrons, our results show that the SSB yields increase with energy. This is an indirect effect of the TSB yield invariance with incident energy and the DSB yield decrease with incident energy. Finally, for all energies, SSB+ clusters are slightly more probable than DSB clusters. This last result is interesting and will be studied further in a future study. We believe it is linked to the fact that the mean distance between two targets on the same strands that can create a SSB+ cluster is shorter than the one between two targets on opposite strands in the DNA model we used. In other words, a second SB (within 10 bps) is more probable to occur in a DNA strand target on the same strand than on the opposite strand. This result contradicts models that use equal probability to hit either strand.[28,29,54] Our simulations show an overall good agreement (Fig. 4–6) with Friedland *et al.* results. A slight overestimation of SSB yields in our work could be explained by the fact that Friedland *et al.* DNA damage scoring algorithm counts as one SSB two adjacent SBs, whereas they are counted as 2 SSBs (and an SSB+) in our work. Our work underestimates DSB yields compared to those reported by Friedland *et al.* and we hypothesize that this is the consequence of different sizes of the DNA strand targets and of nonuniform dose distributions used in their work. Bernal and Liendo's results overall underestimate SSB and DSB yields and show a decreasing SSB yield with energy, in contradiction to the expected behaviour. Their results are a direct effect of the use of a pencil beam at high energies and of a small volumetric source for low energies, that creates inhomogeneous dose distributions in the ROI, with possibly high dose regions created outside of DNA strand target locations. Finally, we remind the reader that

this study did not compare differences in MCTS codes, cross-sections or models, nor did it compare simulations to experimental results, but it was rather focused on end-user controlled parameters and their impacts on end results.

4.5 Conclusion

The main focus of this study was to compare choices of various user-defined MCTS simulation parameters and understand their impact on calculated direct DNA damage yields. Our findings suggest that significant differences arise from subtle modifications of definitions, algorithms, or parameter values, which may or may not impact the validation by experimental data. More specifically, we showed that achieving a uniform dose is ideal for multiple energy studies, and it can be obtained by accounting for the range of the secondary electrons created when defining the incident particle source size, or achieving similar dose distributions for all energies. We compared the impact of the SB definition on TSB yields. We also showed that differences of up to 54 % for SSB yields and of up to 32 % for DSB yields, can result from slight variations in SB clustering algorithms implemented in the literature. This paper shows quantitatively why we need a forum for "standardization" of MCTS simulation parameters.

4.6 Acknowledgments

We would like to acknowledge support from the Fonds de recherche du Québec - Santé, support from the CREATE Medical Physics Research Training

Network grant of NSERC (Grant number: 432290) and support from CIHR-MOP-114910. M Bernal thanks the FAPESP foundation in Brazil for financing his research activities through the 2011/51594-2 project. We would also like to thank the anonymous referees for their valuable insight.

CHAPTER 5
Proton and Light Ion RBE for the Induction of Direct DNA Double Strand Breaks

In Chapter 4, it was shown that user-defined parameters and algorithms have a significant effect on the DNA damage yields obtained with MCTS simulations. Notably, it was found that uniform dose throughout the volume of interest is a pre-requisite for comparisons of DNA damage yields at multiple energies. However, these computation were costly in terms of time for higher energy electron beams. Therefore, in this chapter, a novel algorithm to overlay pregenerated particle tracks on a cell nucleus is presented. The methodology is applied to the study of DNA damage produced by low energy protons and light ions. The DNA model from Chapter 4 was updated to represent a micron-sized cell nucleus. Tracks of protons and other light ions were pregenerated and overlaid on a geometrical DNA model consisting of more than 6×10^9 bps. In this work, the electron tracking cutoff was lowered to 9 eV, but the SB creation energy was kept at 10.79 eV. This change was made in part to be consistent with results simulated using another code, but also to include the lower excitation potentials of liquid water. The SSB and DSB cluster yields (and sizes), as well as the direct portion of RBE_{DSB} (RBE_{DSB}^{direct}) with reference to ^{60}Co was compared for all particles. It was particularly found that the RBE_{DSB}^{direct} increases with LET and is higher for protons than for $^{12}C^{6+}$ ions of the same LET.

Importantly, comparisons of RBE_{DSB}^{direct} with RBE_{DSB} from experimental measurements and other published simulations revealed similar trends, suggesting that considerations of the initial pattern of ED sites play a rather large role in predicting biological effectiveness.

Authors: Piotr Pater, Gloria Bäckstöm, Fernanda Villegas, Anders Ahnesjö, Shirin A. Enger, Jan Seuntjens, and Issam El Naqa.

Published in: *Medical Physics*, **45**, 5, pp 2131–2140, (2016).

Abstract

Purpose: To present and characterize a MC tool for the simulation of the RBE_{DSB}^{direct} for protons and light ions.

Methods: The MC tool uses a pregenerated event-by-event tracks library of protons and light ions that are overlaid on a cell nucleus model. The cell nucleus model is a cylindrical arrangement of nucleosome structures consisting of 198 DNA bps. An algorithm relying on *k*-dimensional trees and cylindrical symmetries is used to search coincidences of ED sites with volumes corresponding to the SPB of the DNA molecule. SBs are scored when energy higher than a threshold is reached in these volumes. Based on the number of affected strands, they are categorized into either SSB or DSB clusters. The number of SBs composing each cluster (i.e., its size) is also recorded. RBE_{DSB}^{direct} is obtained by taking the ratio of DSB yields of a given radiation field to a ^{60}Co field. The MC tool was used to obtain SSB yields, DSB yields and RBE_{DSB}^{direct} as a function of LET for protons ($^{1}H^{+}$), $^{4}He^{2+}$, $^{7}Li^{3+}$, and $^{12}C^{6+}$ ions.

Results: For protons, the SSB yields decreased and the DSB yields increased with LET. At $\approx$ 24.5 keV μm^{-1}, protons generated 15% more DSB clusters than $^{12}C^{6+}$ ions. The RBE_{DSB} varied between 1.24 and 1.77 for proton fields between 8.5 keV μm^{-1} and 30.2 keV μm^{-1}, and it was higher for iso-LET ions with lowest atomic number. The SSB and DSB cluster sizes showed significant differences for all radiation fields. Generally, the yields of SSB clusters of sizes ≥ 2 and the yields of DSB clusters of sizes ≥ 3 increased with LET and increased for iso-LET ions of lower atomic number. On the other hand, the ratios of SSB to DSB clusters of sizes 2–4 did not show variability with LET nor projectile atomic number, suggesting that these metrics are independent of the radiation quality. Finally, a variance of up to 8% in the DSB yields was observed as a function of the particle incidence angle on the cell nucleus. This simulation effect is due to the preferential alignment of ion tracks with the DNA nucleosomes at specific angles.

Conclusions: The MC tool can predict SSB and DSB yields for light ions of various LET and estimate RBE_{DSB}^{direct}. In addition, it can calculate the frequencies of different DNA lesion sizes, which is of interest in the context of biologically relevant absolute dosimetry of particle beams.

5.1 Introduction

Proton and light ion beams are rapidly gaining popularity as a radiation therapy treatment option due to their ability to conform dose to the target while sparing surrounding healthy tissues.[33] These beams have different clinical effects for equivalent physical doses as compared to conventional x-rays. Therefore, the RBE is a useful measure in comparing their effectiveness. The RBE is defined with

respect to a given biological endpoint, many of which have been studied.[15,115] One of these endpoints is the induction of DNA DSB clusters believed to be a precursor to cell death. DSB clusters are formed when SBs on both DNA strands are created within a short distance, hence inducing a high risk for repair corruption. The induction of DNA DSB clusters can be measured experimentally,[116–119] calculated with MC simulations,[63,70,86,90] or inferred from measured ICSDs in volumes equivalent to a few nanometers using nanodosimeters and combinatorics approaches.[32] In this context, this work aims to present a simulation framework for the calculation of the $\mathrm{RBE}_{\mathrm{DSB}}^{\mathrm{direct}}$ at clinically relevant doses and energies for protons and light ions. It can be formally defined as

$$\mathrm{RBE}_{\mathrm{DSB}}^{\mathrm{direct}} = \frac{\overline{D_{N_{\mathrm{DSB}}}}|_{^{60}\mathrm{Co}}}{\overline{D_{N_{\mathrm{DSB}}}}|_{Q}}, \tag{5.1}$$

where $\overline{D_{N_{\mathrm{DSB}}}}$ is the average dose required to induce N_{DSB} direct DSB clusters per cell and is evaluated for the ^{60}Co reference radiation quality and the radiation quality of interest (Q). Direct DSB clusters are exclusively composed of SBs produced by direct radiation effects in the DNA SPB volumes, such as ionizations and excitations. They only form a subset of all DSB clusters, and do not include those due to indirect effects. Focusing only on the direct DSB component allows investigating biologically relevant initial patterns of radiation damage, and attempting to relate these patterns to physically measurable parameters using nanodosimeters. Therefore, $\mathrm{RBE}_{\mathrm{DSB}}^{\mathrm{direct}}$ is a quantity of interest in the context of biologically relevant absolute dosimetry of particle beams,[33] with the goal of differentiating physical

effects of different radiation qualities. This quantity differs from RBE obtained from *in vivo* DSB assays, which includes radical damage and DNA repair.

It is widely assumed that differences in the initial physical energy deposition patterns can be linked to subsequent relevant biological outcomes. It has been postulated that a multiscale framework based on MCTS simulations could subsequently be linked to biological outcomes of interest such as cell kill, mutations or even, at a later stage, to tumor control probability and normal tissue complication probability.[3] Several MC codes have been developed for event-by-event track simulation of particles with energy below 0.1 keV.[85] These codes rely on experimental, theoretical or semi-empirical models for interaction cross-sections at very low energies. Some limits of low energy simulations have been discussed elsewhere,[35,36,120] as well as difficulties in obtaining valid models for biological or condensed matter.[102] Nevertheless, these codes allow investigations of coincidences of ED sites and DNA SPB volumes. The underlying hypothesis of this approach is that initial yields of DSB clusters and their complexities can serve as building blocks of a multistage framework of radiation response.[121]

The current study introduces a simulation framework to estimate RBE_{DSB}^{direct} as a function of LET for protons, several light ions ($^{4}He^{2+}$, $^{7}Li^{3+}$ and $^{12}C^{6+}$), and ^{60}Co photons for clinical doses by calculating direct DSB yields. The framework makes use of a pregenerated tracks library to decrease simulation time, which in turn allows gathering larger statistics. In particular, this study covers the algorithms used to create the tracks, to irradiate a cell nucleus and to compute DNA damage.

Tracks generated with any MCTS code can be used, which makes the simulation framework well suited for comparison studies.

Furthermore, the simulation framework was used to study the RBE_{DSB}^{direct} trends for protons and light ions of various radiation qualities. Biological effects of radiation are related to LET, which gives the linear density of energy depositions along the particle track.[17] However, LET alone is not enough to explain all differences in RBE, as, for instance, different ions at the same LET (iso-LET ions) can exhibit different RBE_{DSB}.[115] Due to large experimental uncertainties, the 10–20% variation in RBE_{DSB} between different ion species might be considered unimportant. Nevertheless, it is expected to see an increased biological effectiveness for particles with lower kinetic energy. In fact, for different iso-LET ions, the kinetic energy increases as a function of the projectile's atomic number Z, and thus, the RBE_{DSB} is expected to decrease for iso-LET ions of higher Z. The simulation framework was also used to computeSSB and DSB cluster yields and their associated sizes (i.e., the number of SBs composing each type of cluster). In this context, "single" and "double" refer only to the number of DNA strands affected in the cluster and not to the number of individual SBs. Short deletions or other types of potentially irreparable damage may arise with increased probability, as the DNA lesions get more complex. The complexity of DNA lesions following proton and ion irradiation has been studied by several authors using MCTS[18,69] and probabilistic approaches.[32,54,70] Probabilistic approaches assume that SBs have an equal probability of occurring on either strand; therefore, simple combinatorics can be used to estimate DNA

lesion types and sizes. SSB and DSB cluster sizes are measures of the DNA damage severity, which can also be compared to experimentally obtained ICSDs and derived quantities.[122] Finally, an angular artifact arising when overlaying tracks on a fixed DNA description was highlighted in this study. To our knowledge, this was not previously studied, yet it can significantly affect simulation results.

5.2 Methods and materials

5.2.1 Simulation framework

The simulation framework overlays particle tracks generated by an MCTS code on a cell nucleus model and determines the yields, types and sizes of the DNA clusters created.

Generation of particle tracks

The simulation framework can use tracks generated from any MCTS code. For this study, the GEANT4.10.01.p02 simulation toolkit[23] with the *G4EmDNAPhysics* physics list available through the GEANT4-DNA project[24] was used to generate tracks and populate the pregenerated tracks library. GEANT4-DNA allows the simulation of electron, proton, and light ion tracks in an event-by-event mode. Electrons were tracked to a kinetic energy of $E_{\text{cut}} = 9\,\text{eV}$, at which point their energy was deposited locally. For every ED site, several parameters were stored in a binary file, including the particle type, process type, position, kinetic energy, energy deposited, and energy transferred. The simulation framework requires tracks generated under transient electronic equilibrium in the incidence direction. Therefore, all tracks were generated in a semi-infinite slab of liquid water with density of $1.0\,\text{g}\,\text{cm}^{-3}$. The slab's thickness ($t_{\text{slab}}$) was chosen according to the

particle type and kinetic energy in order to guarantee that a central portion of length $t_{\text{ion}} = 10\,\mu\text{m}$ is in transient CPE. For every radiation quality, $t_{\text{slab}} = t_{\text{ion}} + 2R(Q)$, where $R(Q)$ is a padding distance equal to the secondary electron maximum range for the given radiation quality Q (see Figure 5–1). The padding distances account for electronic buildup and backscatter in the central length t_{ion}. Only the ED sites within the t_{ion} central portion are stored in the pregenerated track library. For protons and light ions, 500 tracks with an initial kinetic energy T_0 were generated. Some radiation qualities were specifically chosen as to yield iso-LET fields of different ions types, referred as iso-LET ions (see Table 5–1).

For ^{60}Co, the Livermore physics list of GEANT4 was used to score the primary and scattered photon spectrum in a $20\,\text{cm}^3$ liquid water cube irradiated by a pencil beam of 1.17 and 1.33 MeV photons. In a subsequent simulation, 1000 energies were sampled from the spectrum and used to generate electron tracks in a large liquid water volume using GEANT4-DNA processes. The number of generated ^{60}Co tracks was doubled compared to other particles due to a lower dose per track deposited at this energy. On average contributions from 375 tracks per Gray of dose to the nucleus were needed.

Irradiation of the cell nucleus

Cell irradiation is simulated by overlaying tracks from the pregenerated library onto a cell nucleus modeled as a cylinder (all cylinders referred to in this work are right circular cylinders) of radius (r_n) and half-height ($h_n/2$) both equal to 2112.375 nm. For each radiation field studied, 100 fractions at ten polar angles θ equally spaced from 0° to 90° are generated. In each fraction, tracks are

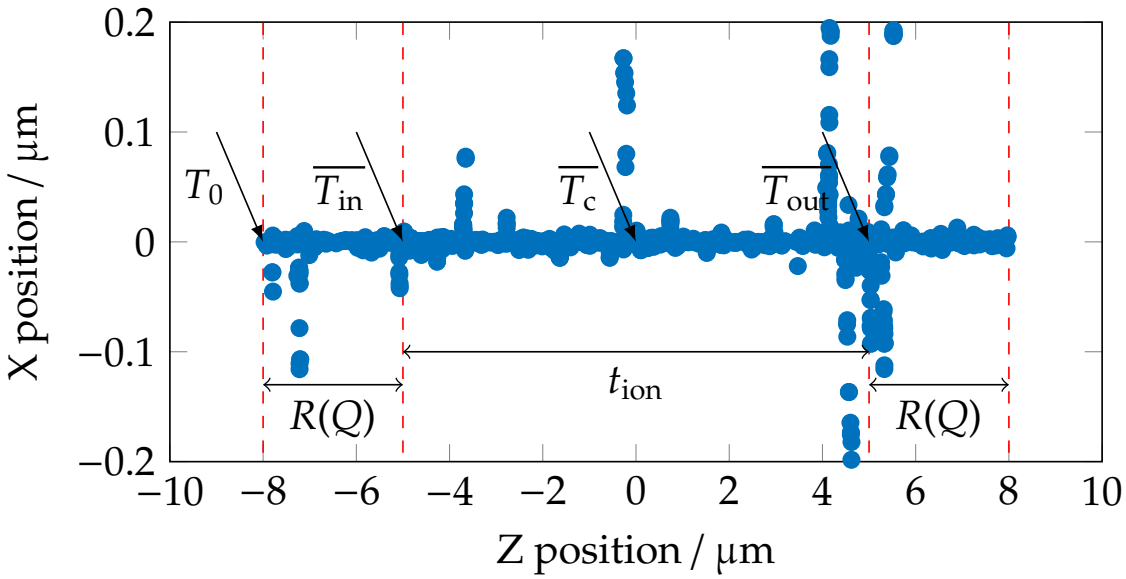

Figure 5–1: Example proton track with initial kinetic energy $T_0 = 5.03\,\text{MeV}$. The central segment t_{ion} is shown as well as the padding distances $R(Q)$. The positions of the initial (T_0), entrance ($\overline{T_{\text{in}}}$), central ($\overline{T_{\text{c}}}$), and exit ($\overline{T_{\text{out}}}$) kinetic energies of the proton are also represented.

Table 5–1: Source specific parameters. LET was estimated as LET = $(\overline{T_{\text{in}}} - \overline{T_{\text{out}}})/t_{\text{ion}}$. The radiation fields referred to as iso-LET ions are grouped at the end of the table.

Q	$R(Q)$ μm	T_0 MeV	$\overline{T_{\text{in}}}$ MeV	$\overline{T_{\text{c}}}$ MeV	$\overline{T_{\text{out}}}$ MeV	LET keV μm^{-1}
^{60}Co	—	N/A	—	—	—	0.30
Protons	3.00	5.03	5.00	4.96	4.92	8.70
Protons	1.70	3.64	3.62	3.57	3.51	11.14
Protons	1.39	3.25	3.23	3.17	3.11	12.18
Protons	0.51	1.82	1.81	1.72	1.62	19.15
Protons	0.40	1.52	1.51	1.40	1.29	22.25
Protons	0.25	1.15	1.14	1.01	0.86	28.17
Protons	0.20	1.07	1.06	0.92	0.76	30.22
^{4}He^{2+}	3.00	20.00	19.91	19.77	19.62	29.33
^{7}Li^{3+}	27.00	103.20	102.46	102.32	102.18	27.56
Protons	0.40	1.35	1.34	1.22	1.09	24.65
^{4}He^{2+}	5.50	25.32	25.31	25.19	25.07	24.22
^{7}Li^{3+}	27.00	120.40	119.75	119.63	119.50	24.25
^{12}C^{6+}	553.00	1185.00	1171.41	1171.29	1171.17	24.33

overlaid until a dose of 2 Gy is reached in the nucleus. Then, the positions and energies associated with each ED site occurring in the cell nucleus are registered for investigation of coincidences with SPB volumes.

For ion beams, a disk source is positioned to be tangent to a sphere of radius $t_{\text{ion}}/2$, with the azimuthal coordinate ϕ randomly and uniformly sampled between 0 and 2π rad and with the polar coordinate θ fixed *a priori* (see Fig. 5–2.) The fixation of θ allows to study the influence of the beam's incidence angle on the DNA damage yields. The radius of the source is $r_s = R(Q) + \sqrt{2}r_n$ to account for secondary electrons from distant, nontraversing tracks. Tracks are randomly chosen from the pregenerated tracks library and positioned at randomly and uniformly distributed points on the surface of the disk. In detail, the radial position r_{track} is sampled between 0 and r_s ($r_{\text{track}} = r_s\sqrt{U}$, where $U \in [0,1]$ is a randomly and uniformly sampled real number) and the angular position γ is randomly and uniformly sampled between 0 and 2π rad. In addition, each track is rotated around its axis by an angle τ, which is randomly and uniformly sampled between 0 and 2π rad. On average for proton fields, less than 15 tracks contribute to the cell nucleus dose per Gray. Therefore, using a pregenerated library of 500 tracks with directions (θ and ϕ), initial positions (r_{track} and γ) and rotation (τ) randomly sampled to further increase the tracks' variability is a sufficient population for this study.

For ^{60}Co, this procedure would be inefficient due to the size of the source needed to achieve CPE. Instead, dose is accumulated by random weighted selection of a track and positioning of the cell nucleus in a dose-receiving position. Track

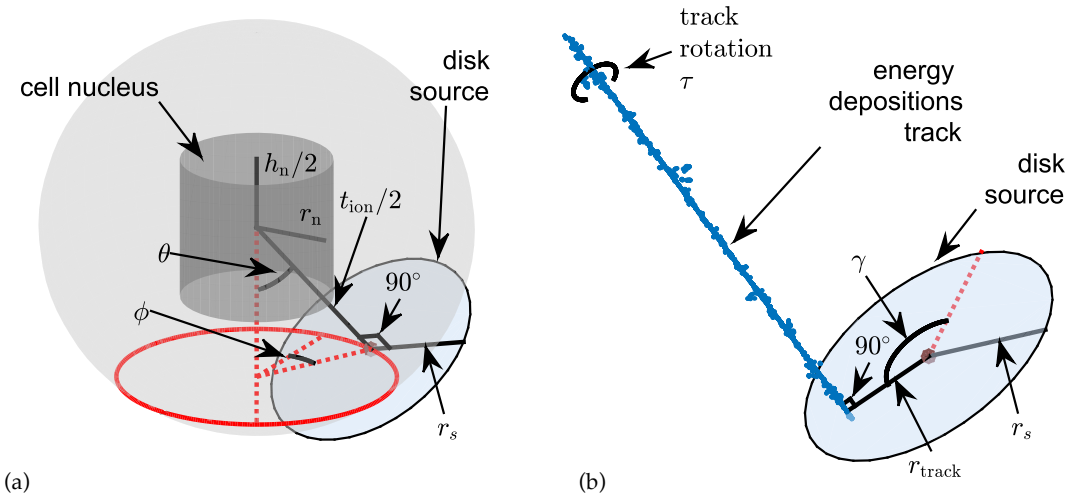

Figure 5–2: Algorithm for overlaying proton and light ion tracks on the cell nucleus.(a) A disk source of radius r_s is positioned around a cell nucleus at a distance $t_{ion}/2$ with a randomly sampled azimuthal angle ϕ and a polar angle θ fixed *a priori*. (b) A track is randomly selected from the pregenerated tracks library and positioned normally to the disk source. The track's radial position r_{track}, angular position γ, and angular rotation τ are randomly sampled. Additional tracks are overlaid on the cell nucleus until a dose of 2 Gy is attained.

weights are proportional to the volume encompassing all possible nuclear positions where dose can be accumulated. Therefore, longer tracks have a larger dose contribution, as expected from microdosimetric consideration of μ-randomness.[123] For each selected track, the cell nucleus is placed at random positions until a nonzero dose is scored. Track selection and nucleus positioning is repeated until a cumulative dose of 2 Gy is reached. The angles θ, ϕ, and τ are sampled as previously described.

Calculation of DNA damage

The mammalian cell nucleus model introduced in this work (see Fig. 5–3) uses a previously published nucleosome model.[40,86] However, the total number of bps and the overall geometry were modified to encompass the whole genome. The nucleus consists of 13 446 instances of the chromatin fiber (cf), which is modeled as an elongated cylinder of radius $r_{\text{cf}} = 15\,\text{nm}$ and height $h_{\text{cf}} = h_n$. The cf instances are arranged along 64 equally spaced shells and a central cf. The radial position of the ith shell is $r^i_{\text{shell}} = (2r_{\text{cf}} + g) * i$, where $g = 2.5\,\text{nm}$ is a gap between shells, with the central cylinder corresponding to $i = 0$. The number of cfs in the ith shell (for $i \geq 1$) is $n^i_{\text{shell}} = \frac{2\pi}{\Theta_i}$, with $\cos(\Theta_i) = \frac{i^2 - 0.5}{i^2}$ giving the minimum angle between adjacent cfs on the ith shell. The actual angular positions of each cf in the ith shell are given by multiples of $\frac{2\pi}{n^i_{\text{shell}}}$ rad, with the initial position randomly chosen. Each cf is composed of 393 levels of a 6-nucleosome structure (6N), modeled as a cylinder of radius $r_{6N} = r_{\text{cf}}$ and height $h_{6N} = 10.5\,\text{nm}$. Each 6N is composed of six independent nucleosomes (N), which were previously described in detail.[40,86] Each nucleosome is composed of two independent turns (T) of 99 bps arranged in a double-stranded fashion. The SPB thickness is $h_{\text{bp}} \approx 0.183\,\text{nm}$, and its volume is $V_{\text{sp}} \approx 0.13\,\text{nm}^3$. The total genomic content of the cell nucleus model ($N_{\text{bps}} \approx 6.28 \times 10^9$ bps) and the cell nucleus volume ($V_{\text{nucl}} \approx 59.2\,\mu\text{m}^3$), are roughly representative of the values for a human cell.

SBs are scored in a given SPB volume if the total energy deposit exceeds $E_{\text{th}} = 10.79\,\text{eV}$, corresponding to the lowest possible ionization potential of liquid water. It is equivalent to state that every ionization occurring in the SPB volume

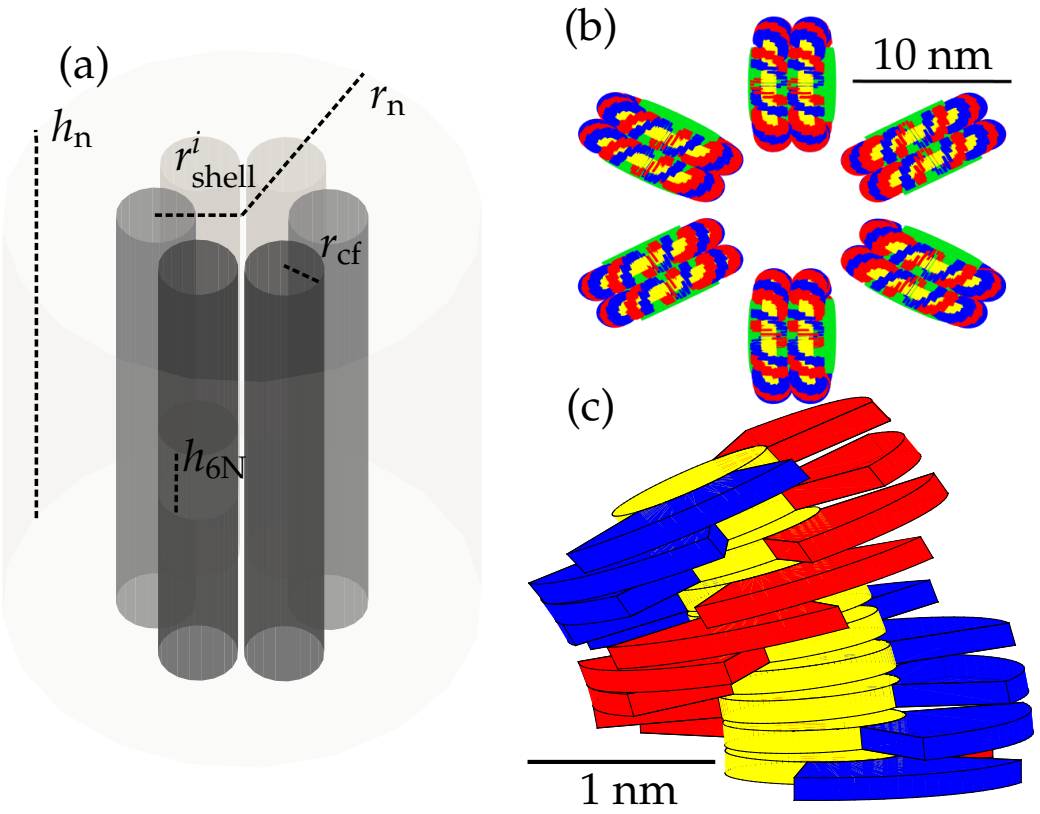

Figure 5–3: (a) Representation of the cell nucleus with the six cf's from the first shell. In addition a $6N$ structure is shown on one cf. (not to scale) (b) A detailed view of a $6N$ structure composed of six nucleosomes. Each nucleosome is formed of two turns of 99 bps wound up around a histone molecule (green cylinder). (c) Representation of a 10 bps segment with dimensions described in detail in Bernal *et al.*[40,86]

leads to an SB. The value of E_{th} largely influences the absolute SSB and DSB yields obtained.[1] Other SB induction algorithms and E_{th} values were also previously used by different authors,[63,111,112,124] based on best-fit approaches to experimental data-sets.

A *k*-dimensional tree algorithm (a space partitioning data structure that determines, for each point in a data-set, its closest neighbor from a set of test points) based on the nuclear cylindrical symmetries is used to find the SB genomic positions. This algorithm improved the DNA damage search times when compared to our own implementation of serial searches over all ED sites or over all SPB volumes. In detail, the algorithm determines the cf closest to each ED site and stores its instance index in memory (1–13 446). Then, a coordinate transformation is performed to move all ED sites to the central cf. ED sites with a radial position larger than r_{cf} are discarded, because they fall outside of any cf and do not participate in the creation of strand breaks. A similar procedure is then used to find the 6N instance index (1–393), the nucleosome index (1–6), the nucleosome turn index (1 or 2), the bp index (1–99), and the SPB volume index (1 or 2) closest to each ED site. At each step, the instance index is stored, all ED sites are moved and those that fall outside of the organizational level of interest are discarded.

SBs in close proximity are classified by type into SSB or DSB clusters. Definitions and algorithms used to classify these cluster types differ among various authors, which may lead to large absolute differences.[1] In this work, the algorithms from the MCDS code[28] were adapted in order to count SSB and DSB clusters. All SBs occurring within a genomic distance smaller or equal to 10 bps in a nucleosome

turn are grouped into the same cluster, otherwise they are grouped into different clusters. Then, each cluster is classified as either a DSB cluster if both strands exhibit at least one SB, or into an SSB cluster if all SBs occur on the same strand. In addition to the cluster type, its size is also recorded as a quantitative measure of the complexity of the DNA damage.

5.2.2 Yields of SSB and DSB lesions

For each radiation quality, DNA damage yields are obtained by normalizing the number of SSB ($N^{i,\theta}_{\mathrm{SSB}}$) and DSB ($N^{i,\theta}_{\mathrm{DSB}}$) clusters in the ith fraction at a polar angle θ by the total cell nucleus dose, $D^{i,\theta}_{n}$,

$$y^{i,\theta}_{\mathrm{SSB}} = N^{i,\theta}_{\mathrm{SSB}}/D^{i,\theta}_{n}, \tag{5.2}$$

$$y^{i,\theta}_{\mathrm{DSB}} = N^{i,\theta}_{\mathrm{DSB}}/D^{i,\theta}_{n}, \tag{5.3}$$

where $D^{i,\theta}_{n}$ depends on the residual dose over 2 Gy given by the last overlaid track. Furthermore, mean angular SSB and DSB yields ($\overline{y^{\theta}_{\mathrm{SSB}}}$ and $\overline{y^{\theta}_{\mathrm{DSB}}}$) are calculated for each polar angle θ by averaging over all fractions. Isotropically weighted mean SSB and DSB yields ($\overline{Y_{\mathrm{SSB}}}$ and $\overline{Y_{\mathrm{DSB}}}$) are obtained by taking a weighted average of $\overline{y^{\theta}_{\mathrm{SSB}}}$ and $\overline{y^{\theta}_{\mathrm{DSB}}}$, with weighting factors (w_{θ}) such that $w_{\theta} = \frac{\sin(\theta)}{\sum_{\theta=0}^{\theta=\pi/2}\sin(\theta)}$. The $\overline{Y_{\mathrm{SSB}}}$ and $\overline{Y_{\mathrm{DSB}}}$ represent an isotropic average of various beam orientations with respect to the cell nucleus orientation. In other words, this data-set represents the irradiation of many cell nuclei, oriented randomly with respect to a parallel beam. All yields are expressed in units of $(\mathrm{Gy\,Gbp})^{-1}$ and are independent of the fraction dose up to hundreds or thousands of Grays, where intertrack effects become nonnegligible.[87]

Finally, our yield definitions specifically show that each fraction data-set is analyzed collectively and not in a per-track fashion, as was previously done to simplify computation.[125] For instance, DSB yields per track are calculated by dividing the number of DSB clusters produced by the jth track of the ith fraction ($N^{i,j}_{\mathrm{DSB}}$) by the track dose ($D^{i,j}_{n}$). Although the number of DSB in the ith fraction (N^{i}_{DSB}) is approximately equal to the average number of DSB per track in the ith fraction times the number of tracks in the ith fraction (n^{i}_{t}),

$$N^{i}_{\mathrm{DSB}} \approx \Big(\frac{\sum_{j=1}^{n^{i}_{t}} N^{i,j}_{\mathrm{DSB}}}{n^{i}_{t}}\Big) n^{i}_{t}, \tag{5.4}$$

it can be shown that the average DSB yield per-track times n^{i}_{t} does not equal to y^{i}_{DSB} for a given polar angle θ,

$$y^{i}_{\mathrm{DSB}} = \frac{N^{i}_{\mathrm{DSB}}}{D^{i}_{n}} \neq \frac{\sum_{j=1}^{n^{i}_{t}} \frac{N^{i,j}_{\mathrm{DSB}}}{D^{i,j}_{n}}}{n^{i}_{t}} n^{i}_{t}. \tag{5.5}$$

The inequality in Eq. (5.5) arises from the fact that the sum of ratios is not equal to the ratio of sums, a common statistical misconception. The approximation sign in Eq. (5.4) is due to the neglect of intertrack effects.

5.2.3 RBE for the induction of direct DSB

The $\mathrm{RBE}^{\mathrm{direct}}_{\mathrm{DSB}}$ was defined in Eq. (5.1). From the yield definition, $\overline{D_{N_{\mathrm{DSB}}}} = N_{\mathrm{DSB}}/\overline{Y_{\mathrm{DSB}}}$, therefore,

$$\mathrm{RBE}_{\mathrm{DSB}} = \frac{\overline{Y_{\mathrm{DSB}}}\,|_{Q}}{\overline{Y_{\mathrm{DSB}}}\,|_{^{60}\mathrm{Co}}}. \tag{5.6}$$

RBE_{DSB}^{direct} represents the induction of direct DSB clusters only and our simulation framework ignores all indirect effects. It is independent of the absolute number of DSB clusters (N_{DSB}) as seen in Eq. (5.1).

5.3 Results

5.3.1 $y_{SSB}^{i,\theta}$ and $y_{DSB}^{i,\theta}$ distributions as a function of θ

The angular distributions in Fig. 5–4 resemble normal distributions as the p values yielded by the one-sample Kolmogorov–Smirnov test for normality[126] are above 0.05, except for the 0° distribution. At this angle, tracks are parallel to the cf axes and have a greater probability of creating DSB clusters due to their preferential alignment with the nucleosome turns. This preferential alignment is a geometrical simulation artifact, which increases the DSB yields as θ aligns with the nucleosome turns, due to the linear paths followed by proton and light ion beams. It is also partly due to the ordering of the DNA model itself, which allows for these preferential directions. Therefore, the $\overline{y_{DSB}^{\theta}}$ significantly increases from 6.03 DSB (Gy Gbp)$^{-1}$ to 6.56 DSB (Gy Gbp)$^{-1}$ from 90° to 10° (two-sample t-test with p value < 0.005). The $y_{SSB}^{i,\theta}$ distributions (not shown) do not exhibit any angular dependencies as SSB clusters of size 1 dominate these distributions and are independent of the geometrical alignment. Similar trends are observed for all radiation qualities except for ^{60}Co, where electron tracks do not tend to follow straight lines.

Additionally, the boxplots from Fig. 5–4 demonstrate the stochastic variability in DSB creation at any angle. Each boxplot represents 100 2 Gy irradiations of a cell nucleus with tracks of the same LET, in the same conditions. For instance, the

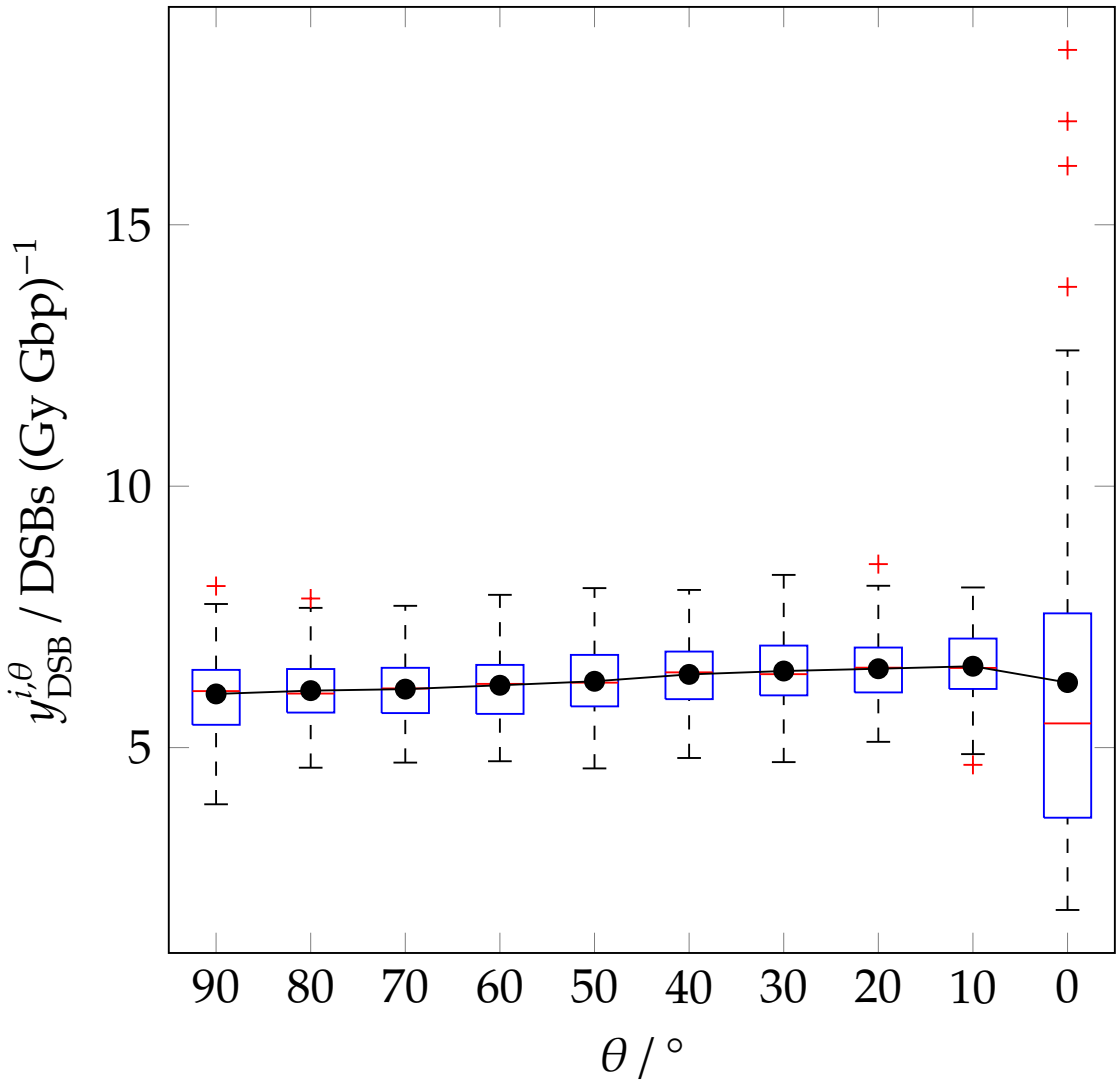

Figure 5–4: Angular distributions $y^{i,\theta}_{\mathrm{DSB}}$ for protons with LET = 30.22 keV µm^{-1}. The distribution average $\overline{y^{\theta}_{\mathrm{DSB}}}$ (●), medians (—), 25$^{\mathrm{th}}$ and 75$^{\mathrm{th}}$ percentiles (▯), ≈ 3σ (⊤), and outliers (+) are shown.

simulated $y^{i,\theta}_{\mathrm{DSB}}$at θ = 90° varied between 3.92 DSB (Gy Gbp)$^{-1}$ and 8.09 DSB (Gy Gbp)$^{-1}$ due to the stochastic nature of energy deposition in the cell nucleus.

5.3.2 $\overline{Y_{SSB}}$ and $\overline{Y_{DSB}}$ as a function of LET

All Y_{SSB} and Y_{DSB} distributions resemble normal distributions as the p values yielded by the one-sample Kolmogorov–Smirnov test for normality are above 0.05. Therefore, Fig. 5–5 presents $\overline{Y_{SSB}}$ and $\overline{Y_{DSB}}$ with error bars representing ±1.96 times the standard error on the mean (SEM). In this case, the SEM of the weighted average was calculated using the Cochran method.[127] For protons (♦), a decreasing trend for $\overline{Y_{SSB}}$ and an increasing trend for $\overline{Y_{DSB}}$ as a function of LET were obtained. These trends are significant for all proton fields simulated as confirmed by the two-sample t-test for unequal variances yielding $p < 0.001$ in all cases. Irradiation with ^{60}Co creates significantly lower DSB yields than for all other particles, in agreement with higher expected RBE for ions. Finally, for the iso-LET ions at $\approx 24.55\,\mathrm{keV\,\mu m^{-1}}$, SSB yields are maximal for $^{12}C^{6+}$ ions, which have the highest atomic number Z in our simulations, and significantly decrease as projectile Z falls to 1. The inverse trend is seen for DSB yields with a maximum for protons and a minimum for $^{12}C^{6+}$ ions. It is noteworthy to mention that the average total number of direct SBs was of 86.8 ± 0.5 SB/(Gy Gbp)$^{-1}$ and was invariant with respect to radiation quality.

5.3.3 Relative contribution of lesion sizes to $\overline{Y_{SSB}}$ and $\overline{Y_{DSB}}$

Table 5–2 presents the relative contributions of cluster sizes to $\overline{Y_{SSB}}$, $\overline{Y_{DSB}}$, and $\overline{Y_{SSB}}/\overline{Y_{DSB}}$. For protons, SSB clusters of size 1 represent between 90 % to 95 % of all SSB clusters and their number tends to decrease with increasing LET. DSB clusters of size 2 represent between 71 % to 78 % of all DSB clusters and their number tends to decrease with increasing LET. Therefore, for proton fields, as the

LET increases, both SSB and DSB clusters tend to be more complex. However, the absolute yield of SSB clusters tends to decrease whereas it increases for DSB clusters. In the case of iso-LET ions, cluster complexity tends to decrease with increasing projectile atomic number Z, as shown in the bottom four rows of Table 5–2. Therefore, the SSB and DSB clusters created by protons have higher sizes than those created by $^{12}C^{6+}$ ions of the same LET. In addition, the $\overline{Y_{SSB}}/\overline{Y_{DSB}}$ ratio decreases with increasing LET and decreasing atomic number. However the ratios of the yields of clusters of sizes 2–4 do not show clear trends as a function of LET or atomic number. For all radiation fields, they stay constant around 1.29, 0.42, and 0.18 for clusters of sizes 2, 3 and 4 respectively, suggesting that these ratios are independent of the radiation fields.

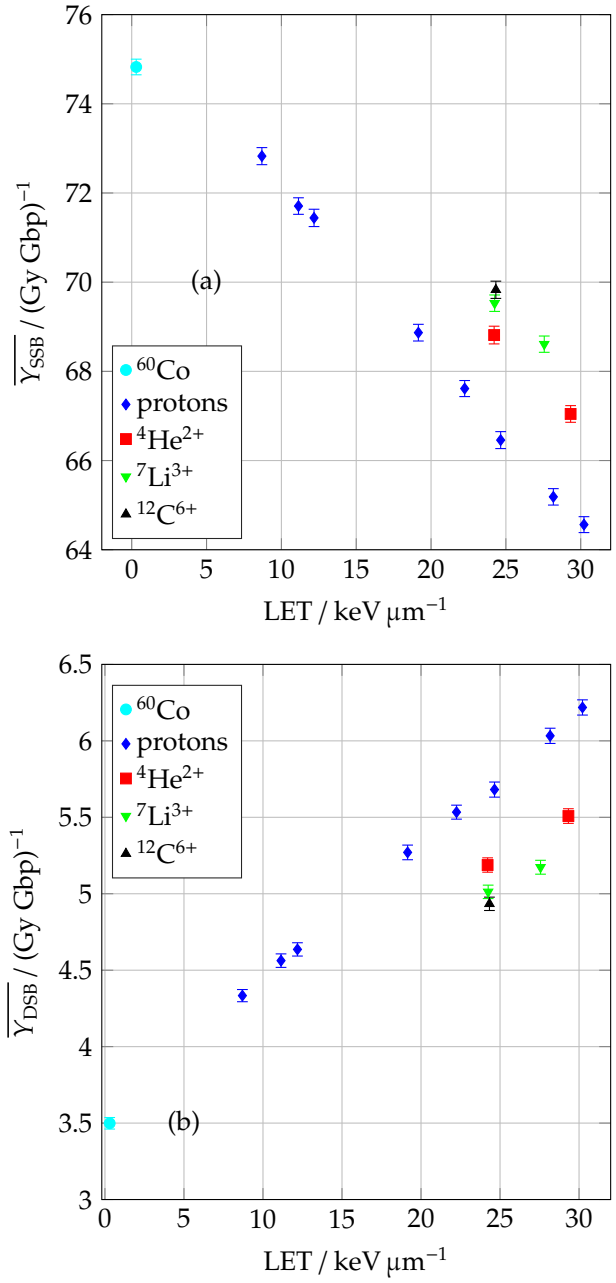

Figure 5–5: $\overline{Y_{SSB}}$ (a) and $\overline{Y_{DSB}}$ (b) as a function of LET. The error bars represent 1.96 times the SEM. For protons, $\overline{Y_{SSB}}$ decreases and $\overline{Y_{DSB}}$ increases with LET. For iso-LET ions at $\approx 24.5\,\mathrm{keV\,\mu m^{-1}}$, $\overline{Y_{SSB}}$ increases and $\overline{Y_{DSB}}$ decreases with the ion's atomic number.

Table 5–2: Average SSB and DSB yields and relative contributions from different lesion sizes. In addition, the ratios of $\overline{Y_{SSB}}$ to $\overline{Y_{DSB}}$ as well as the ratios of the average yields of SSB to DSB lesions of sizes 2–4 are presented. The number of individual strand breaks (SBs) composing a lesion is defined as its size, with a minimum of 1 SB for SSB and 2 SBs for DSB. The SEM is expressed in compact notation, i.e. 325.6(5) = 325.6 ±0.5. Lesions of sizes ≥5 occurred infrequently.

		SSB yields			DSB yields			SSB/DSB ratios			
Ion	LET keV µm^{-1}	$\overline{Y_{SSB}}$ (Gy Gbp)$^{-1}$	Size 1 %	Size ≥2 %	$\overline{Y_{DSB}}$ (Gy Gbp)$^{-1}$	Size 2 %	Size ≥3 %	$\overline{Y_{SSB}}/\overline{Y_{DSB}}$	Size 2	Size 3	Size 4
^{60}Co	0.3	74.8(1)	94.7(1)	5.27(4)	3.50(2)	77.9(5)	22.0(4)	21.40(1)	1.34(1)	0.43(1)	0.21(2)
Protons	8.7	72.8(1)	93.5(1)	6.54(4)	4.33(2)	77.1(4)	22.9(4)	16.80(8)	1.31(1)	0.42(1)	0.18(1)
Protons	11.1	71.7(9)	93.2(1)	6.84(4)	4.56(2)	77.0(4)	23.0(4)	15.72(8)	1.29(1)	0.41(1)	0.18(1)
Protons	12.2	71.4(1)	93.0(1)	7.02(4)	4.64(2)	76.7(4)	23.3(3)	15.41(8)	1.29(1)	0.43(1)	0.18(1)
Protons	19.2	68.9(1)	92.0(1)	8.03(5)	5.27(2)	74.8(4)	25.2(3)	13.07(6)	1.28(1)	0.41(1)	0.16(1)
Protons	22.2	67.6(9)	91.5(1)	8.50(5)	5.53(2)	74.2(4)	25.8(3)	12.22(5)	1.27(1)	0.44(1)	0.18(1)
Protons	28.2	65.2(9)	90.7(1)	9.27(5)	6.03(3)	71.8(3)	28.2(4)	10.81(5)	1.26(1)	0.41(1)	0.16(1)
Protons	30.2	64.6(9)	90.4(1)	9.60(5)	6.22(3)	71.1(3)	28.9(4)	10.38(4)	1.26(1)	0.41(1)	0.18(1)
^{4}He^{2+}	29.3	67.0(1)	91.4(1)	8.61(5)	5.51(2)	73.9(4)	26.1(3)	12.17(6)	1.29(1)	0.43(1)	0.17(1)
^{7}Li^{3+}	27.6	68.6(9)	92.0(1)	7.99(5)	5.17(2)	74.8(4)	25.2(3)	13.26(6)	1.29(1)	0.42(1)	0.18(1)
Protons	24.7	66.5(1)	91.2(1)	8.75(5)	5.68(3)	73.0(4)	27.0(4)	11.70(5)	1.27(1)	0.42(1)	0.17(1)
^{4}He^{2+}	24.2	68.8(1)	92.0(1)	7.99(5)	5.19(2)	75.1(4)	24.9(4)	13.26(6)	1.29(1)	0.43(1)	0.17(1)
^{7}Li^{3+}	24.2	69.5(9)	92.3(1)	7.69(5)	5.01(2)	75.8(4)	24.2(3)	13.87(6)	1.29(1)	0.43(1)	0.19(1)
^{12}C^{6+}	24.3	69.8(1)	92.4(1)	7.65(5)	4.93(2)	75.8(4)	24.2(3)	14.15(7)	1.30(1)	0.44(1)	0.18(1)

5.3.4 RBE_{DSB} as a function of LET

In Fig. 5–6, the RBE_{DSB}^{direct} for protons (♦), $^{4}He^{2+}$ ions (■), $^{7}Li^{3+}$ ions (▼), and $^{12}C^{6+}$ ions (▲) were calculated using Eq. (5.6) and exhibit increasing trends as a function of increasing LET and decreasing projectile atomic number. The RBE_{DSB}^{direct} for iso-LET ions with LET $\approx 24.5\,keV\,\mu m^{-1}$ are equal to 1.62, 1.48, 1.43, and 1.41 for protons, $^{4}He^{2+}$, $^{7}Li^{3+}$, and $^{12}C^{6+}$ ions, respectively. For protons, the RBE_{DSB}^{direct} trends obtained with our simulation framework compare well with the RBE_{DSB} of the PARTRAC code[63] and the MCDS code version 3.10A.[30] Additionally, experimental measurement of DSB cluster induction for protons and alpha particles show an increase of RBE_{DSB} with respect to LET, with a lower RBE_{DSB} for alpha particles than protons at the same LET.

5.4 Discussion

The simulation framework presented in this work allows for the calculation of direct SSB and DSB yields. It can differentiate between initial biological effects of various incident radiation qualities. The increasing trends in RBE_{DSB}^{direct} as a function of LET are consistent with experimental measurements, MCTS simulations, and theoretical observations. However, simulations in general, and our framework, in particular, rely on many parameters, definitions, cross-sections, and an exact absolute simulation of the DNA damage yields is challenging for several reasons.

First off all, accurate event-by-event tracking of electrons with kinetic energies below ≈ 100 eV is difficult, due to an increase of the uncertainties associated with electron position and momentum.[35,36] Also, processes other than ionizations are known to be effective at inducing SBs, even below the ionization threshold for

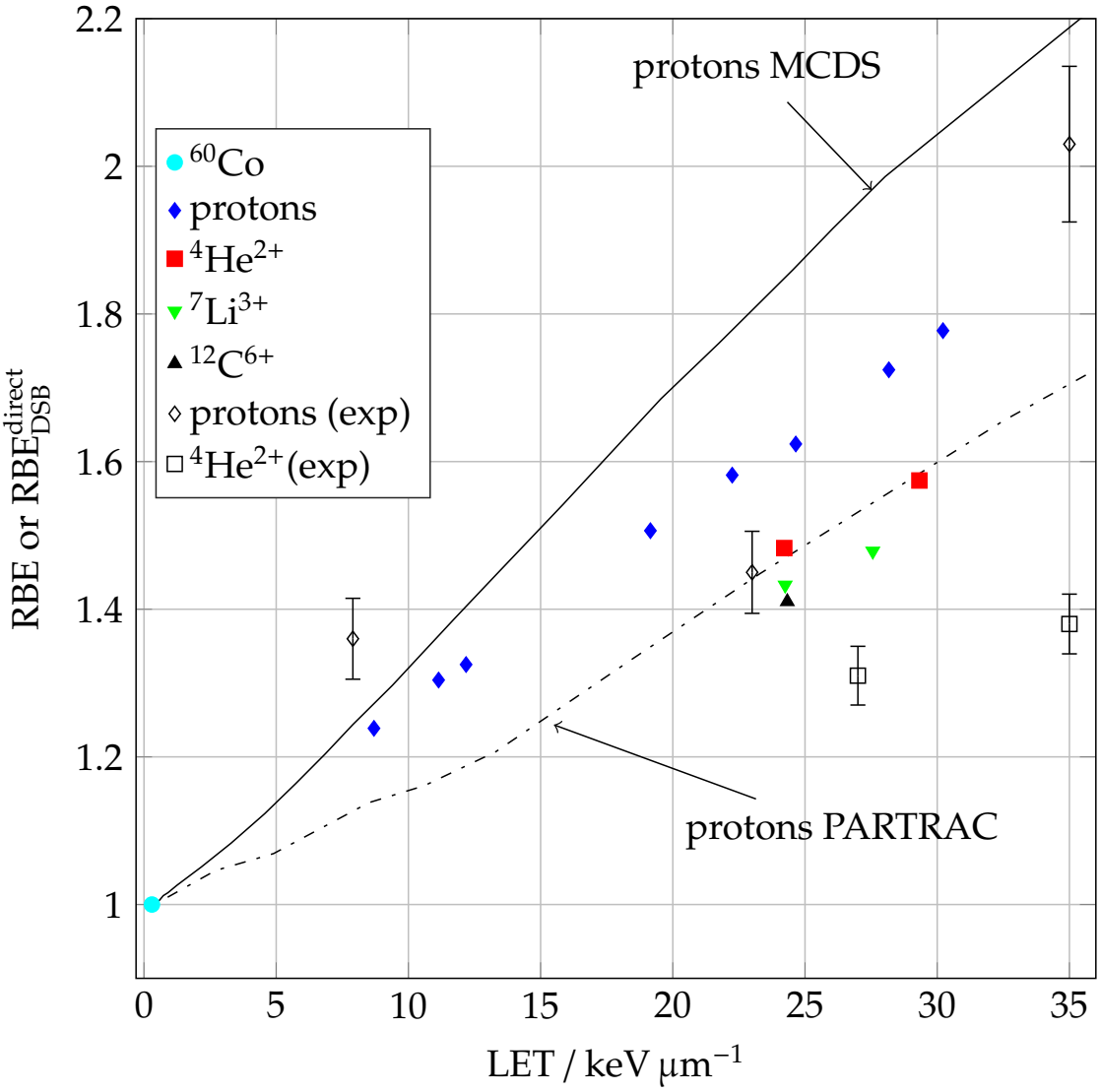

Figure 5–6: RBE_{DSB} or RBE as a function of LET calculated using Eq. (5.6) for this work (solid symbols), PARTRAC simulations (Ref. 63) (dashed-dotted line), MCDS simulations version 3.10A in aerobic conditions (Ref. 30) (solid line), and experimental measurements for protons and $^{4}He^{2+}$ ions (Ref. 116). This work shows RBE_{DSB} whereas all other points are RBE for the induction of DSBs (including indirect effects). Relative errors of our simulations are all below 1 % and smaller than the symbols.

liquid water[65] and they are currently not adequately simulated. In addition, accurate cross-sections for DNA constituents at low electron energies surrounding liquid water were not measured until very recently and are not integrated in most MC codes.[128,129] Moreover, inelastic cross-section model choices can influence simulated absolute SSB and DSB yields.[130] Second, the modeling of DNA damage induction requires a number of adjustable parameters, definitions, and algorithms for the DNA model itself, for the induction of SBs, and for the classification of the clusters, which influence the absolute simulated yields to a large degree.[1] Therefore, SSB and DSB yields reported should not be seen as valid in an absolute sense.

The cell nucleus geometrical model presented in this work has a volume of about $60\,\mu m^3$ and is composed of 6.2×10^9 bps, which are good estimates for a typical human cell nucleus. For instance, the human genome is constituted of approximately 3×10^9 bps and most human cells are diploid.[131] In addition, the average nucleus size of HeLa type cells is $374\,\mu m^3$,[132] and the horizontal cross-section of V79 cells is $168\,\mu m^2$.[133] These dimensions are about 6 to 27 times larger than the cell nucleus model; however, they comprise all nuclear subcompartments and not only DNA, and this size difference does not affect simulation results. One weakness of this model is that all nucleosome turns, nucleosomes, 6Ns and cfs are independent, whereas DNA is a long continuous molecule. This simplification should not have a large impact on the yields of SSB and DSB clusters, as each nucleosome turn is sufficiently long. However, the proposed model does not allow tallying genomic distances between DNA damage lesions, which might be

of interest in estimating the biological effectiveness of radiation quality.[84] A more realistic model, with linker DNA might lead to slightly different absolute yields.[89]

Our results showed that the $y_{\mathrm{DSB}}^{i,\theta}$ is sensitive to θ. There is a synergy between the particle's incidence angle and the nucleosomes orientations. On average, the number of individual SBs created in the cell nucleus is independent of the particle incidence angle. However, the angle affects the SB genomic location. More DSB clusters are created when primary particles pass across the nucleosomes, with a lower θ. In addition, at 0°, the distribution enlarges because particles can pass either through the center of a cf depositing dose, but not creating any DNA damage, or through many nucleosomes which in turn increases DSB cluster production. In this work, we chose a weighting method to obtain average results equivalent to an isotropic irradiation. In itself, this result demonstrates potential pitfalls that need to be considered in scenarios involving proton or light ion tracks overlaid on a nucleus model.

The track overlaying methodology uses sources large enough to cover the cell nucleus in a uniform fluence of parallel tracks. Fig. 5–4 shows that for any given polar angle, the $y_{\mathrm{DSB}}^{i,\theta}$ distributions are rather wide. This variation is due to the stochastic probability of tracks interacting with DNA targets and should be distinguished from the spread in microdosimetric energy deposition for populations of cells irradiated to the same macroscopic dose.[134] In fact, to obtain a full distribution of the average number of DSB clusters per cell in a macroscopic cell population, the yields shown here have to be convolved with the microdosimetric spread in the dose to cells.

As shown in Fig. 5–5, $\overline{Y_{\mathrm{SSB}}}$ decreases and $\overline{Y_{\mathrm{DSB}}}$ increases for proton fields as a function of LET. As the proton energy decreases, the ED sites become bunched closer together, which effectively increases the DSB yields, as well as the yields of SSB of size ≥2. For iso-LET ions, $\overline{Y_{\mathrm{SSB}}}$ decreases whereas $\overline{Y_{\mathrm{DSB}}}$ increases as a function of decreasing projectile atomic number Z. Experimentally, these trends are difficult to see;[115] however, they are theoretically expected. Iso-LET ions with higher Z have a higher kinetic energy; therefore, the secondary electrons they release are more energetic, travel further and create sparser ED site patterns than ions of lower Z (and lower kinetic energy).[135] It explains why protons create more DSB clusters and less SSB clusters than $^{12}C^{6+}$ ions of the same LET. This result is in agreement with PARTRAC simulations, where lower DSB yields for He ions than for protons were found for an LET higher than 20 keV µm^{-1}.[69] Direct DSB yields obtained in this work represent 35–40 % of PARTRAC's total DSB yields, which is expected due to their indirect damage parametrisation.[63]

The invariance of the total direct SB yield is consistent with previous literature, where invariant values of 130 SB/(Gy Gbp)$^{-1}$ (Ref. 40) and 75 SB/(Gy Gbp)$^{-1}$ (Ref. 63) were found for various radiation qualities. The absolute difference between these values is determined mainly by E_{th} and differences in the SB definitions.[1]

SSB and DSB cluster size distributions were shown in Table 5–2. Overall, as LET increases for protons, or as the projectile Z decreases for iso-LET ions, more DSB clusters are created, and the SSB and DSB cluster sizes tend to increase. Therefore, the DNA damage becomes more complex, and potentially more difficult to repair. The ratios of SSB/DSB of sizes 2–4 stayed constant for all radiation qualities

suggesting they are not a function of the radiation field. This is the assumption used in probabilistic approaches[32,54,70] were each strand is an equally-likely target for SBs (50% probability on each strand). According to these probabilistic approaches, clusters with i SBs have a $\tilde{p}(i) = 1 - (\frac{1}{2})^{i-1}$ probability of not being on the same strand, which gives SSB/DSB ratios of 1, 0.33, and 0.14 for clusters of sizes 2, 3, and 4, respectively. These ratios are about 27 to 29 % lower than the ones obtained using the full geometrical DNA model of nucleosomes. We believe that the equiprobability of SBs occurring on any strand is only true for isolated clusters, whereas the actual distances between SBs have an impact on the type of cluster (SSB or DSB) for clusters of sizes ≥ 2.

The $\mathrm{RBE}_{\mathrm{DSB}}^{\mathrm{direct}}$ from Fig. 5–6 varied between 1.23 and 1.77 for protons. These are in good agreement with $\mathrm{RBE}_{\mathrm{DSB}}$ trends of PARTRAC simulations,[63] MCDS simulations,[30] and experimental results, where indirect damage is taken into account. Comparisons of $\mathrm{RBE}_{\mathrm{DSB}}^{\mathrm{direct}}$ with $\mathrm{RBE}_{\mathrm{DSB}}$ may be vague at first; however, it is apparent that a large portion of the total RBE increase with respect to LET is primarily contributed by the direct effects of radiation, which in turn are simply dictated by the beam-specific differences in the initial patterns of energy depositions. Continued development of nanodosimetry will allow for accurate measurement of these patterns. Our framework shows significant differences as a function of radiation quality in quantities such as direct SSB and DSB clusters, as well as in their sizes. These quantities have a conceptually closer biological meaning than their nanodosimetric counterparts such as ICSDs and related moments. It is well

understood that direct DNA damage is difficult if not impossible to isolate experimentally and that it cannot explain the full spectrum of measured DNA damage. It is nevertheless a quantity worth looking at because it can solely be explained by physical considerations of a given beam. The fact that RBE_{DSB}^{direct} and RBE_{DSB} trends are consistent suggests that indirect effects or biological factors alone might have an overall limited dependence on the initial radiation quality. In the future, the RBE_{DSB}^{direct} could be isolated away during the beam calibration phase, leaving only the characterization of effects which have a smaller LET dependence. Our simulation results show that RBE_{DSB}^{direct} has a large, significant, and predictable effect on the RBE_{DSB} and is only dependent on the initial patterns of energy depositions.

5.5 Conclusion

This work presents a simulation tool for direct DNA damage patterns of protons and light ions. The results follow expected trends as a function of LET for protons and as a function of projectile atomic number for iso-LET ions. Direct effects of radiation are of interest as they are created by patterns of energy deposition, which can also be measured using experimental methods. Moreover, these patterns are intrinsic characteristics of the radiation fields and show significant differences with particle type and LET. Further studies of these biologically relevant energy deposition patterns are interesting in designing novel dosimetry standards for ion beams, which could potentially include physically measurable biological effects in the absolute dose specification.

5.6 Acknowledgments

We would like to acknowledge support from the Fonds de recherche du Québec – Santé (FRQS), from the CREATE Medical Physics Research Training Network grant (number 432290) of the Natural Sciences and Engineering Research Council (NSERC), support from NSERC under grants RGPIN 397711-11 and RGPIN-2014-06475 and support from the Canadian Institutes of Health Research (CIHR) under grants MOP-114910, MOP-136774 and MOP-102550.

CHAPTER 6
Technical note : Estimation of clinical photon beam biological effects using event-by-event electron spectra

Chapters 4 and 5 used MC generated tracks of electron, proton, and light ions, coupled with a geometrical model of DNA, to obtain direct DNA damage yields. In both these chapters, indirect effects were not modeled. Even though indirect effects can count for more than 50 % of the total DNA damage, they are not needed to study DNA damage trends as a function of LET. In addition, simulations of indirect effects require a separate set of adjustable parameters and algorithms that complicate the calculations. Moreover, the track overlay of Chapter 5 assumes equal energy particles incident on the cell nucleus, ignoring scattering and straggling in the medium. In this chapter, the track overlay methodology is extended for the simulation of realistic clinical situations that incorporate depth and position of the cell in the phantom, clinical beam sizes, and adequate scattering conditions. This is done by computing the spectrum of electron incident on and created inside a cell irradiated by photon beams. These spectra could be further used in the generation of tracks for overlay on a geometrical description of DNA to obtain direct DSB clusters. Instead, in this study, the MCDS code is used to obtain the RBE_{DSB} for clinical photon fields, including CBCT fields, an EBT x-ray source, and several radio-isotopes used in brachytherapy. Importantly is was found that

the RBE_{DSB} for CBCT and EBT x-ray fields ranged from 1.14–1.16 comparatively to a 1 MeV photon field.

Authors: Piotr Pater, Gabriel Famulari, Issam El Naqa, and Jan Seuntjens.
Submitted to: *Physics in Medicine and Biology*, on May 27th, 2016.

Abstract

With the goal of estimating the relative biological effectiveness for the induction of double strand breaks (RBE_{DSB}) in clinically relevant photon irradiations, we devised a method for obtaining and studying the detailed, event-by-event, spectra of electrons incident on and of primary recoil electrons created inside a μm-sized cell irradiated by photons of energy below 1 MeV. These spectra are necessary intermediate outputs and inherently dictate the change of RBE_{DSB}. First, we tallied the spectrum of photons interacting near a spherical scoring volume (SV) irradiated by large beams in large phantoms. It is used in the generation of electron tracks using the event-by-event mode of Geant4-DNA, which are then overlaid on the cell according to the *associated volume* concept. Finally, a tally of the electron spectra is performed and $RBE_{DSB}$is estimated using the Monte Carlo Damage Simulation code. Our methodology contrasts with previously reported ones using the photon fluence to obtain a primary recoil electron spectrum, which is an approximation that does not hold for clinical situations where incident electrons are dominant. In a benchmark simulation, our methodology correctly reproduced experimental and simulated RBE_{DSB} for ultrasoft x-rays. The RBE_{DSB} for a 50 kVp miniature x-rays, and for 80 and 125 kVp cone beam CT fields were found to range

between 1.14 and 1.16. For brachytherapy isotopes, the simulated RBE_{DSB} were equal to 1.19, 1.12, and 1.04 for I-125, Yb-169 and Ir-192 respectively. This study confirms that the RBE_{DSB} generally increases with decreasing photon energy, and presents a consistent methodology for studying its changes as a function of the simulation geometry.

6.1 Introduction

The yield of DNA double strand breaks (DSBs) induced by photon irradiation increases at lower beam energies.[1,28,30,31,70,71,73,136,137] Monte Carlo track structure simulations can offer an effective and fast alternative to experimental measurements for studying the dependence of the relative biological effectiveness (RBE) for the induction of DSBs, RBE_{DSB}, on energy, depth, beam size, tissue type and other treatment parameters. In these simulations, a microscopic scoring volume (SV) representative of a cell nucleus is irradiated and DNA damage is deduced either from coincidences of energy deposition with an embedded geometrical DNA model or from semi-empirical models relating the incident radiation field to probabilistic distributions of DNA lesions. The absorbed dose in the SV is due to two distinct sources of electrons (see figure 6–1); electrons originating from photon interactions in the SV form the primary recoil electron spectrum (Φ_{pr}) and electrons incident on the SV after being slowed-down in the environment form the incident electron spectrum (Φ_{inc}). Both spectra lead to DNA damage through their interactions in the SV. Earlier simulation studies of RBE_{DSB} focused on single-cell exposures where the totality of the dose was delivered by Φ_{pr} and it's progeny within the SV, effectively ignoring Φ_{inc}. This approximation was necessary in order

to limit the simulation volumes and diminish simulation times. Event-by-event tracking of electrons in macroscopic volumes are not practical computationally. The challenges consist in speeding up simulations while preserving the dosimetric and microdosimetric properties of clinical beams, which include scattering conditions, representative phantom and SV sizes, dose uniformity and angular incidence. The intent of this work is therefore to present a precise multiscale methodology for obtaining Φ_{pr} and Φ_{inc} spectra using event-by-event electron tracking for clinical photon sources in large simulation geometries. The methods are applied to compute electron spectra of ultrasoft x-rays, a 50 kVp intro-operative radiation therapy source, 80 and 125 kVp cone-beam computed tomography (CBCT) fields, iridium-192, iodine-125 and Ytterbium-169 brachytherapy sources and a cobalt-60 reference field. Furthermore, with the use of the Monte Carlo Damage Simulation (MCDS) code,[30,31] RBE_{DSB} for these fields are estimated.

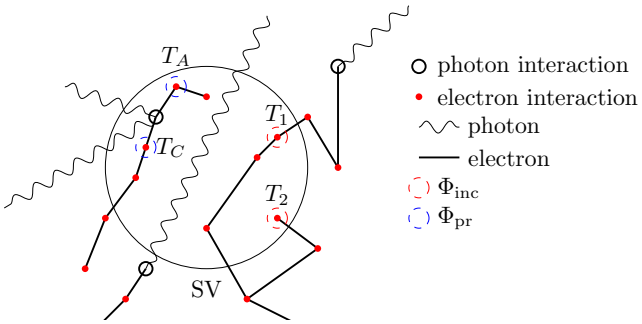

Figure 6–1: Visual representation of electrons counted in Φ_{inc} or Φ_{pr}. The Compton electron and the Auger electron with kinetic energies T_C and T_A, respectively, are counted in Φ_{pr}, because they were liberated by a photon interaction in the SV. The electrons with kinetic energies T_1 and T_2 are part of Φ_{inc}, as they are incident on the SV. All other electrons are not scored. T_C, T_A, T_1, and T_2 correspond to the kinetic energy of the electrons just prior to the interaction.

6.2 Methods and Materials

6.2.1 Multiscale methodology for obtaining the electron spectra

A multiscale methodology was designed to compute Φ_{inc} and Φ_{pr} for a spherical SV of radius r_{SV}. First, the photon interaction spectrum, $\Phi_i(h\nu)$, was tallied using Geant4, using the standard Livermore physics, with fluorescence active and electron tracking turned off.[23] The $\Phi_i(h\nu)$ consists of the energies of photons generating electrons energetic enough to reach the SV (see figure 6–2a). It was scored in the sphere V around the SV with dimensions consistant with the secondary electron maximal range (R_{max}). In a second simulation, electrons tracks were generated in an event-by-event mode down to 9 eV (this corresponds to the lowest excitation potential of liquid water) with Geant4-DNA physics.[24]This was performed for 10^3 initial photons with energy sampled from $\Phi_i(h\nu)$ in an infinite liquid water phantom (see figure 6–2b). The progeny of all electrons generated by the first compton or photoelectric interaction was stored in a pre-generated tracks library. In the third and final step, Φ_{inc} and Φ_{pr}were computed by overlaying each track on the SV, while preserving the microdosimetric properties of the beam. Instead of naive sampling consisting in random placement of a track around a fixed SV, the SV was randomly placed within the track's associated volume (AV), corresponding to the union of spheres of radius r_{SV} centered on the track's energy deposition (ED) sites (see figure 6–2d).[123,138–140] Computationally, it was achieved by placing the SV at 10^3 different positions randomly sampled in a spherical volume of radius equal to r_{SV} around a random ED site in the track. The kinetic energy T of all electrons

incident on the SV was scored in Φ_{inc} and of all electrons generated by photon interactions inside the SV were scored in Φ_{pr}.

The multiscale methodology is computationally efficient even for energetic tracks, as all sampled positions lead to increments in the spectra. However, it introduces two biases that need to be addressed with appropriate weights for every electron energy scored. Firstly, tracks with larger AV have an increased probability of intersecting the SV and have to be sampled more frequently. Instead, each track was sampled 10^3 times, and a weight, $w_{\text{av}}(T)$, proportional to the AV (estimated by numerical simulations using a rejection method) was introduced. Secondly, the AV method biases towards regions of high ED density by randomly selecting an ED site in the track. Therefore, a weight, $w_{\text{ED}}(T)$, proportional to the inverse of the total number of ED sites inside the SV is introduced. In summary, a weight $w(T) = w_{\text{ed}}(T)w_{\text{av}}(T)$ is associated to every incident or primary recoil electron.

6.2.2 Estimation of DNA damage

The average DSB yield in a SV, Δ, is obtained with:

$$\Delta = \frac{\sum_{i=1}^{n_{\text{inc}}+n_{\text{pr}}} \Delta(T_i)\overline{z_F}(T_i)w(T_i)}{\sum_{i=1}^{n_{\text{inc}}+n_{\text{pr}}} \overline{z_F}(T_i)w(T_i)} \tag{6.1}$$

where n_{inc} and n_{pr} are the total number of electrons incident and created inside the SV, respectively, $\overline{z_{\text{F}}}(T)$ is the frequency-mean specific energy per event obtained from separate Geant4-DNA simulations, and $\Delta(T_i)$ is the DSB yield due to monoenergetic electrons of energy T_i obtained from published data generated with the Monte Carlo damage simulation (MCDS) code for aerobic cell conditions.[31,71]

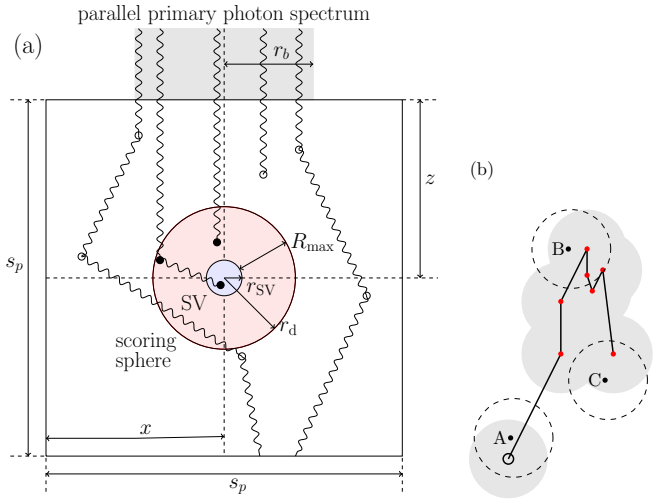

Figure 6–2: (a) The first part of the simulation consists on scoring the energies of photons interacting in the SV or in the sphere surrounding the SV (filled circles) to generate $\Phi_i(h\nu)$. (b) In the second part, electron tracks are generated and overlaid on the SV in order to score Φ_{inc} and Φ_{pr}. Here is a representation of an electron track composed of energy depositing interactions (red circles), starting from a photon interaction (open circle). The track's AV corresponds to the shaded area. Also shown are 3 different positions of the SV (A, B, C) randomly positioned in the AV as described in the text. This placement guarantees dose deposition for each SV position.

The RBE_{DSB} of a given source is calculated by dividing Δ by the average DSB yield of a reference Co-60 radiation source.

Overall, more than 10^3 electron tracks were generated for each simulation, and sampled upwards of 10^3 times each. We estimate our DSB statistical uncertainties to be below 0.1 % in all cases. This estimate is based on convergence of the average, but does not take into consideration other uncertainties related to our input data.

The Co-60 reference beam was simulated in Geant4-DNA as a source of 1.25 MeV photons. However, Geant4-DNA electron cross-sections are limited to below 1 MeV. Therefore, in the Co-60 simulations, a small proportion of electrons

generated above 1 MeV had their energy changed to 1 MeV. We tested that this change has no impact on DSB yields. Our DSB yields for the Co-60 simulations was found to be 8.36 and did not vary significantly with depth. This value is in good agreement with the generally accepted value of 8.3 for Co-60.[71]

6.3 Results

6.3.1 Validation

The multiscale methodology was validated by comparing its performance against brute force calculations for 1 and 10 keV photon beams in a cubic phantom of sides equal to 20 μm. The brute force approach consists in simulating all photon and electron interactions in the phantom using an event-by-event mode and is thus limited to small phantoms and low energies. Several series of simulations of Φ_{inc}, Φ_{pr}, and $\overline{z_F}$ revealed good agreement between both methods. As an example, figure 6–3a compares Φ_{inc} in a SV with r_{SV} = 5 μm and 6–3b compares the $\overline{z_F}$ for increasing r_{SV} between 0.1 and 10 μm. The agreement between both methods is well within the statistical uncertainty, for all sizes studied, which confirms the microdosimetric consistency of the multiscale methodology. It is important to note that the brute force approach is computationally intensive, and impractical for high photon energies, large phantoms and large beams.

Additionally, a validation for several ultra-soft x-rays sources in the single-cell setup was carried out. In these simulations, the cell is in vacuum; therefore, only Φ_{pr} contributes to the damage. Overall, excellent agreement (below 5% difference) is observed for all soft x-ray sources between the DSB yields obtained with the multiscale model estimates and previously published experimental and simulated

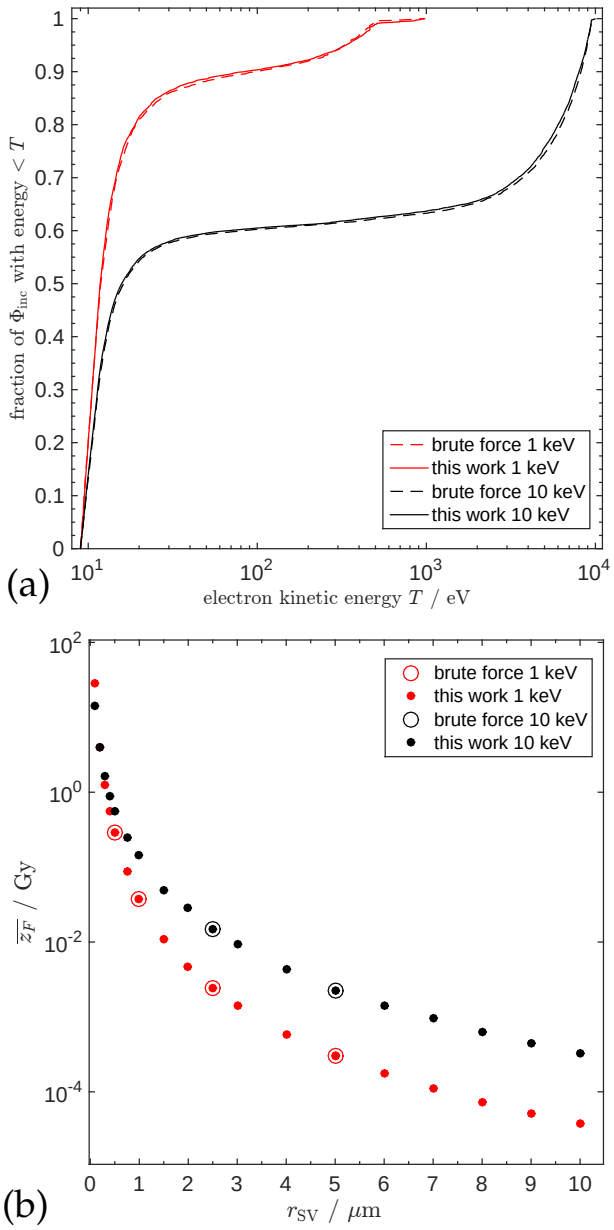

Figure 6–3: (a) Comparison of Φ_{inc} in cumulative form (a) and $\overline{z_F}(T)$ (b) obtained with the brute force and the multiscale model in 1 and 10 keV photon sources incident on a SV of radius equal to 5 μm. Standard deviations are smaller than the symbols and not shown.

datasets (see Table 6–1).[71,141] For Co-60, the experimental data point acquired with the pulse field gel electrophoresis technique is 10–13% lower than both simulations. Nevertheless, these differences are well within experimental uncertainties. Taken together, both the brute force comparisons and the ultra-soft x-ray benchmarking confirm that the multiscale model performs as expected in these conditions.

Table 6–1: Benchmarking of simulated DSB yields (in DSBs Gy^{-1} Gbp^{-1}) with the multi-scale model (Δ_{msm}) against simulations (Δ_{sim})[71] and experiments (Δ_{exp}).[141]

type	hν (keV)	Δ_{msm}	Δ_{sim}	Δ_{exp}
C_K	0.28	20.0	19.8	20.7
Cu_L	0.969	16.6	16.6	17.4
Al_K	1.49	14.2	14.2	14.3
Ti_K	4.55	10.3	10.8	10.4
Co-60	1250	8.73	8.4	7.6

The differences between the photon interaction spectrum ($\Phi_i(h\nu)$) and the track-length photon fluence, $\Phi(h\nu)$,[80] are depicted in figure 6–4a. Calculations at depths of 2, 5, and 15 cm in a cubic phantom of sides equal to 20 cm irradiated by 1 MeV monoenergetic photon beams were performed. The increased scatter contribution with depth is reflected in both the $\Phi(h\nu)$ and $\Phi_i(h\nu)$ spectra. However, $\Phi_i(h\nu)$ is significantly different from the photon fluence, and weighted towards lower energies for all depths. This is expected as higher energy photons pass more often through the volume without interaction. The sharp rise exhibited by all curves around 0.2 MeV is attributed to Compton backscattered photons.

Figure 6–4b compares Φ_{inc} and Φ_{pr} for a SV of 5 μm at 5 cm depth irradiated by 1 MeV photons. These spectra were obtained by sampling tracks obtained from either $\Phi_i(h\nu)$ or $\Phi(h\nu)$(see legend). Overall, the use of $\Phi(h\nu)$ instead of $\Phi_i(h\nu)$

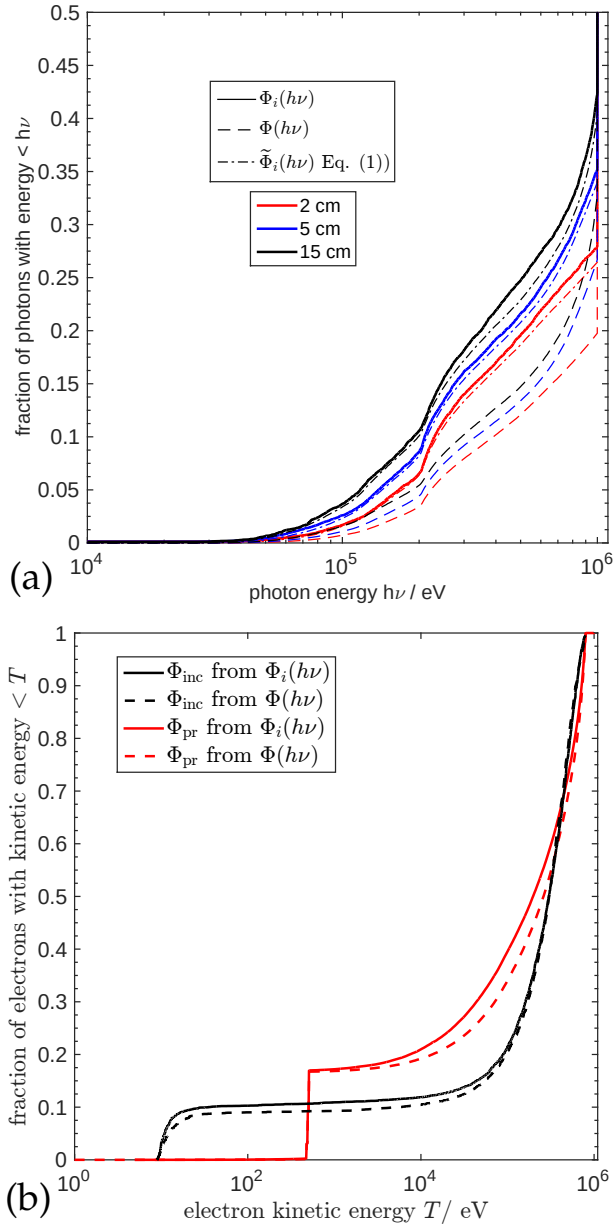

Figure 6–4: (a) Comparison of $\Phi_i(h\nu)$ and $\Phi(h\nu)$ for a 1 MeV monoenergetic photon beam at depths of 2, 5, and 15 cm on the central axis. (b) Comparison of $\Phi_{\rm pr}$ and $\Phi_{\rm inc}$ calculated using $\Phi_i(h\nu)$ or $\Phi(h\nu)$ for 1 MeV photons for the 5 cm depth simulation.

has a small but significant influence on the simulated electron spectra seen by the SV. The Φ_{pr} exhibits a sharp rise at 500 eV due to the creation of Auger electrons, followed by a continuous increase up to the maximum Compton recoil electron energy around 0.8 MeV. Additionally, a small percentage is due to photo-electrons just below 1 MeV. Further, Φ_{inc} differs significantly from Φ_{pr}. For instance, electrons above 150 keV represent 70 % of all electrons in Φ_{inc}, whereas they account for only 55 % of the electrons in Φ_{pr}. Additionally, the fraction $F_{pr}(h\nu, r_{SV}) = n_{pr}/(n_{inc} + n_{pr})$ gives the proportion of primary recoil electrons in the SV out of all electrons incident and created in the SV by photon interactions. As the r_{SV} decreases, relatively more electrons are incident on the SV, whereas the proportion of electrons created inside the SV diminishes (see table 6–2). The importance of Φ_{pr} diminishes as the energy increases or the SV size decrease. Particularly, at 1 MeV and for a 5 μm SV, primary recoil electrons account for only 15.4% of all electrons incident or created in the SV by photon interactions.

Table 6–2: F_{pr} as a function of $h\nu$ and r_{SV} expressed in %.

	$F_{pr}(h\nu, r_{SV})$ (%)		
r_{SV} (μm)	$h\nu$ = 1 keV	$h\nu$ = 10 keV	$h\nu$ = 1 MeV
0.1	37.7	5.6	2.4
0.5	73.5	13.4	6.8
1.0	84.6	25.2	8.8
2.5	92.6	50.7	13.6
5.0	96.9	66.2	15.4
10	98.4	80.4	17

Figure 6–5a presents $\overline{z_F}(T)$ obtained by irradiating SVs of 0.1, 1, and 5 μm radius with monoenergetic electrons using Geant4-DNA. The total absorption

limit, $\overline{z_F}(T) = 4T/3\pi\rho r^3$, is given by the dotted lines and a continuous slowing down approximation (CSDA), $\overline{z_F}(T) = 0.204L_0/4r^2$, is given by the dashed blue lines.[30,142] For a given electron energy, $\overline{z_F}(T)$ decreases with increasing sphere volume. The maximum occurs when the range associated with T aligns with the mean chord length through the sphere. Small differences in $\overline{z_F}(T)$ are also seen when comparing the types of electron sources. Notably, for T below the maximum, the surface isotropic source does not reach full absorption as is the case for the volumetric source. Figure 6–5b presents $\overline{z_F}(T)\Delta(T)$ for all simulations. Although the DSB yield increases with decreasing electron energy, it is generally not true to say that a 10 eV electron creates more DSBs than a 100 eV electron. Correct comparison of the DNA damage of different electron energies needs to be weighted by the frequency-mean specific energy $\overline{z_F}(T)$ (figure 6–5). From these figures, it is evident that the most damaging electron energy is one at which the electron range aligns with the mean chord length through the SV, hence 0.9, 5-8, and 15-20 keV for 0.1, 1 and 5 μm spheres, respectively. The DSBs Gbp^{-1} can be read directly from figure 6–5b for any electron energy. Low energy electrons may have the highest DSBs yield per Gy, however they lead to two order of magnitude less DSBs per incident electron than 1 to 20 keV electrons.

Table 6–3 shows estimates of DSB yields obtained with the multiscale methodology for Co-60 fields at different depths, for 80 and 125 kVp CBCT fields (spectra obtained from personal communication with Charles Kirkby[71]), for a 50 kVp intraoperative radiotherapy beam (spectrum obtained from Peter Watson[143,144]), and for several isotopes relevant to brachytherapy.

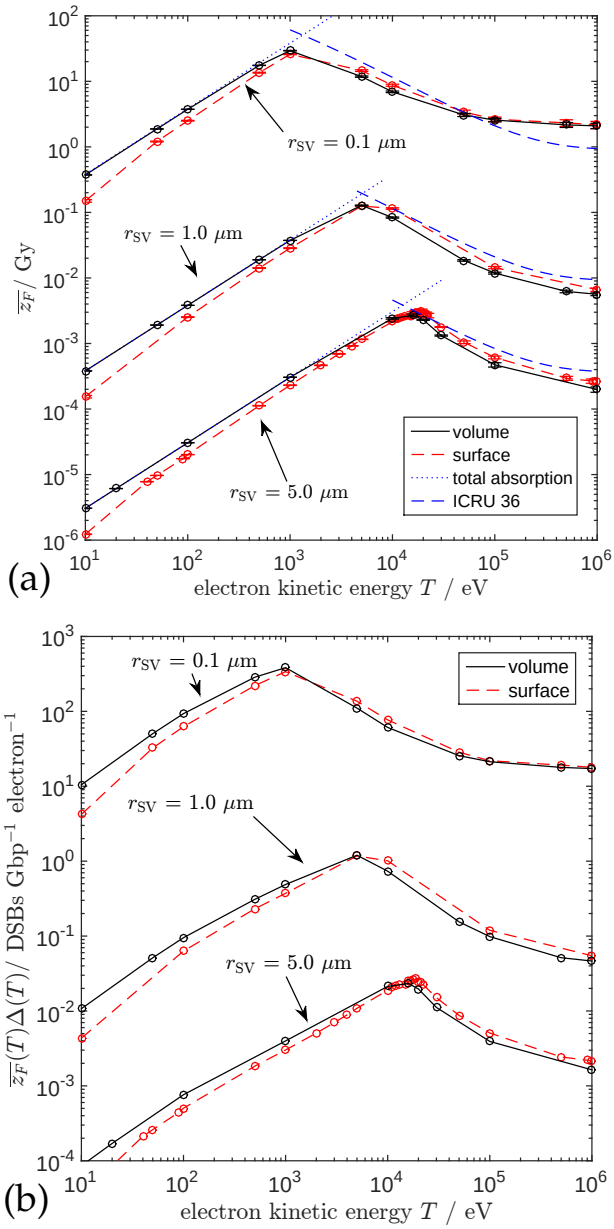

Figure 6–5: (a) $\overline{z_F}(T)$ for 0.1, 1.0, and 5.0 μm spheres irradiated by monoenergetic electrons from 10 to 10^6 eV. Electrons were either started uniformly in the volume (solid black lines) or on the surface (red dashed lines) of the SV. (b) Absolute DSB yield per Gbp and per incident electron.

The multiscale model RBE_{DSB} was in good agreement with previously published values for 80 and 125 kVp CBCT beams.[71] In particular, there was no change in RBE_{DSB} between the two kVp settings. Additionally, insignificant differences in DSB yields were found between the surface and 5 cm depth for those fields. For the 50 kVp intra-operative source, DSB yields were 35-38% lower than previously reported.[73] For the Ir-192 source, we found an RBE_{DSB} of 1.04-1.06, increasing further away from the source. Generally, good agreement was obtained for all other brachytherapy sources.

Table 6–3: DSB yields and RBE_{DSB} using the multiscale methodology for clinical photon fields.

source	depth	Δ_{msm}	RBE_{DSB}	
	cm	DSB Gy^{-1} Gbp^{-1}	multiscale model	literature
Co-60	5.0	8.36	1	1[71]
Co-60	15.0	8.36	1	1[71]
CBCT 125 kVp	0.0	9.49	1.14	1.11[71]
CBCT 125 kVp	5.0	9.49	1.14	1.08[71]
CBCT 80 kVp	0.0	9.60	1.15	1.11[71]
CBCT 80 kVp	5.0	9.69	1.16	1.08[71]
IORT 50 kVp	0.0	9.72	1.16	1.54[73]
IORT 50 kVp	1.0	9.65	1.15	1.54[73]
IORT 50 kVp	2.0	9.51	1.14	1.53[73]
Ir-192	0.5	8.67	1.04	1[70]
Ir-192	5.0	8.68	1.04	-
Ir-192	10.0	8.72	1.04	-
Ir-192	20.0	8.87	1.06	-
Yb-169	0.5	9.35	1.12	1.06[70]
I-125	0.5	9.96	1.19	1.16[70]

6.4 Discussion

In this work, a multiscale method for simulating the event-by-event electron spectra in a cell nucleus is presented and used to estimate $RBE_{DSB}$for various photon fields. Compared to previous simulations, the effect of the surrounding of the cell nucleus are incorporated in the DNA damage estimates. The novelty of the method resides in the use of the photon interaction spectrum instead of the photon fluence, the generation of the incident electron fluence and it's effect on radiation damage, and the application of the associated volume concept for acceleration of the generation of electron spectra.

In several previous studies, the fluence instead of the $\Phi_i(h\nu)$ was used to generate electron tracks.[73,134,137] However, using fluence, which describes a passage through a surface makes little sense when calculating a primary recoil spectrum, which inherently assumes an interaction. The photon interaction spectrum only considers the photon that generate electrons that can affect the SV at the given position in the phantom. As seen in figure 6–4, the spectra are different and have an impact on the calculated RBE_{DSB}.

The Φ_{inc} and Φ_{pr} spectra are two distinct electron sources with different spectral characteristics dependant on the beam/phantom size/shape, the SV radius, and its position in the phantom. In single-cell irradiations (where the SV is located in vacuum), the dose and DNA damage exclusively comes from Φ_{pr}. However, when a phantom surrounds the SV Φ_{inc} becomes dominant. In previous studies,[73] the effects of Φ_{inc} were completely neglected and the RBE_{DSB} was estimated for only the primary recoil electrons, generated uniformly in the SV. Because Φ_{inc} is

dominant at clinical energies, it is relevant to simulate it using event-by-event electron tracking down to below ionization potential energies. For instance, the simulated DSB yield for the Co-60 photon source in the single-cell geometry was 4% higher than in a phantom, due to the lack of accounting for Φ_{inc}.

The use of the track's AV to accelerate dose deposition inside the SV, with proper use of the w_{ed} weight, is not new in itself.[138,139,145] However, to our knowledge, the w_{av} component was previously ignored. Villegas et al.[134] voxelized each track and scored the specific energy in each non-zero dose voxel. Such a method, approximates w_{av} by the number of non-zero dose voxels in each track. This approximation is further improved by using the AV as a weight, and sampling an equal number of SV positions for each track. Even in the case of monoenergetic electrons, each track's AV, although not drastically different from one another, should be used for consistency.

Additionally, in several previous studies,[73,146] equation 6.1 was weighted with the electron fluence rather than the specific energy. For MV photon fields this has little to now effect, because the DSB yields are almost invariant with energy in that range.[146] However, at lower energies, the DSB yields are significantly affected.[71] Notably, for the miniature x-ray source at 50 kVp, the $RBE_{DSB}$is expected to be in the 1.15-1.30 range.[137] In the study by White et al.,[73] the authors acknowledged that the RBE_{DSB} around 1.4-1.6 they obtained is high, and they attributed this, in part, to the simulation of Auger electrons. Instead, the multiscale model presented in this work suggests that the lack of accounting for the specific energy of the electrons,

coupled with the omission of the Φ_{inc} component, and the use of the photon fluence instead of the interaction spectrum artificially increased the RBE_{DSB}(see table 6–3).

In previous work, the reciprocity theorem was used to decrease simulation time in clinical photon beams.[71,76] Instead of irradiating a 5 μm SV with a large beam, a semi-infinite slab 5 μm thick was irradiated by a pencil beam. This technique is well adapted for DSB tally directly in the MC code.[31,71] However, the reciprocity theorem has some limits. Dose from electrons generated parallel to the semi-infinite dimension is artificially increased compared to the spherical case. Additionally, the exact scattering conditions of a given beam size r_b in a given phantom of size s_p are not accurately preserved.

The multiscale method could be applied to arbitrarily large phantoms and beams for energies below 1 MeV. This energy threshold could be extended to higher photon energy if corresponding event-by-event ionization and excitation cross-sections were available in Geant4-DNA. In that case, activation of condensed history electron transport to simulate bremsstrahlung would probably also be required in the $\Phi_i(h\nu)$ generation step.

6.5 Conclusion

A multiscale methodology to obtain the incident and primary recoil electron spectra inside scoring volumes irradiated by clinical photon beams is presented. It uses event-by-event simulations to estimate the DSB yields and RBE_{DSB} differences between the various spectra. The results show that the primary recoil electron in the cell have little influence on DNA damage - incident electron have a dominant effect and need to be accounted for. An approximate RBE_{DSB} of 1.15 was found for

the 50 kVp miniature x-rays source and 80-125 kVp CBCT fields, invariant with depth or energy. The RBE_{DSB} for several isotopes ranged from 1.04-1.06 for Ir-192 to 1.19 for I-125, in good agreement with previous results. If RBE_{DSB} information is to be included in any way at the clinical planning or assessment stage for photons, a better standardization of the simulation techniques of RBE_{DSB} is still needed.

6.6 Acknowledgments

We sincerely thank Dr. Robert Stewart, Dr. Charles Kirkby, and Mr. Peter Watson for providing us with the MCDS monoenergetic DSB yields, the CBCT and miniature X-ray primary spectra, and for fruitful discussions. We would like to acknowledge support from the CREATE Medical Physics Research Training Network grant (number 432290) of the Natural Sciences and Engineering Research Council (NSERC).

CHAPTER 7
Conclusion

7.1 Summary

Numerical modelling of DNA damage using event-by-event MC methods was investigated throughout this thesis. More specifically, improvements to methodologies for simulating DNA SSB and DSB cluster yields were presented for photon, electron, proton and other light ions beams. The RBE_{DSB} or RBE_{DSB}^{direct} for these particles at various energies were studied extensively, and inaccuracies in conventional methodologies were reported and corrected.

In Chapter 4, the effects of user-defined parameters and algorithms on the computed SSB and DSB cluster yields of monoenergetic electron beams was investigated. This was achieved by tracking electrons in an event-by-event mode through a detailed geometrical DNA model comprised of millions of DNA nucleosomes. It was found that simple details in the implementation of search and classification algorithms may lead to large discrepancies in the absolute DNA damage yields. Complementarily, it was shown that some previously obtained results using pencil and plane beams were not consistent with the expected trends of SSB and DSB cluster yields, because CPE was not attained in the region containing the DNA model. Additionally, SB definitions and classification algorithms and the overall terminology used need further standardization.

Chapter 5 further developed the geometrical DNA model to fully include the entire genome in a 5 μm cell nucleus. In addition, a novel track overlay algorithm based on a pregenerated track library was presented in order to decrease simulation time while conserving CPE laterally and longitudinally in the cell nucleus. Conservation of CPE was previously shown to be necessary when comparing DNA damage yields for different energies. The methodology presented in this chapter was used to obtain the RBE_{DSB}^{direct} for various proton and other light ion beams with respect to ^{60}Co. It was found that the RBE_{DSB}^{direct} increased with LET for protons and that it was higher for protons than other higher atomic number ions of equal LET. Moreover, it was found that the ratio of SSB to DSB clusters composed of a given number of SBs is invariant with particle type or kinetic energy. This ratio seems to be a characteristic of the DNA model itself, and it suggests that additional damage is more probable on the same strand than on the opposite strand.

The track overlay methodology was adapted in Chapter 6 to compute the spectra of electrons incident and created inside a cell irradiated by clinical photon beams. These spectra contain all the information on average radiation type experienced by a cell in CPE and under the full influence of phantom scatter. They can be used as inputs to an MCTS and geometrical DNA model approach for the calculation of DNA damage. However, in this chapter, they were instead put into a probabilistic quasi-phenomenological code that estimates RBE_{DSB}. This approach was validated against a brute force approach in limited size phantoms, and it was compared to conventional methodology in a single-cell setup, neglecting phantom scatter. The RBE_{DSB} was computed for CBCT fields of 80 and 125 kVp, for a 50

kVp miniature EBT x-ray source and several brachytherapy isotopes. Particularly, for the miniature EBT x-ray source, the computed RBE_{DSB} was radically different from previous estimates. Nevertheless, it was consistent with the expected values based on experiments with similar spectra.

Taken as a whole, Chapters 4–6 consistently improved on the conventional methodologies used in MCTS simulations of radiation induced DNA damage. The algorithms for overlaying tracks on the DNA model to create CPE in the regions of interest were adapted for large-scale MC simulations.

7.2 Future perspectives

This thesis answered many questions regarding numerical models for radiation-induced DNA damage. At the same time, however, the data provided welcomes new questions and research avenues.

It was found that the ratios of SSB to DSB clusters are independent of the energy or the type of the incident particles. Experimental measurements of these ratios could help in developing more adequate SB creation and classification algorithms. Additionally, probabilistic models based on equal-hit probability on both strands should be modified to take this new information into account.

Although DSB clusters are widely accepted as the main precursor of cell death, it is neither clear nor trivial to relate the DSB cluster yield to a cell survival outcome. Recent studies on DNA repair predict that the complexity of the DSB clusters define which repair pathway will be undertaken, if any. Simple DSB clusters tend to be repaired by NHEJ while more complex ones are repaired by homologous recombination (HR), which is a much longer process. These repair mechanisms with

their inherent miss-repair rates need to be integrated into the simulation. Besides DSB clusters, other forms of genomic damage such as chromosome aberrations, mutations, and creation of short DNA fragments can be observed experimentally and simulated with MCTS codes.

Standardization of the definitions used in the field is still needed. For instance, the DNA *lesions* should be reserved to the elemental damage types, namely all sorts of SBs and BD. The word *clusters* should refer to local groups of lesions that can exclusively be DSB clusters, SSB clusters, or BD clusters. The adjectives *simple* and *complex* should be used with respect to clusters of lesions only and depict the number of lesions composing a cluster, according to a predefined threshold. Moreover, details in the definition of the SB and in the classification algorithms should also be clearly stated and respected. For instance, two directly adjacent SBs in a SSB cluster are sometimes only counted as one SB (because experimentally there would be no fragmentation). In other work, SBs can be part of more than one cluster at the same time. These small differences have a noticeable impact on the published SSB and DSB cluster yields and complicate intercomparisons between simulation codes.

Finally, the overlaying technique presented in this work could be easily adapted for computation of ICSDs. This purely physical metric is an inherent characteristic of a radiation field, just as LET or lineal energy. It was recently proposed that biologically relevant absolute dosimetry could eventually replace conventional absolute dosimetry for protons and light ions.[33] RBE varies as a function of particle type and energy, even for the same absorbed doses. A large

portion of these variations are simply caused by differences in the patterns of ionizations of these radiation modalities. One could think of separating the relative biological effect into a contribution from the ICSDs and a contribution from all other factors (cell type, oxygen concentration, genetics, etc.) The track overlaying methodology could score the ICSDs for various clinical situations using proton and light ion beams in a fast and efficient way. Several publications discussed how the characteristics of these distributions can be used to estimate the biological effects of an increased clustering, but further efforts are needed for a complete clinical implementation.

REFERENCES

[1] Piotr Pater, Jan Seuntjens, Issam El Naqa, and Mario A Bernal. On the consistency of Monte Carlo track structure DNA damage simulations. *Medical Physics*, 41(12):121708, December 2014.

[2] Piotr Pater, Gloria Bäckstöm, Fernanda Villegas, Anders Ahnesjö, Shirin A. Enger, Jan Seuntjens, and Issam El Naqa. Proton and light ion RBE for the induction of direct DNA double strand breaks. *Medical Physics*, 43(5):2131–2140, 2016.

[3] Issam El Naqa, Piotr Pater, and Jan Seuntjens. Monte Carlo role in radiobiological modelling of radiotherapy outcomes. *Physics in Medicine and Biology*, 57(11):R75–R97, May 2012.

[4] Canadian Cancer Society's Advisory Committee on Cancer Statistics. *Canadian Cancer Statistics 2015*. Canadian Cancer Society, Toronto, ON, 2015.

[5] American Cancer Society. *Cancer Treatment and Survivorship Facts & Figures 2014-2015*. American Cancer Society, Atlanta, USA, 2014.

[6] Kienböck R. Über die einwirkung des röntgen-lichtes auf die haut. *Wiener Klinische Wochenschrift*, 13:1153–1166, 1900.

[7] Joachim Widder. The origins of radiotherapy: Discovery of biological effects of X-rays by Freund in 1897, Kienböck's crucial experiments in 1900 and still it is the dose. *Radiotherapy and Oncology*, 112(1):150–152, 2014.

[8] Task Group 1 of the Professional Information and Clinical Relations Committee. *The role of a physicist in radiation oncology*. American Association of Medical Physicists, 1993.

[9] Weidong Wang and Jinyi Lang. Strategies to optimize radiotherapy based on biological responses of tumor and normal tissue (Review). *Experimental and Therapeutic Medicine*, 4(2):175–180, 2012.

[10] J A Bonnell. ICRU Report 19. Radiation Quantities and Units. *British Journal of Industrial Medicine*, 29(4):464, oct 1972.

[11] Richard Bernhard Richardson and Mary-ellen Harper. Mitochondrial stress controls the radiosensitivity of the oxygen effect: Implications for radiotherapy. *Oncotarget*, 7(April), 2015.

[12] L. Pray. Discovery of dna structure and function: Watson and crick. *Nature Education*, 1(1), 2008.

[13] H Nikjoo, D. E. Charlton, and D. T. Goodhead. Monte Carlo track structure studies of energy deposition and calculation of initial DSB and RBE. *Advances in Space Research*, 14(10):161–180, October 1994.

[14] Reza Taleei, Peter M. Girard, and Hooshang Nikjoo. {DSB} repair model for mammalian cells in early s and {G1} phases of the cell cycle: Application to damage induced by ionizing radiation of different quality. *Mutation Research/Genetic Toxicology and Environmental Mutagenesis*, 779:5 – 14, 2015.

[15] H Paganetti. Relative biological effectiveness (RBE) values for proton beam therapy. Variations as a function of biological endpoint, dose, and linear energy transfer. *Physics in Medicine and Biology*, 2014.

[16] J. F. Ward. Biochemistry of DNA lesions. *Radiation Research*, 104(2s):S103–S111, 1985.

[17] Pavel Kundrát and Robert D Stewart. On the biophysical interpretation of lethal DNA lesions induced by ionising radiation. *Radiation Protection Dosimetry*, 122(1–4):169–172, 2006.

[18] H Nikjoo, P O Neill, W E Wilson, and D T Goodhead. Computational Approach for Determining the Spectrum of DNA Damage Induced by Ionizing Radiation. *Radiation Research*, 156(5):577–583, 2001.

[19] Chun S.Y. Badavi F.F. Townsend L.W. Lamkin S.L Wilson, J.W. *HZETRN: A heavy ion and nucleon transport code for space radiations. NASA TP-3146*. NASA Langley Research Center, 1991.

[20] Koji Niita, Tatsuhiko Sato, Hiroshi Iwase, Hiroyuki Nose, Hiroshi Nakashima, and Lembit Sihver. Phits–a particle and heavy ion transport code system. *Radiation Measurements*, 41(9âĂŞ10):1080–1090, 2006. Space Radiation Transport, Shielding, and Risk Assessment Models.

[21] Robert Katz. The parameter-free track structure model of scholz and kraft for heavy-ion cross sections. *Radiation Research*, 160(6):724–728, 2003.

[22] Francesc Salvat, José M Fernández-Varea, and Josep Sempau. PENELOPE-2008: A code system for Monte Carlo simulation of electron and photon transport. In *Workshop Proceedings 30 June-3 July*, page 7, Barcelona, Spain, 2006.

[23] S Agostinelli, John Allison, K Amako, J. Apostolakis, H Araujo, P Arce, M Asai, D Axen, S Banerjee, and G Barrand. GEANT4—a simulation toolkit. *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3):250–303, 2003.

[24] S Incerti, A Ivanchenko, M Karamitros, A Mantero, P Moretto, H N Tran, B Mascialino, C Champion, V N Ivanchenko, M A Bernal, Z Francis, C Villagrasa, G Baldacchino, P Guèye, R Capra, P Nieminen, and C Zacharatou. Comparison of GEANT4 very low energy cross section models with experimental data in water. *Medical Physics*, 37(9):4692, 2010.

[25] W Friedland, P Jacob, H G Paretzke, M Merzagora, and A Ottolenghi. Simulation of DNA fragment distributions after irradiation with photons. *Radiation and Environmental Biophysics*, 38(1):39–47, 1999.

[26] M.J. Berger. *Monte-Carlo calculations of the penetration and diffusion of fast charged particles.* Academic Press, 1963.

[27] J E Turner, J L Magee, H A Wright, A Chatterjee, R N Hamm, and R H Ritchie. Physical and Chemical Development of Electron Tracks in Liquid Water. *Radiation Research*, 96(3):437–449, 1983.

[28] V A Semenenko and R D Stewart. A fast Monte Carlo algorithm to simulate the spectrum of DNA damages formed by ionizing radiation. *Radiation research*, 161(4):451–457, apr 2004.

[29] V A Semenenko and R D Stewart. Fast Monte Carlo simulation of DNA damage formed by electrons and light ions. *Physics in Medicine and Biology*, 51(7):1693–1706, 2006.

[30] Robert D Stewart, Victor K Yu, Alexandros G Georgakilas, Constantinos Koumenis, Joo Han Park, and David J Carlson. Effects of Radiation Quality and Oxygen on Clustered DNA Lesions and Cell Death. *Radiation Research*, 176:587–602, 2011.

[31] Robert D Stewart, Seth W Streitmatter, David C Argento, Charles Kirkby, John T Goorley, Greg Moffitt, Tatjana Jevremovic, and George A Sandison. Rapid MCNP simulation of DNA double strand break (DSB) relative biological effectiveness (RBE) for photons, neutrons, and light ions. *Physics in medicine and biology*, 60(21):8249–8274, November 2015.

[32] G Garty, R Schulte, S Shchemelinin, C Leloup, G Assaf, A Breskin, R Chechik, V Bashkirov, J Milligan, and B Grosswendt. A nanodosimetric model of radiation-induced clustered DNA damage yields. *Physics in Medicine and Biology*, 55(3):761–781, January 2010.

[33] H Palmans, H Rabus, A L Belchior, M U Bug, S Galer, U Giesen, G Gonon, G Gruel, G Hilgers, D Moro, H Nettelbeck, M Pinto, A Pola, S Pszona, G Schettino, P H G Sharpe, P Teles, C Villagrasa, and J J Wilkens. Future development of biologically relevant dosimetry. *The British Journal of Radiology*, 88(1045):20140392, January 2015.

[34] Werner Friedland, Peter Jacob, Herwig G Paretzke, and Tobias Stork. Monte Carlo Simulation of the Production of Short DNA Fragments by Low-Linear Energy Transfer Radiation Using Higher-Order DNA Models. *Radiation Research*, 150(2):170, August 1998.

[35] Rowan M Thomson and Iwan Kawrakow. On the Monte Carlo simulation of electron transport in the sub-1 keV energy range. *Medical Physics*, 38(8):4531, 2011.

[36] D Liljequist and H Nikjoo. On the validity of trajectory methods for calculating the transport of very low energy (<1 keV) electrons in liquids and amorphous media. *Radiation Physics and Chemistry*, 99:45–52, June 2014.

[37] J Allison, K Amako, J Apostolakis, H Araujo, P Arce Dubois, M Asai, G Barrand, R Capra, S Chauvie, R Chytracek, G A P Cirrone, G Cooperman, G Cosmo, G Cuttone, G G Daquino, M Donszelmann, M Dressel, G Folger, F Foppiano, J Generowicz, V Grichine, S Guatelli, P Gumplinger, A Heikkinen, I Hrivnacova, A Howard, S Incerti, V Ivanchenko, T Johnson, F Jones, T Koi, R Kokoulin, M Kossov, H Kurashige, V Lara, S Larsson, F Lei, O Link, F Longo, M Maire, A Mantero, B Mascialino, I McLaren, P Mendez Lorenzo, K Minamimoto, K Murakami, P Nieminen, L Pandola, S Parlati, L Peralta, J Perl, A Pfeiffer, M G Pia, A Ribon, P Rodrigues, G Russo, S Sadilov, G Santin, T Sasaki, D Smith, N Starkov, S Tanaka, E Tcherniaev, B Tome, A Trindade, P Truscott, L Urban,

M Verderi, A Walkden, J P Wellisch, D C Williams, D Wright, and H Yoshida. Geant4 developments and applications. *Ieee Transactions on Nuclear Science*, 53(1):270–278, 2006.

[38] M A Bernal, M C Bordage, J M C Brown, M Davídková, E Delage, Z El Bitar, S A Enger, Z Francis, S Guatelli, V N Ivanchenko, M Karamitros, I Kyriakou, L Maigne, S Meylan, K Murakami, S Okada, H Payno, Y Perrot, I Petrovic, Q T Pham, A Ristic-Fira, T Sasaki, V Štěpán, H N Tran, C Villagrasa, and S Incerti. Track structure modeling in liquid water: A review of the Geant4-DNA very low energy extension of the Geant4 Monte Carlo simulation toolkit. *Physica Medica*, 31(8):861–74, 2015.

[39] C Villagrasa, Z Francis, and S Incerti. Physical models implemented in the GEANT4-DNA extension of the GEANT-4 toolkit for calculating initial radiation damage at the molecular level. *Radiation Protection Dosimetry*, 143(2-4):214–218, 2011.

[40] M A Bernal, C E deAlmeida, C Sampaio, S Incerti, C Champion, and P Nieminen. The invariance of the total direct DNA strand break yield. *Medical Physics*, 38(7):4147–4153, 2011.

[41] M. Karamitros, S. Luan, M.A. Bernal, J. Allison, G. Baldacchino, M. Davidkova, Z. Francis, W. Friedland, V. Ivantchenko, A. Ivantchenko, A. Mantero, P. Nieminem, G. Santin, H.N. Tran, V. Stepan, and S. Incerti. Diffusion-controlled reactions modeling in geant4-dna. *Journal of Computational Physics*, 274:841 – 882, 2014.

[42] DE Cullen, JH Hubbel, and L Kyssel. *EPDL97: The evaluated photon data library, '97 version*. Lawrence Livermore National Laboratory, 1997.

[43] D Emfietzoglou, G Papamichael, K Kostarelos, and M Moscovitch. A Monte Carlo track structure code for electrons (approximately 10 eV-10 keV) and protons (approximately 0.3-10 MeV) in water: partitioning of energy and collision events. *Physics in medicine and biology*, 45(11):3171–3194, 2000.

[44] D J Brenner and M Zaider. A computationally convenient parameterisation of experimental angular distributions of low energy electrons elastically scattered off water vapour. *Physics in Medicine and Biology*, 29(4):443–447, 1983.

[45] C Champion, S Incerti, H. Aouchiche, and D. Oubaziz. A free-parameter theoretical model for describing the electron elastic scattering in water in the Geant4 toolkit. *Radiation Physics and Chemistry*, 78(9):745–750, 2009.

[46] J Neuefeind, C J Benmore, B Tomberli, and P A Egelstaff. Experimental determination of the electron density of liquid H2O and D2O. *Journal of Physics: Condensed Matter*, 14(23):L429, 2002.

[47] R. D. Birkhoff J. M. Heller, R. N. Hamm and L. R. Painter. Collective oscillation in liquid water. *Journal of chemical physics*, 60:3483âĂŞ–3486, 1974.

[48] D. Emfietzoglou. Inelastic cross-sections for electron transport in liquid water: a comparison of dielectric models. *Radiation Physics and Chemistry*, 66(6):373 – 385, 2003.

[49] D Emfietzoglou and H Nikjoo. The effect of model approximations on single-collision distributions of low-energy electrons in liquid water. *Radiat Res*, 163(1):98–111, 2005.

[50] M. Michaud and L. Sanche. Absolute vibrational excitation cross sections for slow-electron (1 – 18 ev) scattering in solid h_2o. *Phys. Rev. A*, 36:4684–4699, Nov 1987.

[51] P. Rowntree, L. Parenteau, and L. Sanche. Electron stimulated desorption via dissociative attachment in amorphous H2O. *The Journal of Chemical Physics*, 94(12):8570, 1991.

[52] S Incerti, C Champion, H N Tran, M Karamitros, M. Bernal, Z Francis, V Ivanchenko, A Mantero, and Members of the Geant4-DNA collaboration. Energy deposition in small-scale targets of liquid water using the very low energy electromagnetic physics processes of the Geant4 toolkit. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 306(C):158–164, July 2013.

[53] M U Bug, E Gargioni, S Guatelli, S Incerti, H Rabus, R Schulte, and A B Rosenfeld. Effect of a magnetic field on the track structure of low-energy electrons: a Monte Carlo study. *The European Physical Journal D*, 60(1):85–92, 2010.

[54] Ziad Francis, Carmen Villagrasa, and Isabelle Clairand. Simulation of DNA damage clustering after proton irradiation using an adapted DBSCAN algorithm. *Computer methods and programs in biomedicine*, 101(3):265–270, March 2011.

[55] Daniele Alloni, Alessandro Campa, Werner Friedland, Luca Mariotti, and Andrea Ottolenghi. Track structure, radiation quality and initial radiobiological events: Considerations based on the PARTRAC code experience. *International Journal of Radiation Biology*, 88(1-2):77–86, 2012.

[56] Werner Friedland, Pavel Kundrát, and Peter Jacob. Stochastic modelling of DSB repair after photon and ion irradiation. *International Journal of Radiation Biology*, 88(1-2):129–136, 2012.

[57] W. Friedland and P. Kundrát. Chromosome aberration model combining radiation tracks, chromatin structure, DSB repair and chromatin mobility. *Radiation Protection Dosimetry*, 166(1-4):71–74, 2015.

[58] Michael Dingfelder, Detlev Hantke, Mitio Inokuti, and Herwig G. Paretzke. Electron inelastic-scattering cross sections in liquid water. *Radiation Physics and Chemistry*, 53(1):1–18, 1999.

[59] M Dingfelder, R H Ritchie, J E Turner, W Friedland, H G Paretzke, and R N Hamm. Comparisons of Calculations with PARTRAC and NOREC: Transport of Electrons in Liquid Water. *Radiation Research*, 169(5):584–594, 2008.

[60] H W Barkas. *Nuclear research emulsions*. Academic Press, 1963.

[61] Marion U Bug, Elisabetta Gargioni, Heidi Nettelbeck, Woon Yong Baek, Gerhard Hilgers, Anatoly B Rosenfeld, and Hans Rabus. Ionization cross section data of nitrogen, methane, and propane for light ions and electrons and their suitability for use in track structure simulations. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 88(4):43308, oct 2013.

[62] Werner Friedland, Herwig G Paretzke, Francesca Ballarini, Andrea Ottolenghi, Gregor Kreth, and Christoph Cremer. First steps towards systems radiation biology studies concerned with DNA and chromosome structure within living cells. *Radiation and Environmental Biophysics*, 47(1):49–61, January 2008.

[63] Werner Friedland, Peter Jacob, Philipp Bernhardt, Herwig G Paretzke, and Michael Dingfelder. Simulation of DNA Damage after Proton Irradiation. *Radiation Research*, 159(3):401–410, 2002.

[64] Werner Friedland, Peter Jacob, Herwig G Paretzke, Andrea Ottolenghi, Francesca Ballarini, and Marco Liotta. Simulation of light ion induced DNA damage patterns. *Radiation protection Dosimetry*, 122(1-4):116–120, 2006.

[65] Badia Boudaïffa, Darel Hunting, and P Cloutier. Induction of single- and double-strand breaks in plasmid DNA by 100–1500 eV electrons. *International Journal of Radiation Biology*, 76(9):1209–1221, January 2000.

[66] K M Prise, M Folkard, B D Michael, B Vojnovic, B Brocklehurst, a Hopkirk, and I H Munro. Critical energies for SSB and DSB induction in plasmid DNA by low-energy photons: action spectra for strand-break induction in plasmid DNA irradiated in vacuum. *International journal of radiation biology*, 76(7):881–890, 2000.

[67] P Bernhardt, W Friedland, P Jacob, and H G Paretzke. Modeling of ultrasoft X-ray induced DNA damage using structured higher order DNA targets. *International Journal of Mass Spectrometry*, 223-224:579–597, 2003.

[68] A Campa, F Ballarini, M Belli, R Cherubini, V Dini, G Esposito, W Friedland, S Gerardi, S Molinelli, A Ottolenghi, H Paretzke, G Simone, and M A Tabocchini. DNA DSB induced in human cells by charged particles and gamma rays: experimental results and theoretical approaches. *International journal of radiation biology*, 81(11):841–854, November 2005.

[69] Werner Friedland, Michael Dingfelder, Peter Jacob, and Herwig G Paretzke. Calculated DNA double-strand break and fragmentation yields after irradiation with He ions. *Radiation Physics and Chemistry*, 72:279–286, 2005.

[70] Y Hsiao and R D Stewart. Monte Carlo simulation of DNA damage induction by x-rays and selected radioisotopes. *Physics in medicine and biology*, 53(1):233–244, jan 2008.

[71] C Kirkby, E Ghasroddashti, Y Poirier, M Tambasco, and R D Stewart. RBE of kV CBCT radiation determined by Monte Carlo DNA damage simulations. *Physics in medicine and biology*, 58(16):5693–5704, aug 2013.

[72] Brigitte Reniers, Derek Liu, Thomas Rusch, and Frank Verhaegen. Calculation of relative biological effectiveness of a low-energy electronic brachytherapy source. *Physics in Medicine and Biology*, 53(24):7125–7135, 2008.

[73] Shane A White, Brigitte Reniers, Evelyn E C de Jong, Thomas Rusch, and Frank Verhaegen. A comparison of the relative biological effectiveness of low energy electronic brachytherapy sources in breast tissue: a Monte Carlo study. *Physics in medicine and biology*, 61(1):383–399, jan 2016.

[74] Ya-Yun Hsiao, Tzu-Hsiang Hung, Shu-Ju Tu, and Chuan-Jong Tung. Fast Monte Carlo simulation of DNA damage induction by Auger-electron emission. *International Journal of Radiation Biology*, pages 1–9, 2014.

[75] Chun-Ching Wang, Y Hsiao, Chung-Chi Lee, Tsi-Chian Chao, Chun-Chieh Wang, and Chuan-Jong Tung. Monte Carlo simulations of therapeutic proton beams for relative biological effectiveness of double-strand break. *International journal of radiation biology*, 88(1-2):158–163, jan 2012.

[76] R. M. Sternheimer. Multiple Scattering Correction for Counter Experiments. *Review of Scientific Instruments*, 25(11):1070, 1954.

[77] H. E. Johns and J. R. Cunningham. *The Physics of Radiology*. C.C. Thomas, 1983.

[78] T. M. Jenkins, W. R. Nelson, and A. Rindi. *Monte Carlo Transport of Electrons and Photons*. Springer US, 1988.

[79] ICRU 2011. Fundamental quantities and units for ionizing radiation (revised). *Journal of the ICRU*, 11, 2011.

[80] Lech Papiez and Jerry J Battista. Radiance and particle fluence. *Physics in medicine and biology*, 39(6):1053–62, 1994.

[81] J.P. Seuntjens B.R.B. Walters D.W.O. Rogers, I. Kawrakow and E. Mainegra-Hing. *NRC User Codes for EGSnrc*. Ionizing Radiation Standards, National Research Council Canada, 2016.

[82] Roland B Hawkins. A Microdosimetric-Kinetic Model for the Effect of Non-Poisson Distribution of Lethal Lesions on the Variation of RBE with LET. *Radiation Research*, 160(1):61–69, July 2003.

[83] Roland B Hawkins and Taku Inaniwa. A Microdosimetric-Kinetic Model for Cell Killing by Protracted Continuous Irradiation Including Dependence on

LET I: Repair in Cultured Mammalian Cells. *Radiation Research*, 180(6):584–594, December 2013.

[84] Thomas Friedrich, Marco Durante, and Michael Scholz. Modeling Cell Survival after Photon Irradiation Based on Double-Strand Break Clustering in Megabase Pair Chromatin Loops. *Radiation Research*, 178(5):385–394, November 2012.

[85] H Nikjoo, S Uehara, D Emfietzoglou, and F. A. Cucinotta. Track-structure codes in radiation research. *Radiation Measurements*, 41(9-10):1052–1074, October 2006.

[86] M A Bernal and J A Liendo. An investigation on the capabilities of the PENELOPE MC code in nanodosimetry. *Medical Physics*, 36(2):620–625, February 2009.

[87] H Nikjoo, D. T. Goodhead, D. E. Charlton, and H G Paretzke. Energy deposition in small cylindrical targets by ultrasoft x-rays. *Phys Med Biol*, 34(6):691–705, June 1989.

[88] W Friedland, P Bernhardt, P Jacob, and H G Paretzke. Simulation of DNA damage after proton and low LET irradiation. *Radiation Protection Dosimetry*, 99(1-4):99–102, 2002.

[89] M A Bernal, D Sikansi, F Cavalcante, S Incerti, C Champion, V Ivanchenko, and Z Francis. An atomistic geometrical model of the B-DNA configuration for DNA–radiation interaction simulations. *Computer Physics Communications*, 184(12):2840–2847, December 2013.

[90] H Nikjoo, P O'Neill, M Terrissol, and DT Goodhead. Quantitative modelling of DNA damage using Monte Carlo track structure method. *Radiation and Environmental Biophysics*, 38(1):31–38, April 1999.

[91] William R Holley and Aloke Chatterjee. Clusters of DNA damage induced by ionizing radiation: Formation of short DNA fragments. I. Theoretical modeling. *Radiation Research*, 145(2):188–199, 1996.

[92] Werner Friedland, Michael Dingfelder, Pavel Kundrát, and Peter Jacob. Track structures, DNA targets and radiation effects in the biophysical Monte Carlo simulation code PARTRAC. *Mutation Research*, 711(1-2):28–40, June 2011.

[93] W. E. Wilson and Hooshang Nikjoo. A Monte Carlo code for positive ion track simulation. *Radiation and Environmental Biophysics*, 38(2):97–104, 1999.

[94] José M Fernández-Varea, Gloria González-Muñoz, Mariel E Galassi, Kristin Wiklund, Bengt K Lind, Anders Ahnesjö, and Nina Tilly. Limitations (and merits) of PENELOPE as a track-structure code. *International Journal of Radiation Biology*, 88(1-2):66–70, January 2012.

[95] Jun-Li Li, Chun-Yan Li, Rui Qiu, Cong-Chong Yan, Wen-Zhang Xie, Zhi Zeng, and Chuan-Jong Tung. Comparison of direct DNA strand breaks induced by low energy electrons with different inelastic cross sections. *Nuclear Inst. and Methods in Physics Research*, 311(C):27–36, September 2013.

[96] Hooshang Nikjoo and Peter Girard. A model of the cell nucleus for DNA damage calculations. *International Journal of Radiation Biology*, 88(1-2):87–97, January 2012.

[97] Anil Kumar and Michael D Sevilla. Radiation Effects on DNA: Theoretical Investigations of Electron, Hole and Excitation Pathways to DNA Damage. In Manoj K Shukla and Jerzy Leszczynski, editors, *Radiation Induced Molecular Phenomena in Nucleic Acids*, pages 577–617. Springer Netherlands, Dordrecht, 2008.

[98] C Champion, H Lekadir, M E Galassi, O Fojón, R D Rivarola, and J Hanssen. Theoretical predictions for ionization cross sections of DNA nucleobases impacted by light ions. *Physics in Medicine and Biology*, 55(20):6053–6067, September 2010.

[99] Isabel Abril, Rafael Garcia-Molina, Cristian D Denton, Ioanna Kyriakou, and Dimitris Emfietzoglou. Energy Loss of Hydrogen- and Helium-Ion Beams in DNA: Calculations Based on a Realistic Energy-Loss Function of the Target. *Radiation Research*, 175(2):247–255, February 2011.

[100] Isabel Abril, Cristian D Denton, Pablo de Vera, Ioanna Kyriakou, Dimitris Emfietzoglou, and Rafael Garcia-Molina. Effect of the Bethe surface description on the electronic excitations induced by energetic proton beams in liquid water and DNA. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 268(11-12):1763–1767, June 2010.

[101] Isabel Abril, Rafael Garcia-Molina, Pablo de Vera, Ioanna Kyriakou, and Dimitris Emfietzoglou. Inelastic Collisions of Energetic Protons in Biological Media. In *Advances in Quantum Chemistry*, pages 129–164. Elsevier Inc., 2013.

[102] Zhenyu Tan, Yueyuan Xia, Xiangdong Liu, Mingwen Zhao, Yanju Ji, Feng Li, and Boda Huang. Cross sections of electron inelastic interactions in DNA. *Radiation and Environmental Biophysics*, 43(3):173–182, August 2004.

[103] David Liljequist, Thiansin Liamsuwan, and Hooshang Nikjoo. Elastic scattering cross section models used for Monte Carlo simulation of electron tracks in media of biological and medical interest. *International Journal of Radiation Biology*, 88(1-2):29–37, January 2012.

[104] D Emfietzoglou, I Kyriakou, I Abril, R Garcia-Molina, and H Nikjoo. Inelastic scattering of low-energy electrons in liquid water computed from optical-data models of the Bethe surface. *International Journal of Radiation Biology*, 88(1-2):22–28, January 2012.

[105] Liming Zhang and Zhenyu Tan. A new calculation on spectrum of direct DNA damage induced by low-energy electrons. *Radiation and Environmental Biophysics*, 49(1):15–26, December 2009.

[106] A Akkerman, T Boutboul, A Breskin, R Chechik, and A Gibrekhterman. Low-energy electron transport in alkali halides. *Journal of Applied Physics*, 76(8):4656, 1994.

[107] J Kuhr and H-J Fitting. Monte Carlo simulation of electron emission from solids. *Journal of Electron Spectroscopy and Related Phenomena*, 105(2):257–273, 1999.

[108] Pablo de Vera, Eugene Surdutovich, Isabel Abril, Rafael Garcia-Molina, and Andrey V Solov'yov. Analytical model of ionization and energy deposition by proton beams in subcellular compartments. *The European Physical Journal D*, 68(4):96, April 2014.

[109] Radmila Panajotovic, Frédéric Martin, Pierre Cloutier, Darel Hunting, and Léon Sanche. Effective Cross Sections for Production of Single-Strand Breaks in Plasmid DNA by 0.1 to 4.7 eV Electrons. *Radiation Research*, 165:452–459, 2006.

[110] Badia Boudaïffa, Pierre Cloutier, Darel Hunting, Michael A Huels, and Léon Sanche. Resonant formation of DNA strand breaks by low-energy (3 to 20 eV) electrons. *Science*, 287(5458):1658–1660, 2000.

[111] D. E. Charlton, H Nikjoo, and J L Humm. Calculation of Initial Yields of Single- and Double-strand Breaks in Cell Nuclei from Electrons, Protons and Alpha Particles. *International Journal of Radiation Biology*, 56:1–19, 1989.

[112] D. E. Charlton and J L Humm. A Method of Calculating Initial DNA Strand Breakage Following the Decay of Incorporated 125I. *International Journal of Radiation Biology*, 53(3):353–365, April 1988.

[113] M A Bernal and J A Liendo. Erratum: "An investigation on the capabilities of the PENELOPE MC code in nanodosimetry" [Med. Phys. 36, 620–625 (2009)]. *Medical Physics*, 37(8):4520–4520, 2010.

[114] Alexandros G Georgakilas, Peter O'Neill, and Robert D Stewart. Induction and Repair of Clustered DNA Lesions: What Do We Know So Far? *Radiation Research*, 180(1):100–109, July 2013.

[115] Brita Singers Sørensen, Jens Overgaard, and Niels Bassler. In vitro RBE-LET dependence for multiple particle types. *Acta Oncologica*, 50(6):757–762, August 2011.

[116] D Frankenberg, H J Brede, U J Schrewe, C Steinmetz, M Frankenberg-Schwager, G Kasten, and E Pralle. Induction of DNA double-strand breaks by 1H and 4He lons in primary human skin fibroblasts in the LET range of 8 to 124 keV/microm. *Radiation research*, 151(5):540–549, 1999.

[117] Francesco Tommasino, Thomas Friedrich, Burkhard Jakob, Barbara Meyer, Marco Durante, and Michael Scholz. Induction and Processing of the Radiation-Induced Gamma-H2AX Signal and Its Link to the Underlying Pattern of DSB: A Combined Experimental and Modelling Study. *PLoS ONE*, 10(6):e0129416, 2015.

[118] H M Dang, M J van Goethem, E R van der Graaf, S Brandenburg, R Hoekstra, and T Schlatholter. Plasmid DNA damage by heavy ions at spread-out Bragg peak energies. *The European Physical Journal D*, 60(1):51–58, 2010.

[119] Dalong Pang, James E Rodgers, Barry L Berman, Sergey Chasovskikh, and Anatoly Dritschilo. Spatial distribution of radiation-induced double-strand breaks in plasmid DNA as resolved by atomic force microscopy. *Radiation Research*, 164(6):755–765, 2005.

[120] Larry H Toburen. Challenges in Monte Carlo track structure modelling. *International Journal of Radiation Biology*, 88(1-2):1–9, May 2011.

[121] Thomas S Deisboeck, Zhihui Wang, Paul Macklin, and Vittorio Cristini. Multiscale cancer modeling. *Annual review of biomedical engineering*, 13:127–155, August 2011.

[122] P Lazarakis, M U Bug, E Gargioni, S Guatelli, H Rabus, and A B Rosenfeld. Comparison of nanodosimetric parameters of track structure calculated by the Monte Carlo codes Geant4-DNA and PTra. *Physics in Medicine and Biology*, 57(5):1231–1250, 2012.

[123] A M Kellerer. Chord-Length Distributions and Related Quantities for Spheroids. *Radiation Research*, 98(3):425–437, 1984.

[124] H Nikjoo, P O'Neill, W. E. Wilson, and D. T. Goodhead. Computational Approach for Determining the Spectrum of DNA Damage Induced by Ionizing Radiation. *Radiation Research*, 156(5):577–583, 2001.

[125] J Fulford. *Quantification of complex DNA damage by ionising radiation: an experimental and theoretical approach*. PhD thesis, Brunel University, London, 2000.

[126] George Marsaglia, Wai Wan Tsang, and Jingbo Wang. Evaluating Kolmogorov's Distribution. *Journal of Statistical Software*, 8(18):1–4, November 2003.

[127] Donald F Gatz and Luther Smith. The standard error of a weighted mean concentration–i. bootstrapping vs other methods. *Atmospheric Environment*, 29(11):1185–1193, 1995.

[128] W. Y. Baek, M. Bug, H. Rabus, E. Gargioni, and B. Grosswendt. Differential elastic and total electron scattering cross sections of tetrahydrofuran. *Physical Review A*, 86(3):1–15, 2012.

[129] W. Y. Baek, a. Arndt, M. U. Bug, H. Rabus, and M. Wang. Total electron-scattering cross sections of pyrimidine. *Physical Review A - Atomic, Molecular, and Optical Physics*, 88(3):1–8, 2013.

[130] Jun-Li Li, Chun-Yan Li, Rui Qiu, Cong-Chong Yan, Wen-Zhang Xie, Zhi Zeng, and Chuan-Jong Tung. Comparison of direct DNA strand breaks induced by low energy electrons with different inelastic cross sections. *Nuclear Instruments and Methods in Physics Research Section B: Beam Interactions with Materials and Atoms*, 311:27–36, 2013.

[131] International Human Genome Sequencing Consortium. Finishing the euchromatic sequence of the human genome. *Nature*, 431(7011):931–945, October 2004.

[132] G G Maul and L Deaven. Quantitative determination of nuclear pore complexes in cycling cells with differing DNA content. *The Journal of cell biology*, 73(3):748–760, June 1977.

[133] Munetoshi Maeda, Noriko Usami, and Katsumi Kobayashi. Low-dose hypersensitivity in nucleus-irradiated V79 cells studied with synchrotron X-ray microbeam. *Journal of Radiation Research*, 49(2):171–180, March 2008.

[134] Fernanda Villegas, Nina Tilly, and Anders Ahnesjö. Monte Carlo calculated microdosimetric spread for cell nucleus-sized targets exposed to brachytherapy 125I and 192Ir sources and 60Co cell irradiation. *Physics in Medicine and Biology*, 58(17):6149–6162, August 2013.

[135] S Incerti, P Gillet, P. Barberet, M Bardiès, M A Bernal, M Psaltaki, and The Geant-DNA collaboration. Simulating radial dose of ion tracks in liquid water simulated with Geant4-DNA: A comparative study. *Nuclear Instruments & Methods in Physics Research Section B-Beam Interactions with Materials and Atoms*, 333:92–98, 2014.

[136] Hooshang Nikjoo and Lennart Lindborg. RBE of low energy electrons and photons. *Physics in Medicine and Biology*, 55:R65–R109, 2010.

[137] M A Bernal, C E DeAlmeida, M David, and E Pires. Estimation of the RBE of mammography-quality beams using a combination of a Monte Carlo code with a B-DNA geometrical model. *Physics in Medicine and Biology*, 56(23):7393–7403, 2011.

[138] D.E. Lea. Action of Radiation on Living Cells. *Cambridge University Press*, 1946.

[139] Marco Zaider. Microdosimetry and Katz's Track Structure Theory. *Radiation research*, 124(1):16–22, 1990.

[140] H H Rossi and M Zaider. *Microdosimetry and Its Applications*. Springer, Berlin, Heidelberg, 1996.

[141] C M de Lara, M A Hill, T J Jenner, D Papworth, and P O'Neill. Dependence of the yield of DNA double-strand breaks in Chinese hamster V79-4 cells on the photon energy of ultrasoft X rays. *Radiation research*, 155(3):440–448, 2001.

[142] ICRU 1984. Microdosimetry. *ICRU report 36*, (Bethesda, MD: International Commission on Radiation Units and Measurements), 1984.

[143] Peter G. F. Watson and Jan Seuntjens. A monte carlo model of a miniature low-energy x-ray tube using egsnrc. *Medical Physics*, 41:24, 2014.

[144] Peter G. F. Watson and Jan Seuntjens. Technical Note: Effect of explicit M and N-shell atomic transitions on a low-energy x-ray source. *Medical Physics*, 43(4):1760–1763, 2016.

[145] C S Wuu, P Kliauga, and H I Amols. Microdosimetric evaluation of relative biological effectiveness of Pd-132, I-125, Am245, and Ir-192 brachytherapy sources. *International journal of radiation oncology, biology, physics*, 36(3):689–697, 1996.

[146] A O Ezzati, Y Xiao, M Sohrabpour, and M T Studenski. The effect of energy spectrum change on DNA damage in and out of field in 10-MV clinical photon beams. *Medical & biological engineering & computing*, 53(1):67–75, jan 2015.

Glossary

$OH^{\bullet}$ OH free radical.

RBE_{DSB} relative biological effectiveness for the induction of DNA double-strand breaks.

RBE_{DSB}^{direct} relative biological effectiveness for the induction of direct DNA double-strand breaks.

6N 6-nucleosome structure.

BD base damage.

BioQuaRT Biologically weighted Quantities in Radiation Therapy.

bp base pair.

CBCT cone-beam computed tomography.

cf chromatin fiber.

CPE charged particle equilibrium.

CT computed tomography.

DBSCAN density-based spatial clustering of applications with noise.

DSB double-strand break.

EBRT external beam RT.

EBT electronic brachytherapy.

ED energy deposition.

EPDL97 Evaluated Photon Data Library 1997.

HR homologous recombination.

ICRU international Commission on radiation Units and measurements.

ICSD ionization clusters size distribution.

IORT intraoperative RT.

IR ionizing radiation.

LEE low energy electrons.

LEPR low energy process region.

LET linear energy transfer.

LINAC clinical linear accelerator.

MC Monte Carlo.

MCDS Monte Carlo Damage Simulation.

MCNP Monte Carlo N-Particle.

MCTS Monte Carlo track structure.

NHEJ non-homologous end joining.

NTCP normal tissue complication probability.

PARTRAC PARticle TRACking.

PENELOPE Penetration and ENErgy LOss of Positrons and Electrons.

PET positron emission tomography.

RBE relative biological effectiveness.

ROI region of interest.

RT radiation therapy.

SB strand break.

SEM standard error on the mean.

SP stopping power.

SPB sugar-phosphate backbone.

SSB single-strand break.

SSB+ complex SSB.

TCP tumor control probability.

TRT targeted radionuclide therapy.

TSB total strand break.