Development and application of the Geant4-DNA toolkit for the simulation of radiobiological effects at the sub-cellular scale

Wook Geun Shin

THÈSE EN COTUTELLE PRÉSENTÉE

POUR OBTENIR LE GRADE DE

DOCTEUR

DE L'UNIVERSITÉ DE BORDEAUX ET DE L'UNIVERSITÉ DE YONSEI

ÉCOLE DOCTORALE UNIVERSITÉ DE BORDEAUX

ÉCOLE DOCTORALE UNIVERSITÉ DE YONSEI

SPÉCIALITÉ : Astrophysique, Plasmas, Nucléaire

Par **Wook-Geun SHIN**

Development and application of the Geant4-DNA toolkit for the simulation of radiobiological effects at the sub-cellular scale

Sous la direction de **Sebastien INCERTI**
et de **Chul Hee MIN**

Soutenue le 22 décembre 2020

Membres du jury :

M. INCERTI, Sébastien	Directeur de recherche	CENBG / CNRS	Directeur de thèse
M. MIN, Chul Hee	Professeur associé	Yonsei University	Codirecteur de thèse
M. EL BITAR, Ziad	Directeur de recherche	IPHC / CNRS	Rapporteur
M. CHO, Kihyeon	Professeur	UST / KISTI	Rapporteur
M. TAKASHI, Sasaki	Professeur	KEK	Examinateur
Mme BORDAGE, Marie-Claude	Chargée de recherche	CRCT / CNRS	Examinatrice

Développement et application de Geant4-DNA pour la simulation des effets radiobiologiques à l'échelle sub-cellulaire

Résumé : Prévoir les effets biologiques induits par les rayonnements ionisants est un défi scientifique majeur de la radiobiologie actuelle, en particulier pour essayer de mieux comprendre les effets des faibles doses sur le milieu vivant ainsi que la cancérogénèse. L'approche computationnelle basée sur les codes de simulation des structures de traces dans le milieu biologique par la technique Monte Carlo est aujourd'hui la méthode la plus fiable pour calculer les effets précoces des radiations ionisantes sur l'ADN, la cible cellulaire principale des effets des radiations. Parmi les codes existants, l'extension Geant4-DNA de la boîte à outils généraliste Geant4 est la première entièrement ouverte et librement accessible à la communauté. Geant4-DNA peut simuler non seulement l'étape physique mais aussi les étapes physico-chimique et chimique de la radiolyse de l'eau. Ces étapes peuvent être combinées avec des modèles géométriques simplifiés de l'ADN afin d'évaluer les dommages précoces directs et indirects à l'ADN. Dans cette thèse, je propose (1) d'améliorer dans Geant4-DNA la modélisation de la diffusion élastique des électrons dans l'eau liquide pour simuler plus précisément la distribution spatiale des dépôts d'énergie et des espèces moléculaires. Ensuite, (2) l'étape physico-chimique de la radiolyse de l'eau est également améliorée en se basant sur des approches décrites dans la littérature (modélisation, mesures), cette étape affectant fortement l'étape chimique en modifiant les rendements initiaux et la concentration des espèces. (3) La méthode du temps de réaction indépendant (IRT) est en outre implémentée dans Geant4-DNA afin de réduire le temps de calcul pour simuler la cinétique chimique de la radiolyse de l'eau. Enfin, j'évalue (4) les dommages biologiques induits à l'échelle subcellulaire en utilisant une géométrie de l'ADN cellulaire développée dans une étude précédente, en incluant dans la simulation toutes les améliorations développées au cours de cette thèse, jusqu'à la réparation des dommages précoces. Ces développements sont regroupés au sein d'une chaine de simulation complète destinée aux utilisateurs de Geant4 et de son extension Geant4-DNA.

Mots clés : Radiobiologie, Dommage, ADN, simulation Monte Carlo, Geant4-DNA

Development and application of the Geant4-DNA toolkit for the simulation of radiobiological effects at the sub-cellular scale

Abstract : Predicting the biological effects induced by ionizing radiation is a major scientific challenge of current radiobiology, in particular to try to better understand the effects of low doses on living beings as well as carcinogenesis. The computational approach based on codes to simulate trace structures in the biological medium using the Monte Carlo technique is today the most reliable method to calculate the early effects of ionizing radiation on DNA, the main cellular target of radiation effects. Among the existing codes, the Geant4-DNA extension of the Geant4 general purpose simulation toolkit is the first one fully open and freely available to the community. Geant4-DNA can simulate not only the physical but also the physico-chemical and chemical stages of water radiolysis. These stages can be combined with simplified geometric models of DNA to assess direct and indirect early DNA damage. In this thesis, I propose (1) to improve in Geant4-DNA the modeling of the elastic scattering of electrons in liquid water in order to simulate more precisely the spatial distribution of energy deposits and molecular species. Then, (2) the physico-chemical stage of water radiolysis is also improved based on approaches described in the literature (modeling and measurements), this step strongly affecting the chemical stage by modifying the initial yields and the concentration of species. (3) In addition, the Independent Reaction Time (IRT) method is implemented in Geant4-DNA in order to reduce the computational time to simulate the chemical kinetics of water radiolysis. Finally, I evaluate (4) the biological damage induced at the subcellular scale using a cellular DNA geometry developed in a previous study, including in the simulation all the improvements developed during this thesis, up to the repair of early DNA damage. These developments are grouped in a complete simulation chain for users of the Geant4-DNA extension of Geant4.

Keywords : Radiobiology, Damage, DNA, Monte Carlo simulation, Geant4-DNA

Unité de recherche

UMR 5797, CENBG, 33170 Gradignan, France

Table of contents

Chapter 1 Introduction 1

1.1.Context……..……...2

1.2.Monte Carlo track structure simulation……...……..……4

1.3.The Geant4-DNA project……9

1.3.1. Physical stage……..……...…...10

1.3.2. Pre-chemical and chemical stages……..…….…..14

1.3.3. Geometrical models and DNA damage scoring……...….16

1.4.Purpose of this thesis……18

References……...…19

Chapter 2 Physical stage 28

2.1.Development of a new electron elastic scattering cross-section model for Geant4-DNA using ELSEPA for liquid-phase water……..……....31

2.1.1. Description of potentials and selected options……...…...33

2.1.2. Verification of validity by comparing calculated cross-section with experiments and other cross-section models…….……...…..38

2.1.3. Results for determination of optimal options and parameters……..……40

2.1.4. Plausibility of the electron elastic scattering models……46

2.2. The impact of the ELSEPA electron elastic scattering model on Geant4-DNA simulations……...……...……....48

2.2.1. Determination of optimal angle and energy binning in ELSEPA…….48

2.2.2. Geant4-DNA examples……49

2.2.3. Results for the implementation of the elastic scattering cross-section into Geant4-DNA…….....55

2.2.4. Results for the Geant4-DNA simulations…….59

2.3. Conclusions…….…66

References……..……68

Chapter 3 Pre-chemical stage 77

3.1.Evaluation of the influence of the spatial distribution of molecular species on water radiolysis simulations using Geant4-DNA……..…….…….....78

3.1.1. Principles of Geant4-DNA simulation of water radiolysis…….……79

3.1.2. Geant4-DNA elastic scattering models for electrons in liquid water…….81

3.1.3. Geant4-DNA electron thermalization model……..83

3.1.4. Chemistry parameters……88

3.1.5. The new "chem6" example……90

3.1.6. Radiochemical yield simulation in water……....93

3.1.7. LET calculations……94

3.1.8. Influence of electron elastic scattering models……...95

3.1.9. Influence of electron thermalization models……...97

3.1.10. Influence of chemistry parameters……101

3.2.Evaluation of the impact of the pre-chemical processes……...103

3.2.1. Physical and physico-chemical processes…….……103

3.2.2. Dissociation channels……....108

3.2.3. Displacement of hot fragments……111

3.2.4. Validation study by comparing simulation data with literature data……114

3.2.5. Results for ionization and excitation……117

3.2.6. Results for electron attachment process……119

3.2.7. Results for electron-hole recombination and dissociation channel……...122

3.2.8. Influence of the pre-chemical processes on water radiolysis simulation…124

3.3.Conclusions……...…....126

References……...……..128

Chapter 4 Chemical stage 135

4.1.Green Function Diffusion Equation…….……136

4.1.1. Totally diffusion-controlled reactions (Types I and III) ……...138

4.1.2. Partially diffusion-controlled reactions (Types II and IV) ……....140

4.1.3. First-order reactions and background reactions (Type VI reactions) …..146

4.1.4. Contact reactions……...147

4.2.The independent reaction time technique……148

4.2.1. Search range…….150

4.2.2. Sampling of diffusion-controlled reactions……...151
4.2.3. Sampling of type VI reactions and contact reactions……153
4.2.4. Reaction site and position of secondaries……154
4.3.Implementation of IRT in the chemistry framework of Geant4-DNA……155
4.4.Radiochemical yield simulation in water……..156
4.5.Validation of implementation……...157
4.6.Results for G-values versus time……..160
4.7.Results for G-values versus LET……..163
4.8.Conclusions……..…….164
References…….165

Chapter 5 DNA damage 169

5.1.DNA geometry……..……171
5.2.DNA damage and scoring……..……..……172
5.2.1. Source classification…….173
5.2.2. Complexity of breaks……176
5.3.Implementation of IRT method into "molecularDNA" example……..177
5.4.Verification of the applicability of the IRT approach ("cylinders" approach) …178
5.5.Evaluation of DNA damage in a simplified human cell ("human_cell" approach)……...179
5.6.Cell repair model……...181
5.7.Results for verification of the applicability of the IRT approach…….182
5.8.Results for human cell nucleus…….183
5.9.Conclusions……..…….191
References……..…...……192

Conclusions & perspectives……...…….197

References……..…...……200

Chapter 1

Introduction

Table of contents

1.1. Context……..……...2
1.2. Monte Carlo track structure simulation……..…….4
1.3. The Geant4-DNA project……9
1.3.1. Physical stage……...…...10
1.3.2. Pre-chemical and chemical stages…….…..14
1.3.3. Geometrical models and DNA damage scoring……..….16
1.4. Purpose of this thesis…….18
References……19

1.1. Context

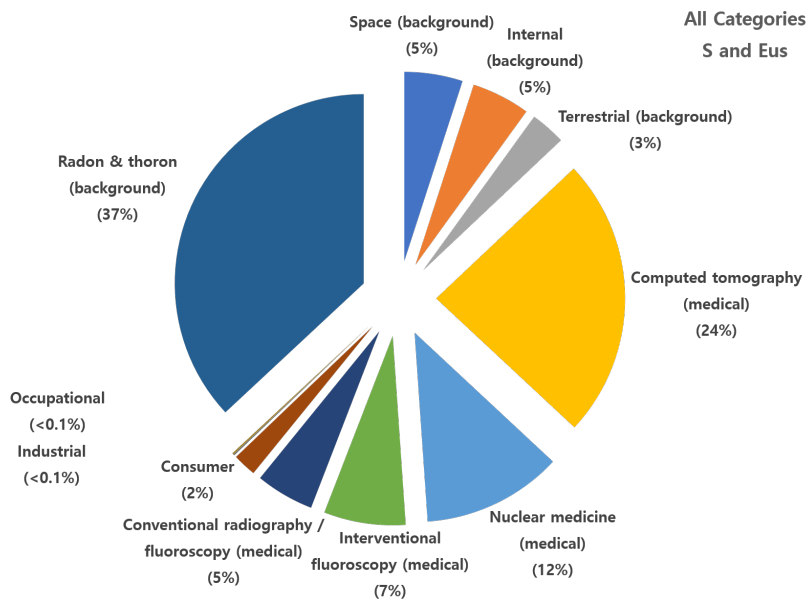


Figure 1.1: Percentage contribution of radiation sources reported by the National Council on Radiation Protection and Measurements (NCRP, 2009) based on collective effective dose S (person-sievert) and effective dose per individual in the U. S. population E_{us} (millisievert).

Human-beings are exposed to various sources of ionizing radiation during their life. According to the NCRP report 160 (NCRP, 2009), the average annual radiation dose in the U. S. is about 6.2 mSv, 50% originating from background radiation and 50% from human-made sources, as illustrated in Figure 1.1. The major contribution of human-made sources is for medical purpose.

Since X-rays were discovered in 1895 and researchers started to investigate the medical uses of ionizing radiation, the deleterious effects on human body have been of global interest for more

than a century (Nias, 1998).

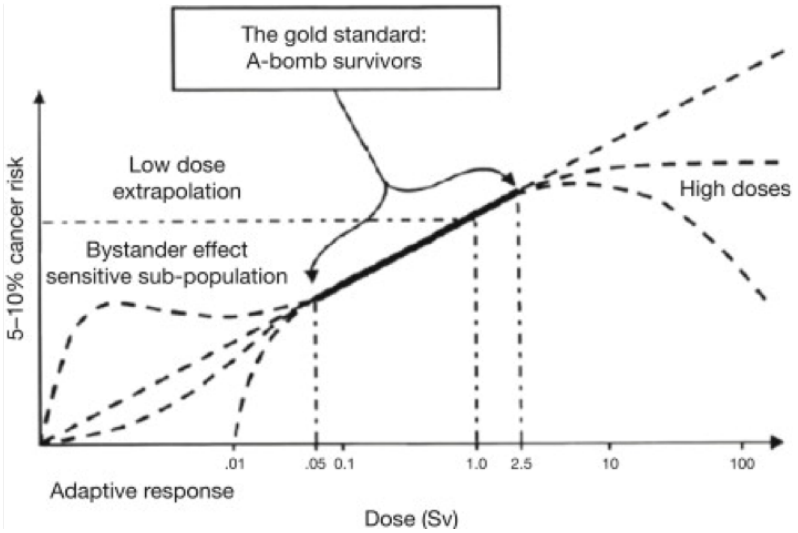


Figure 1.2: The cellular damage risk induced by ionizing radiation as a function of annual dose rate from Hall (2004).

It is possible to epidemiologically predict the radiation risks on humans by following up the atomic bomb survivors (UNSCEAR, 2000, Hall, 2004) at high doses, however the ionizing radiation hazards below the covered dose range (generally a few hundreds of mSv) suffer large uncertainties as illustrated in Figure 1.2. Several risk models in the so called "low-dose" region have been proposed: the Linear-No Threshold (LNT) model, which assumes that the stochastic effect has a linear relationship with dose and no lower dose threshold. This model forms the modern concept of radiation safety, "as low as reasonably achievable" (ALARA) (ICRP, 1977, Prasad et al., 2004), even though it is reported that the linear extrapolation cannot accurately predict the radiation risks (Hooker et al., 2004). Alternatively, the adaptive response model, well-known as radiation hormesis, is a hypothesis that any toxin below threshold stimulates a protective biological response, even in ionizing radiation (UNSCEAR, 1994, Wolff, 1998). There is another interesting experimental observation that unirradiated cells in an irradiated population of cells receive a signal from neighbor irradiated cells and mimic them, the so-called radiation-induced "bystander" effects (Nagasawa and Little, 1992, Seymour and Mothersill, 2004). However, experimental validation of those risk models remains today a scientific challenge.

The mechanistic evaluation of biological effects induced by ionizing radiation is necessary, in order not only to understand low-dose carcinogenesis in many domains but also to improve existing and develop, innovative therapeutic approaches that use ionizing radiation. For example, accurate calculation of relative biological effect (RBE) is important in radiation therapy, especially for charged particles (e.g. proton and carbon therapy) (Paganetti et al., 2002, Frese et al., 2012). Recently, biological effects revealed for nanoparticle-aided, FLASH or mini/microbeam radiotherapies still need to be elucidated (Engels et al., 2016, Dos Santos et al., 2020, Ramos-

Méndez et al., 2020). Also, the low-dose irradiation to patients undergoing radiology and nuclear medicine exams is an important subject of research (Fazel et al., 2009). In addition, in the case of radiation industry and space science, the influence of chronic exposure in the industry (Howe et al., 2004) or during space missions (Mortazavi et al., 2003) should be evaluated for radioprotection of workers.

In order to elucidate the mechanisms involved in ionizing radiation damage, the structure of human cells and their characteristics have been studied for a long time. Mammalian cells, including human cells, are complex biological systems consisting of a nucleus and surrounding cytoplasm. The cells contain several cytoplasmic organelles such as mitochondria, ribosomes, Golgi vesicles, centrioles, and lysosomes. However, it is still admitted today that the most sensitive target to ionizing radiation is the cell nucleus and its deoxyribonucleic acid (DNA) content, which can critically impact the fate of the cell after irradiation (Nias, 1998).

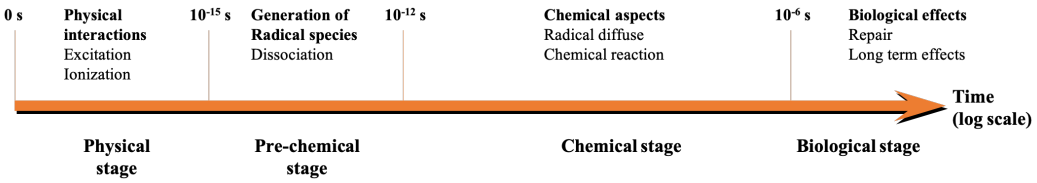


Figure 1.3: The time frame for effects of ionizing radiation reproduced from Turner (2007).

It is classically reported that the radiobiological mechanisms consist of physical, pre-chemical, chemical and biological stages (Turner, 2007) as shown in Figure 1.3. At first, the physical stage takes place in attosecond scale ($< 10^{-15}$ s) and corresponds to excitations and ionizations of molecules which lead to energy deposition. The resulting excited and ionized molecules can fragment into molecular species ($< 10^{-12}$ s) which can chemically react with biomolecules (e.g. DNA and RNA) present in the cell and induce early indirect radiobiological effects ($< 10^{-6}$ s). Nowadays, it is recognized that Monte Carlo (MC) simulations are the most reliable approach in order to estimate early radiobiological effects induced by ionizing radiation (Dingfelder, 2012).

1.2. Monte Carlo track structure simulation

The MC technique is based on random number generation and enables to simulate the stochastic nature of particle-matter interactions. In particular, it is used for the simulation of the transportation of radiation through matter (Metropolis, 1987). However, many general-purpose MC codes and toolkits, such as MCNP (Pelowitz, 2011) and Geant4 (Agostinelli et al., 2003, Allison et al., 2006, Allison et al., 2016), use a "condensed-history" approach, which approximates the multiple particle collisions as one single step accumulating them (Larsen and Tolar, 2001). The step limit is usually too large to simulate accurately particle transportations below the micrometer, which

is typically the sub-cellular scale (Lazarakis et al., 2018). Moreover, it is reported that the contribution of indirect damage is dominant at low linear energy transfer (LET) (Hirayama et al., 2009). Thus, the simulation of physical interactions is not sufficient and a careful modeling of radiolysis is required, including diffusions and chemical reactions of molecular species with biological medium, for the evaluation of indirect DNA damage (O'Neill and Wardman, 2009). In order to overcome such limitations of the MC technique, a number of Monte Carlo Track Structure (MCTS) codes and toolkits have been developed.

In brief, the "MCTS method" represents the MC method simulating every interaction without condensed-history approximation, using a "discrete" approach, which simulates particle transportation step-by-step. Most of the MCTS tools approximate the target as liquid water which composes more than 60% of human body (Mitchell et al., 1945).

The modeling of pre-chemical and chemical stages is more complicated. In the same spirit as for the physical stage, the water radiolysis simulation is performed with the assumption that human biological medium consists of liquid water. The water molecules ionized and excited during the physical stage undergo dissociation processes during the pre-chemical stage. For example, Figure 1.4 shows an overview of dissociation channels proposed by Kreipl et al. (2009).

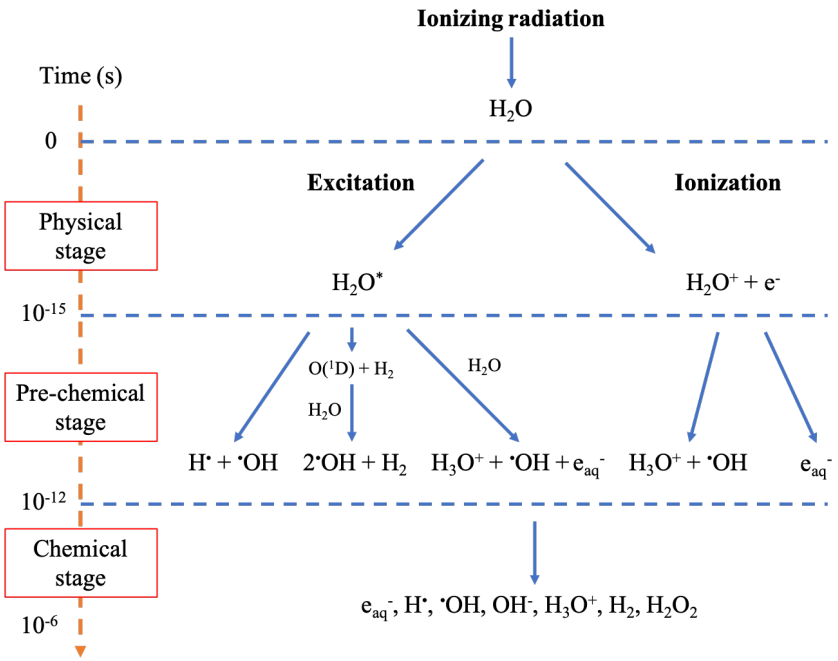


Figure 1.4: Dissociation approach proposed by Kreipl et al. (2009). The figure is reproduced from Buck et al. (2012).

The molecular species (e_{aq}^-, $H^\bullet$, $^\bullet OH$, OH^-, H_3O^+, H_2, H_2O_2) generated by the dissociation process in the pre-chemical stage diffuse following a Brownian motion (Knight, 1962) during the chemical stage. In most MCTS tools, the Brownian transportation of the species is typically modeled using step-by-step (SBS) method (Turner et al., 1983, Michalik et al., 1998, Kreipl et al.,

2009, Karamitros et al., 2011), which diffuses all these molecules at every single time step. However, water radiolysis simulation using the SBS method has a huge computational burden due to the necessity to diffuse all the molecular species and calculate interparticle distances. Due to such limitation, several MCTS tools implemented the independent reaction time (IRT) method (Clifford et al., 1986), which approximates that the reaction probability depends on the initial separation distance and is independent from the diffusion trajectory (Plante and Devroye, 2017).

Simplified geometrical models of biological targets such as DNA, chromatin fibers, cell nuclei are fully tool-specific and usually do not provide a variety of such geometries. For instance, the KURBUC code uses a simplified cylindrical chromosome model developed by Charlton et al. (1989), Nikjoo et al. (1994), and the geometry is extended up to a cell nucleus (Nikjoo and Girard, 2012). In the case of the PARTRAC code (Friedland et al., 2003), more complex rosette structure is modeled up to human fibroblast cell nucleus (Friedland et al., 2011). Both geometrical models can be used to estimate the damages from cell nucleus scale down to chromatin fiber, base pairs, and even biomolecules.

Table 1.1 from Tang (2019) shows the available MCTS simulation tools existing today, the list of particles they can transport, the electron energy range covered, the available biological media, and the capacity of simulating chemical stage. A more detailed review of some of these MCTS codes can be found in Nikjoo et al. (2006). In this thesis, we will focus exclusively on the Geant4-DNA MCTS extension of the Geant4 general purpose and open source MC toolkit.

Table 1.1: List of Monte Carlo track structure codes from Tang (2019).

MCTS tools	Available particle	Energy range of e-	Medium	Indirect damage	Reference
CPA100	e-	Therm.[a] – 256 keV	Water (L[b]), DNA	O	Terrissol and Beaudre (1990)
Delta	e-	10 eV – 10 keV	Water (V[c])	O	Zaider et al. (1983)
EPOTRAN	e-, e+	7.4 eV – 10 keV	Water (L,V)	X	Champion et al. (2012)
ETRACK	e-, p, α	10 eV – 10 keV	Water (V)	O	Ito (1987)
ETS	e-	10 eV – 10 keV	Water (L,V)	O	Hill and Smith (1994)
Geant4-DNA	e-, e+, p, H, α, ions	Therm. – 1 MeV	Water (L), DNA	O	Incerti et al. (2010a)
IONLYS/IONLYS-IRT	e-, ions	10 eV – 100 keV	Water (L)	O	Cobut et al. (1998)
KAPLAN	e-	1 – 10 keV	Water (L, V)	O	Kaplan et al. (1990)
KITrack	e-, ions	10 eV – 100 keV	Water (L)	X	Wiklund et al. (2011)
KURBUC	e-, p, α, ions, carbon	10 eV – 10 MeV	Water (L, V)	O	Uehara et al. (1993)
LEEPS	e-, e+	100 eV – 100 keV	Several materials	X	Fernández-Varea et al. (1996)
LEPTS	e-, e+, p	Therm. – 10 keV e-	Water (V), CH_4, C_2H_4, C_4H_8O, SF_6, $C_4H_4N_2$	X	Sanz et al. (2012), Blanco et al. (2013)
LionTrack	e-, p, ions	> 50 eV	Water (L)	X	Backstrom et al. (2013)
LQD	e-, ions	-	Water (L)	O	Gervais et al. (2005)
MC4	e-, ions	10 eV	Water (L, V)	X	Emfietzoglou et al. (2003)
MOCA8B	e-	10 eV – 100 keV e-	Water (L, V)	O	Paretzke (1987)
NASIC	e-	Therm. – 1 MeV e-	Water (L)	O	Li et al. (2015)
NOTRE DAME	e-, ions	10 eV e-	Water (L, V)	O	Pimblott et al. (1990)
OREC/NOREC	e-	7.4 eV – 1 MeV e-	Water (L)	X	Semenenko et al. (2003)
PARTRAC	e-, e+, p, H, α, ions	1 eV – 10 MeV	Water (L), DNA	O	Friedland et al. (2003)

PITS04	e-, ions	10 eV	Water (L)	X	Wilson et al. (2004)
PITS99	e-, ions	10 eV	Water (V)	O	Wilson and Nikjoo (1999)
Ptra	e-, p, α	1 eV – 10 keV	Water (L, V), DNA	X	Grosswendt and Pszona (2002)
RADAMOL (TRIOL/STOCHECO)	e-, ions	7.4 eV – 2 MeV	Water (L)	O	Bigildeev and Michalik (1996)
RETRACKS/RITRACKS	e-, ions	0.1 eV – 100 MeV	Water (L, V)	O	Plante and Cucinotta (2009)
SHERBROOKE	e-, ions	10 eV	Water (L, V)	O	Cobut et al. (2004)
STBRGEN	e-, ions	10 eV	Water (L, V)	O	Chatterjee and Holley (1993)
TILDA-V	e-, p, H, ions	7.4 eV	Water (L, V), DNA	X	Champion et al. (2005)
TOPAS-nBio	e-, e+, p, H, α, ions	Therm. – 1 MeV	Water (L), DNA	O	Schuemann et al. (2019)
TRAX/TRAX-CHEM	e-, e+, p, ions	1 eV – several MeV	Water (V)	O	Krämer and Kraft (1994)
TRION	e-, ions	10 eV	Water (L, V)	X	Lappa et al. (1993)
TRACEL	e-, ions	10 eV	Water (L, V)	O	Tomita et al. (1997)

[a] Therm. indicates the thermalization energy of electron

[b] Liquid phase of water

[c] Vapour phase of water

1.3. The Geant4-DNA project

GEometry ANd Tracking4 (Geant4 - https://geant4.web.cern.ch) is an open source Monte Carlo toolkit developed in C++ language (Agostinelli et al., 2003, Allison et al., 2006, Allison et al., 2016) initiated in 1994 by an international collaboration for the simulation of high energy physics experiments at the CERN Large Hadron Collider, in Switzerland. Thanks to its object-oriented architecture, Geant4 has been progressively extended over the years for various research fields such as astrophysics, nuclear physics, medical physics, and radiation protection. Today, many international groups collaborate and contribute to the development of the toolkit for these various research topics, as illustrated in Figure 1.5.

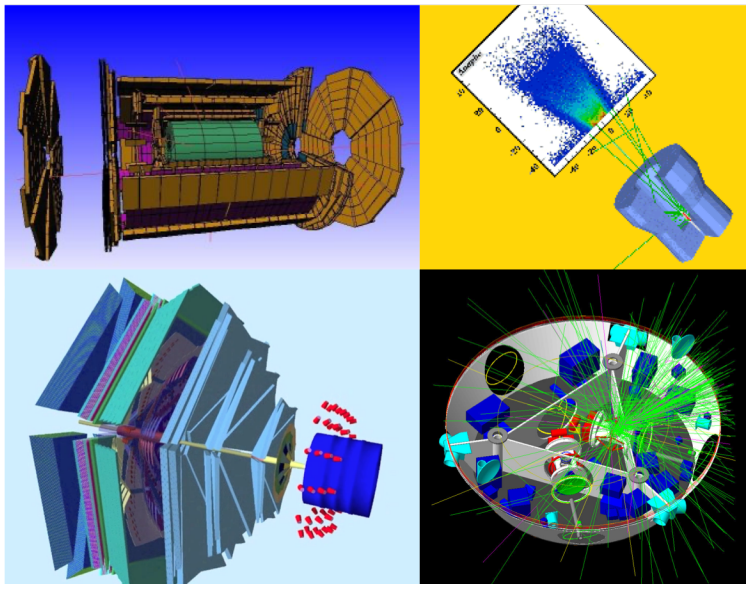


Figure 1.5: Examples of Geant4 applications. The ATLAS project (left upper), a superficial brachytherapy device and corresponding dose distribution (right upper), a modeling of the CLAS12 detector at Jefferson Lab (left below), and the LISA science module spacecraft (right below). All figures are available in the Geant4 website (https://geant4.web.cern.ch).

The Geant4-DNA project (http://geant4-dna.org), fully included in Geant4, was initially launched in 2001 by the European Space Agency (ESA) in order to provide the community with an open access toolkit to evaluate the biological damage induced by ionizing radiation at the subcellular scale (Incerti et al., 2010a, Incerti et al., 2010b, Bernal et al., 2015, Incerti et al., 2018), in the context of space radiation protection studies. It was the first fully open access MCTS code available freely to the community without considering now the TOPAS-nBio extension of TOPAS (Perl et al., 2012, Schuemann et al., 2019), which in particular wrap Geant4-DNA and Geant4 respectively.

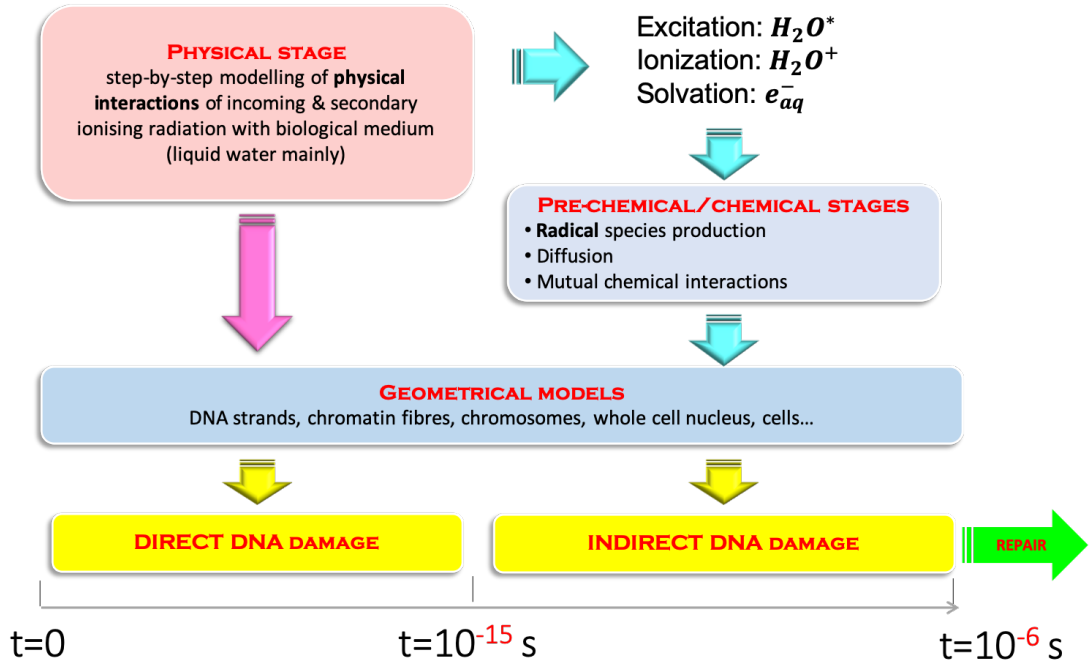


Figure 1.6: Geant4-DNA approach for the simulation of radio-induced biological effects.

Figure 1.6 shows the approach adopted by Geant4-DNA for evaluating DNA damage according to the time evolution and stage. This is a classical approach adopted by other codes (e.g. PARTRAC, KURBUC, TRAX/TRAX-CHEM, etc.). As described above, all the radiobiological stages are available, and the biological damages can be scored according to the source of the damage, direct damage from physical interactions and indirect damage from chemical reactions, respectively. Each stage of Geant4-DNA will be further described in the following paragraphs.

1.3.1. Physical stage

Accurate cross-sections models and descriptions of physical interaction final state (e.g. creation of secondary particles, energy loss and angular deviation of incident particle, etc.) are required for an accurate simulation of step-by-step particle tracking. Especially, low-energy secondary electrons dominantly affect the induction of sub-cellular scale damages (Nikjoo et al., 2016). For instance, inelastic interactions lead to direct energy deposition in the irradiated medium, and elastic interactions influence the energy deposition pattern. One of the main advantages of Geant4-DNA is the possibility to implement several alternative or complementary physics models describing such interactions. This is useful to evaluate the impact of physics models on simulation results. The selection of physics models can be done easily through the usage of "physics constructors" which contain all physics models associated to each particle that Geant4-DNA can handle (e.g. electrons, protons, etc.) for each physical interaction (also called "process", such as ionization, excitation, elastic scattering, etc). Geant4-DNA provides three recommended reference physics constructors

for MCTS simulations in liquid water. These constructors differ only by their electron models (all models for other particles are identical) as shown in Table 1.2. We will describe further the various theoretical or empirical approaches used to calculate such models in Chapter 2, associated references are given for further reading.

Table 1.2: Elastic and inelastic models of electrons employed in Geant4-DNA physics constructors and energy limits of applicability.

Physics constructor	Elastic	Excitation	Ionization	Vibrational excitation	Attachment
G4EmDNAPhysics_ option2*	Champion model (7.4 eV – 1 MeV) (Champion, 2003)	Emfietzoglou dielectric model (9 eV – 1 MeV) (Incerti et al., 2010b)	Emfietzoglou dielectric model (11 eV – 1 MeV) (Incerti et al., 2010b)	Sanche cross-section (2 eV - 100 eV) (Michaud et al., 2003)	Melton cross-section (4 eV – 13 eV) (Melton, 1972)
G4EmDNAPhysics_ option4*	Uehara screened Rutherford model (9 eV – 10 keV) (Uehara et al., 1993)	Emfietzoglou-Kyriakou dielectric model (8 eV – 10 keV) (Emfietzoglou et al., 2005)	Emfietzoglou-Kyriakou dielectric model (10 eV – 10 keV) (Emfietzoglou et al., 2005)	-	-
G4EmDNAPhysics_ option6*	CPA100 model (11 eV – 256 keV) (Bordage et al., 2016)	CPA100 model (11 eV – 256 keV) (Bordage et al., 2016)	CPA100 model (11 eV – 256 keV) (Bordage et al., 2016)	-	-

* Abbreviated later as “option 2”, “option 4”, and “option 6”

In addition to electrons, Geant4-DNA can simulate physical interactions (processes) for protons and alpha particles including their charged states (H^0, H^+, He^0, He^+, He^{2+}). The models for protons and alpha particles are based on the models of Dingfelder et al. (2000). Below 500 keV, the Miller and Green excitation model and Rudd ionization model described in Rudd et al. (1985), Dingfelder et al. (2000) are used. The Born and Bethe theories are used above 500 keV for Rudd et al. (1985), Dingfelder et al. (2000). The model for the charge exchange process (gain or loss of electrons) is also obtained by applying the analytical model of Dingfelder et al. (2000). These models allow to perform simulations in the energy range of 100 eV-100 MeV and 1 keV-400 MeV for protons and alpha particles, respectively. In the case of heavy ions, such as Li, Be, B, C, N, O, Si, Fe, only the discrete ionization model of Booth and Grant (1965) is available (Francis et al., 2011a). Geant4-DNA uses the Livermore physics models for photons, based on the Evaluated Photon Data Library (EPDL97), which is available in Geant4 for the low energy domain (Cullen et al., 1997). All details of Geant4-DNA physics models are well-described in Incerti et al. (2010b), Incerti et al. (2018).

Moreover, Geant4-DNA provides also several examples for evaluating physical quantities which can be used to reproduce previous Geant4-DNA literature results. The list and associated references are shown in Table 1.3.

Table 1.3: List of Geant4-DNA examples available for MCTS simulations in liquid water, taken from Incerti et al. (2018).

Physics example	Purpose	Reference
dnaphysics	Details of tracking, automatic combination with Geant4 standard electromagnetic physics models	Bernal et al. (2015)
microdosimetry	Combination of Geant4 standard electromagnetic and Geant4-DNA processes and models in different regions	Bernal et al. (2015)
range	Range, projected range, penetration	Kyriakou et al. (2016)
spower	Stopping power	Incerti et al. (2017)
mfp	Mean-free-path (MFP)	Incerti et al. (2018)
wvalue	Mean energy required for the creation of an ion pair in liquid water (the so-called "W-value")	Kyriakou et al. (2015)
svalue	Dose to a liquid water target per unit of cumulated activity in a source region (the so-called "S-value")	Bernal et al. (2015), André et al. (2014), Sefl et al. (2015)
slowing	Slowing-down electron spectra	Incerti et al. (2017)
microyz	Microdosimetric distributions (lineal energy y, specific energy z) and related quantities	Kyriakou et al. (2017)
TestEm12	Dose point kernel	Bernal et al. (2015), Kyriakou et al. (2016), Bordes et al. (2017)
TestEm5	Identification of atomic de-excitation products for Geant4-DNA processes	-

1.3.2. Pre-chemical and chemical stages

As the simulation of physical interactions, Geant4-DNA provides a "chemistry constructor" which contains the dissociation probabilities of the ionized and excited water molecules, the list of molecular species, their diffusion coefficients, and the chemical reaction rates. In Geant4-10.3, there was only one chemistry constructor available, "G4EmDNAChemistry", based on the chemistry model of PARTRAC (Kreipl et al., 2009, Karamitros, 2012, Karamitros et al., 2014).

In the pre-chemical stage, the dissociations of molecular species and their probabilities are given according to the excitation and ionization levels of the water molecule as shown in Figure 1.4. And then, the initial positions of molecular species are determined by momentum conservation and empirical root-mean-square distance.

The modeling of chemistry (radiolysis) in Geant4-DNA is based on the SBS approach combining Smoluchowski Brownian diffusion equation (Berg, 1993) describing Brownian motion

and Brownian bridge technique developed by Karamitros (2012), Karamitros et al. (2014). However, the chemistry aspects in reality are continuous. In order to avoid the distortion induced by using discrete steps and to reasonably reduce calculation time, Geant4-DNA dynamically calculates time steps (using the "G4DNAMoleculeEncounterStepper" class) based on an idea initially proposed by Michalik et al. (1998). This technique evaluates the probability of a chemical reaction within the selected statistical confidence (95% confidence level by default). For each time step, the SBS algorithm should find the closest reactant to verify whether the reaction has happened. However, this process requires lots of separation assessments, of order N^2 (square of the reactants number). K-d tree algorithm (de Berg et al., 2008) allows to decrease the time complexity from N^2 to $N \times \ln(N)$ based on a space-partitioning technique for organizing points in a k-dimensional space. Chemical reactions occur when two molecular species are closer than the specific reaction radius. After all possible chemical reactions have been processed, molecular species are diffused based on their diffusion coefficients and Brownian diffusion equation. However, due to the discretization of time steps, we should carefully consider possible reactions that can occur during a discrete step. For example, when the two reactants are separated by more than reaction radius at pre-step and post-step point which are the initial and final times of a time step, respectively. It is indeed possible that the separation is less than reaction radius during a time step. In order to account for these possible reactions, Brownian bridge technique is implemented in Geant4-DNA (in the "G4DNASmoluchowskiReactionModel" class).

Several Geant4-DNA examples are provided in order to test water radiolysis simulation and also to evaluate radiochemical yields, so called "G-values" which are the number of molecular species normalized to a deposited energy amount of 100 eV, using SBS approach, as shown in Table 1.4.

Table 1.4: List of Geant4-DNA examples available for water radiolysis simulations updated from Bernal et al. (2015).

Chemistry example	Purpose	Reference
chem1	Activation of chemical module	Karamitros et al. (2011)
chem2	Selection of time steps	Karamitros et al. (2011)
chem3	Visualization of chemical stage as a function of time	Karamitros et al. (2011)
chem4	Calculation of G-values as a function of time	Karamitros et al. (2011)
chem5	Calculation of G-values as a function of time using specific constructors (released later in Geant4.10.5)	Ramos-Mendez et al. (2018)

1.3.3. Geometrical models and DNA damage scoring

Several approaches to calculate DNA damages induced by ionizing radiation using Geant4-DNA have been proposed so far, as shown in Table 1.5. However, those examples (when released in version 10.3 of Geant4) oversimplify the damage scoring, as for example clustering of energy deposition without considering the simulation of full chemistry for the simulation of indirect damage.

In order to more accurately evaluate DNA damages, two geometrical approaches have been later proposed based on two dedicated examples: the *dnadamage1* and *molecularDNA* examples. The *dnadamage1* example was developed using the external tool *DNAfabric*, a C++ software generating complex DNA geometrical models from the nucleotide scale to cell nucleus (Meylan et al., 2016). This example can simulate not only physical interactions of sub-micrometer scale but also chemistry aspects of biomolecules (Meylan et al., 2017). The version released in Geant4 is limited to a segment of chromatin fiber and uses the step-by-step modelling of chemistry. Alternatively, the *molecularDNA* example developed by Lampe (2017) is also able to simulate both physical and pre-chemical/chemical stages (Lampe et al., 2018a, Lampe et al., 2018b) based on a private version of the IRT approach developed by Karamitros et al. (2020); in this example the geometrical model can be generated using python script, bringing easiness to model different types of geometries such as E. coli bacterium (Lampe et al., 2016). This example has not been released in Geant4. One of the objectives of this thesis is to finalize the development of this example for its future release.

Regarding the simulation of late damage, beyond a few ns, a repair model has not been released yet. Belov et al. (2015) were the first to propose a repair model for human fibroblast cells that can be implemented in Geant4-DNA. The model well-validated by Sakata et al. (2020) will be soon released with the *molecularDNA* example.

Table 1.5: List of Geant4-DNA examples available for geometrical and damage simulation reproduced updated from Bernal et al. (2015).

Damage example	**Purpose**	**Reference**
wholeNuclearDNA	Geometry of the DNA contained in a eukaryotic cell nucleus	Dos Santos et al. (2013)
pdb4dna	Interface to the Protein Data Bank for the implementation of realistic molecular geometries	Delage et al. (2015)
clustering	Pattern of energy deposition	Francis et al. (2011b)
microbeam	Cellular irradiation in single ion mode	Incerti et al. (2007)
neuron	Simulation of a neural network	Belov et al. (2016)
dnadamage1	Geometry generated by DNAfabric tool (Released in Geant4.10.6)	Meylan et al. (2017)
molecularDNA	Geometry generated by python script (Not released yet)	Lampe (2017)

Geant4-DNA is available in open access to the community since 2010, being fully included in Geant4, and every year new developments are distributed in open access. It allows to simulate all the radiobiological stages described above, which can be reused by other simulation tools such as TOPAS-nBio (Ramos-Mendez et al., 2018) or GATE (Pham et al., 2015). However, Geant4-DNA has still important limitations that must be improved:

1) The elastic model, especially for low energy electrons, is difficult to validate because elastic cross-section measurements in liquid water are a technical challenge. Many MCTS tools including Geant4-DNA have taken into account the phase influence between vapour and liquid phase, however, the default elastic scattering model (Champion, 2003) shows deviations with experimental data.
2) The pre-chemical stage of Geant4-DNA is mainly based on that of PARTRAC, which adjusts parameters for matching with experimental G-values (Kreipl et al., 2009). However, Geant4-DNA takes into account additional pre-chemical interactions, such as electron molecular attachment and electron-hole recombination processes, and this leads to disagreements with experimental data.
3) Simulation of radiochemical yields using SBS method is a huge computational burden. For instance, a few days are needed for the proton case performed in the previous study (Meylan et al., 2017).
4) A fully integrated damage simulation chain easily usable at nucleus scale is still missing in Geant4-DNA.

1.4. Purpose of this thesis

The aim of this thesis is to improve Geant4-DNA for a more accurate simulation of DNA damage induced by ionizing radiation, up to late effects. This thesis includes (1) the development of more accurate Geant4-DNA physical and chemical models, and (2) the development of a fully integrated simulation "chain", simultaneously simulating all the radiobiological stages.

For that, at first, we propose to develop a new elastic scattering model for electrons in liquid water in order to accurately simulate the spatial distribution of secondary electrons. The developed model is verified and validated by comparing several physical quantities with experimental and computational results. This is described in Chapter 2. The new elastic scattering model may not impact much on physical quantities because it does not lead to any energy deposition, however, this model could influence the concentration of DNA damages, which has direct correlation with double strand breaks (DSBs), one of the most critical damages (Khanna and Jackson, 2001).

In Chapter 3, the pre-chemical stage of water radiolysis simulation is improved based on the approaches described in the literature. The pre-chemical models such as electron thermalization models, pre-chemical interactions, and dissociation channels employed by several MCTS tools are compared. The influence of the adjustable settings is evaluated for the initial radiochemical yields (which significantly affect the chemical stage), as a function of LET and time, using the new *chem6* example developed in this thesis.

Next, in order to reduce the computational burden in the chemical stage, a new implementation of the independent reaction time (IRT) method is proposed in Chapter 4. The validation of this implementation is performed by comparing predictions with experimental data on G-values. This development is a key component of the simulation chain since it will allow to reach sufficient statistics in reasonable calculation time.

Finally, in Chapter 5, the impact of the above developments is evaluated with the prediction of biological damage using the not-yet-released *molecularDNA* example. Not only SSB and DSB yields but also other quantities, including repair tendency, are calculated using a simplified geometry of a human cell nucleus model. The results are compared with experimental data.

References

Agostinelli, S., Allison, J., Amako, K., et al. 2003, Geant4 - a simulation toolkit. *Nucl. Instrum. Meth. A,* 506**,** 250-303.

Allison, J., Amako, K., Apostolakis, J., et al. 2006, Geant4 developments and applications. *IEEE T. Nucl. Sci.,* 53**,** 270-278.

Allison, J., Amako, K., Apostolakis, J., et al. 2016, Recent developments in Geant4. *Nucl. Instrum. Meth. A,* 835**,** 186-225.

André, T., Morini, F., Karamitros, M., et al. 2014, Comparison of Geant4-DNA simulation of S-values with other Monte Carlo codes. *Nucl. Instrum. Meth. B,* 319**,** 87-94.

Backstrom, G., Galassi, M. E., Tilly, N., et al. 2013, Track structure of protons and other light ions in liquid water: applications of the LIonTrack code at the nanometer scale. *Med. Phys.,* 40**,** 064101.

Belov, O. V., Batmunkh, M., Incerti, S., et al. 2016, Radiation damage to neuronal cells: Simulating the energy deposition and water radiolysis in a small neural network. *Phys. Med.,* 32**,** 1510-1520.

Belov, O. V., Krasavin, E. A., Lyashko, M. S., et al. 2015, A quantitative model of the major pathways for radiation-induced DNA double-strand break repair. *J. Theor. Biol.,* 366**,** 115-30.

Berg, H. C. 1993. *Random walks in biology,* New Jersey, Princeton University Press.

Bernal, M. A., Bordage, M. C., Brown, J. M. C., et al. 2015, Track structure modeling in liquid water: A review of the Geant4-DNA very low energy extension of the Geant4 Monte Carlo simulation toolkit. *Phys. Med.,* 31**,** 861-874.

Bigildeev, E. A. & Michalik, V. 1996, Charged particle tracks in water of different phases. Monte Carlo simulation of electron tracks. *Radiat. Phys. Chem.,* 47**,** 197-207.

Blanco, F., Muñoz, A., Almeida, D., et al. 2013, Modelling low energy electron and positron tracks in biologically relevant media. *Eur. Phys. J. D,* 67.

Booth, W. & Grant, I. S. 1965, The energy loss of oxygen and chlorine ions in solids. *Nucl. Phys.,* 63**,** 481-495.

Bordage, M. C., Bordes, J., Edel, S., et al. 2016, Implementation of new physics models for low energy electrons in liquid water in Geant4-DNA. *Phys. Med.,* 32**,** 1833-1840.

Bordes, J., Incerti, S., Lampe, N., et al. 2017, Low-energy electron dose-point kernel simulations using new physics models implemented in Geant4-DNA. *Nucl. Instrum. Meth. B,* 398**,** 13-20.

Buck, E. C., Wittman, R. S., Skomurski, F. N., et al. 2012. Radiolysis process modeling results for scenarios. PNNL-21554.

Champion, C. 2003, Theoretical cross sections for electron collisions in water: structure of electron tracks. *Phys. Med. Biol.,* 48**,** 2147-68.

Champion, C., L'Hoir, A., Politis, M. F., et al. 2005, A Monte Carlo code for the simulation of heavy-ion tracks in water. *Radiat. Res.,* 163**,** 222-231.

Champion, C., Le Loirec, C. & Stosic, B. 2012, EPOTRAN: a full-differential Monte Carlo code for electron and positron transport in liquid and gaseous water. *Int. J. Radiat. Biol.,* 88**,** 54-61.

Charlton, D. E., Nikjoo, H. & Humm, J. L. 1989, Calculation of initial yields of single- and double-strand breaks in cell nuclei from electrons, protons and alpha particles. *Int. J. Radiat. Biol.,* 56**,** 1-19.

Chatterjee, A. & Holley, W. R. 1993, Computer simulation of initial events in the biochemical mechanisms of DNA damage. *Adv. Radiat. Biol.,* 17**,** 181-226.

Clifford, P., Green, N. J. B., Oldfield, M. J., et al. 1986, Stochastic models of multi-species kinetics in radiation-induced spurs. *J. Chem. Soc. Faraday Trans. 1,* 82**,** 2673-2689.

Cobut, V., Cirioni, L. & Patau, J. P. 2004, Accurate transport simulation of electron tracks in the energy range 1 keV–4 MeV. *Nucl. Instrum. Meth. B,* 215**,** 57-68.

Cobut, V., Frongillo, Y., Patau, J. P., et al. 1998, Monte Carlo simulation of fast electron and proton tracks in liquid water-I. Physical and physicochemical aspects. *Radiat. Phys. Chem.,* 51**,** 229-243.

Cullen, D. E., Hubbell, J. H. & Kissel, L. 1997. EPDL97: the evaluated photo data library `97 version. Lawrence Livermore National Lab., CA (United States).

de Berg, M., van Kreveld, M., Overmars, M., et al. 2008. *Computational geometry,* Berlin, Heidelberg, Springer.

Delage, E., Pham, Q. T., Karamitros, M., et al. 2015, PDB4DNA: Implementation of DNA geometry from the Protein Data Bank (PDB) description for Geant4-DNA Monte-Carlo simulations. *Comput. Phys. Commun.,* 192**,** 282-288.

Dingfelder, M. 2012, Track-structure simulations for charged particles. *Health Phys.,* 103**,** 590-5.

Dingfelder, M., Inokuti, M. & Paretzke, H. 2000, Inelastic-collision cross sections of liquid water for interactions of energetic protons. *Radiat. Phys. Chem.,* 59**,** 255-275.

Dos Santos, M., Delorme, R., Salmon, R., et al. 2020, Minibeam radiation therapy: A micro- and nano-dosimetry Monte Carlo study. *Med. Phys.,* 47**,** 1379-1390.

Dos Santos, M., Villagrasa, C., Clairand, I., et al. 2013, Influence of the DNA density on the number of clustered damages created by protons of different energies. *Nucl. Instrum. Meth. B,* 298**,** 47-54.

Emfietzoglou, D., Cucinotta, F. A. & Nikjoo, H. 2005, A complete dielectric response model for liquid water: a solution of the Bethe ridge problem. *Radiat. Res.,* 164**,** 202-211.

Emfietzoglou, D., Karava, K., Papamichael, G., et al. 2003, Monte Carlo simulation of the energy loss of low-energy electrons in liquid water. *Phys. Med. Biol.,* 48**,** 2355.

Engels, E., Corde, S., McKinnon, S., et al. 2016, Optimizing dose enhancement with Ta2O5 nanoparticles for synchrotron microbeam activated radiation therapy. *Phys. Med.,* 32**,** 1852-1861.

Fazel, R., Krumholz, H. M., Wang, Y., et al. 2009, Exposure to low-dose ionizing radiation from medical imaging procedures. *New Engl. J. Med.,* 361**,** 849-857.

Fernández-Varea, J., Liljequist, D., Csillag, S., et al. 1996, Monte Carlo simulation of 0.1–100 keV electron and positron transport in solids using optical data and partial wave methods. *Nucl. Instrum. Meth. B,* 108**,** 35-50.

Francis, Z., Incerti, S., Karamitros, M., et al. 2011a, Stopping power and ranges of electrons, protons and alpha particles in liquid water using the Geant4-DNA package. *Nucl. Instrum. Meth. B,* 269**,** 2307-2311.

Francis, Z., Villagrasa, C. & Clairand, I. 2011b, Simulation of DNA damage clustering after proton irradiation using an adapted DBSCAN algorithm. *Comput. Meth. Prog. Bio.,* 101**,** 265-70.

Frese, M. C., Yu, V. K., Stewart, R. D., et al. 2012, A mechanism-based approach to predict the relative biological effectiveness of protons and carbon ions in radiation therapy. *Int. J. Radiat. Oncol.,* 83**,** 442-50.

Friedland, W., Dingfelder, M., Kundrat, P., et al. 2011, Track structures, DNA targets and radiation effects in the biophysical Monte Carlo simulation code PARTRAC. *Mutat. Res.,* 711**,** 28-40.

Friedland, W., Jacob, P., Bernhardt, P., et al. 2003, Simulation of DNA Damage after Proton Irradiation. *Radiat. Res.,* 159**,** 401-410.

Gervais, B., Beuve, M., Olivera, G. H., et al. 2005, Production of HO2 and O2 by multiple ionization in water radiolysis by swift carbon ions. *Chem. Phys. Lett.,* 410**,** 330-334.

Grosswendt, B. & Pszona, S. 2002, The track structure of alpha-particles from the point of view of ionization-cluster formation in "nanometric" volumes of nitrogen. *Radiat. Environ. Bioph.,* 41**,** 91-102.

Hall, E. J. 2004, Henry S. Kaplan Distinguished Scientist Award 2003. The crooked shall be made straight; dose-response relationships for carcinogenesis. *Int. J. Radiat. Biol.,* 80**,** 327-37.

Hill, M. A. & Smith, F. A. 1994, Calculation of initial and primary yields in the radiolysis of water. *Radiat. Phys. Chem.,* 43**,** 265-280.

Hirayama, R., Ito, A., Tomita, M., et al. 2009, Contributions of direct and indirect actions in cell killing by high-LET radiations. *Radiat. Res.,* 171**,** 212-8.

Hooker, A. M., Bhat, M., Day, T. K., et al. 2004, The linear no-threshold model does not hold for low-dose ionizing radiation. *Radiat. Res.,* 162**,** 447-52.

Howe, G. R., Zablotska, L. B., Fix, J. J., et al. 2004, Analysis of the mortality experience amongst U.S. nuclear power industry workers after chronic low-dose exposure to ionizing radiation. *Radiat. Res.,* 162**,** 517-526.

ICRP 1977. Recommendations of the ICRP. *Annals of the ICRP.*

Incerti, S., Baldacchino, G., Bernal, M., et al. 2010a, The Geant4-DNA project. *Int. J. Model. Simul. Sci. Comput.,* 1**,** 157-178.

Incerti, S., Ivanchenko, A., Karamitros, M., et al. 2010b, Comparison of GEANT4 very low energy cross section models with experimental data in water. *Med. Phys.,* 37**,** 4692-4708.

Incerti, S., Kyriakou, I., Bernal, M., et al. 2018, Geant4-DNA example applications for track structure simulations in liquid water: a report from the Geant4-DNA Project. *Med. Phys.,* 45**,** e722-e739.

Incerti, S., Kyriakou, I. & Tran, H. N. 2017, Geant4-DNA simulation of electron slowing-down spectra in liquid water. *Nucl. Instrum. Meth. B,* 397**,** 45-50.

Incerti, S., Zhang, Q., Andersson, F., et al. 2007, Monte Carlo simulation of the CENBG microbeam and nanobeam lines with the Geant4 toolkit. *Nucl. Instrum. Meth. B,* 260**,** 20-27.

Ito, A. 1987. Calculation of double strand break probability of DNA for low LET radiations based on track structure analysis. *Nuclear and Atomic Data for Radiotherapy and Related Radiobiology.* International Atomic Energy Agency (IAEA).

Kaplan, I. G., Miterev, A. M. & Sukhonosov, V. Y. 1990, Simulation of the primary stage of liquid water radiolysis. *Radiat. Phys. Chem.,* 36**,** 493-498.

Karamitros, M. 2012. *Extension de l'outil Monte Carlo généraliste Geant4 pour la simulation de la radiolyse de l'eau dans le cadre du projet Geant4-DNA.* Ph.D. thesis, Université Bordeaux 1.

Karamitros, M., Brown, J. M. C., Lampe, N., et al. 2020, Implementing the independent reaction time method in Geant4 for radiation chemistry simulations. *arXiv preprint arXiv:2006.14225*.

Karamitros, M., Luan, S., Bernal, M. A., et al. 2014, Diffusion-controlled reactions modeling in

Geant4-DNA. *J. Comput. Phys.,* 274**,** 841-882.

Karamitros, M., Mantero, A., Incerti, S., et al. 2011, Modeling radiation chemistry in the Geant4 toolkit. *Prog. Nucl. Sci. Tech.,* 2**,** 503-508.

Khanna, K. K. & Jackson, S. P. 2001, DNA double-strand breaks: signaling, repair and the cancer connection. *Nat. Genet.,* 27**,** 247-254.

Knight, F. B. 1962, On the random walk and Brownian motion. *T. Am. Math. Soc.,* 103**,** 218-228.

Krämer, M. & Kraft, G. 1994, Calculations of heavy-ion track structure. *Radiat. Environ. Bioph.,* 33**,** 91-109.

Kreipl, M. S., Friedland, W. & Paretzke, H. G. 2009, Time- and space-resolved Monte Carlo study of water radiolysis for photon, electron and ion irradiation. *Radiat. Environ. Bioph.,* 48**,** 11-20.

Kyriakou, I., Emfietzoglou, D., Ivanchenko, V., et al. 2017, Microdosimetry of electrons in liquid water using the low-energy models of Geant4. *J. Appl. Phys.,* 122.

Kyriakou, I., Incerti, S. & Francis, Z. 2015, Technical Note: Improvements in Geant4 energy-loss model and the effect on low-energy electron transport in liquid water. *Med. Phys.,* 42**,** 3870-6.

Kyriakou, I., Šefl, M., Nourry, V., et al. 2016, The impact of new Geant4-DNA cross section models on electron track structure simulations in liquid water. *J. Appl. Phys.,* 119.

Lampe, N. 2017. *The long term impact of ionising radiation on living systems.* Ph.D. thesis, Université Clermont Auvergne, France.

Lampe, N., Biron, D. G., Brown, J. M. C., et al. 2016, Simulating the impact of the natural radiation background on bacterial systems: Implications for very low radiation biological experiments. *PLoS ONE,* 11**,** e0166364.

Lampe, N., Karamitros, M., Breton, V., et al. 2018a, Mechanistic DNA damage simulations in Geant4-DNA part 1: A parameter study in a simplified geometry. *Phys. Med.,* 48**,** 135-145.

Lampe, N., Karamitros, M., Breton, V., et al. 2018b, Mechanistic DNA damage simulations in Geant4-DNA Part 2: Electron and proton damage in a bacterial cell. *Phys. Med.,* 48**,** 146-155.

Lappa, A. V., Bigildeev, E. A., Burmistorv, D. S., et al. 1993, "Trion" code for radiation action calculations and its application in microdosimetry and radiobiology. *Radiat. Environ. Bioph.,* 32**,** 1-19.

Larsen, E. W. & Tolar, D. R. A "Transport" Condensed History Method. *In:* KLING, A., BARÄO, F. J. C., NAKAGAWA, M., TÁVORA, L. & VAZ, P., eds. Advanced Monte Carlo for Radiation Physics, Particle Transport Simulation and Applications,

2001 Berlin, Heidelberg. Springer, 49-54.

Lazarakis, P., Incerti, S., Ivanchenko, V., et al. 2018, Investigation of track structure and condensed history physics models for applications in radiation dosimetry on a micro and nano scale in Geant4. *Biomed. Phys. Eng. Express,* 4.

Li, J., Li, C., Qiu, R., et al. 2015, DNA strand breaks induced by electrons simulated with Nanodosimetry Monte Carlo Simulation Code: NASIC. *Radiat. Prot. Dosim.,* 166**,** 38-43.

Melton, C. E. 1972, Cross sections and interpretation of dissociative attachment reactions producing OH−, O−, and H− in H2O. *J. Chem. Phys.,* 57**,** 4218-4225.

Metropolis, N. 1987, The beginning of the Monte Carlo method. *Los Alamos Science,* 15**,** 125-130.

Meylan, S., Incerti, S., Karamitros, M., et al. 2017, Simulation of early DNA damage after the irradiation of a fibroblast cell nucleus using Geant4-DNA. *Sci. Rep.,* 7**,** 11923.

Meylan, S., Vimont, U., Incerti, S., et al. 2016, Geant4-DNA simulations using complex DNA geometries generated by the DnaFabric tool. *Comput. Phys. Commun.,* 204**,** 159-169.

Michalik, V., Begusova, M. & Bigildeev, E. A. 1998, Computer-aided stochastic modeling of the radiolysis of liquid water. *Radiat. Res.,* 149**,** 224-236.

Michaud, M., Wen, A. & Sanche, L. 2003, Cross sections for low-energy (1–100 eV) electron elastic and inelastic scattering in amorphous ice. *Radiat. Res.,* 159**,** 3-22.

Mitchell, H. H., Hamilton, T. S., Steggerada, F. R., et al. 1945, The chemical composition of the adult human body and its bearing on the biochemistry of growth. *J. Biol. Chem.,* 158**,** 625-637.

Mortazavi, S. M. J., Cameron, J. R. & Niroomand-rad, A. 2003, Adaptive response studies may help choose astronauts for long-term space travel. *Adv. Space Res.,* 31**,** 1543-1551.

Nagasawa, H. & Little, J. B. 1992, Induction of sister chromatid exchanges by extremely low doses of α-particles. *Cancer Res.,* 52**,** 6394-6396.

NCRP 2009. Ionizing radiation exposure of the population of the United States. *NCRP Report 160.* (Bethesda, MD:NCRP).

Nias, A. H. W. 1998. *An introduction to radiobiology,* Chichester ;New York :, Wiley.

Nikjoo, H., Charlton, D. E. & Goodhead, D. T. 1994, Monte Carlo track structure studies of energy deposition and calculation of initial DSB and RBE. *Adv. Space Res.,* 14**,** 161-180.

Nikjoo, H., Emfietzoglou, D., Liamsuwan, T., et al. 2016, Radiation track, DNA damage and

response-a review. *Rep. Prog. Phys.,* 79**,** 116601.

Nikjoo, H. & Girard, P. 2012, A model of the cell nucleus for DNA damage calculations. *Int. J. Radiat. Biol.,* 88**,** 87-97.

Nikjoo, H., Uehara, S., Emfietzoglou, D., et al. 2006, Track-structure codes in radiation research. *Radiat. Meas.,* 41**,** 1052-1074.

O'Neill, P. & Wardman, P. 2009, Radiation chemistry comes before radiation biology. *Int. J. Radiat. Biol.,* 85**,** 9-25.

Paganetti, H., Niemierko, A., Ancukiewicz, M., et al. 2002, Relative biological effectiveness (RBE) values for proton beam therapy. *Int. J. Radiat. Oncol.,* 53**,** 407-421.

Paretzke, H. G. 1987, Radiation track structure theory. *Kinetics of Non-homogeneous Processes*.

Pelowitz, D. B. 2011, MCNPX user's manual version 2.7.0-LA-CP-11-00438. *Los Alamos National Laboratory*.

Perl, J., Shin, J., Schumann, J., et al. 2012, TOPAS: an innovative proton Monte Carlo platform for research and clinical applications. *Med. Phys.,* 39**,** 6818-37.

Pham, Q. T., Anne, A., Bony, M., et al. 2015, Coupling of Geant4-DNA physics models into the GATE Monte Carlo platform: Evaluation of radiation-induced damage for clinical and preclinical radiation therapy beams. *Nucl. Instrum. Meth. B,* 353**,** 46-55.

Pimblott, S. M., LaVerne, J. A., Mozumder, A., et al. 1990, Structure of electron tracks in water. 1. Distribution of energy deposition events. *J. Phys. Chem.,* 94.

Plante, I. & Cucinotta, F. A. 2009, Cross sections for the interactions of 1 eV–100 MeV electrons in liquid water and application to Monte-Carlo simulation of HZE radiation tracks. *New J. Phys.,* 11.

Plante, I. & Devroye, L. 2017, Considerations for the independent reaction times and step-by-step methods for radiation chemistry simulations. *Radiat. Phys. Chem.,* 139**,** 157-172.

Prasad, K. N., Cole, W. C. & Haase, G. M. 2004, Radiation protection in humans: extending the concept of as low as reasonably achievable (ALARA) from dose to biological damage. *Brit. J. Radiol.,* 77**,** 97-9.

Ramos-Méndez, J., Domínguez-Kondo, N., Schuemann, J., et al. 2020, LET-Dependent Intertrack Yields in Proton Irradiation at Ultra-High Dose Rates Relevant for FLASH Therapy. *Radiat. Res.,* 194**,** 351-362.

Ramos-Mendez, J., Perl, J., Schuemann, J., et al. 2018, Monte Carlo simulation of chemistry following radiolysis with TOPAS-nBio. *Phys. Med. Biol.,* 63**,** 105014.

Rudd, M. E., Goffe, T. V., DuBois, R. D., et al. 1985, Cross sections for ionization of water vapor

by 7-4000-keV protons. *Phys. Rev. A,* 31**,** 492-494.

Sakata, D., Belov, O., Bordage, M. C., et al. 2020, Fully integrated Monte Carlo simulation for evaluating radiation induced DNA damage and subsequent repair using Geant4-DNA. *Sci. Rep.,* 10**,** 20788.

Sanz, A. G., Fuss, M. C., Munoz, A., et al. 2012, Modelling low energy electron and positron tracks for biomedical applications. *Int. J. Radiat. Biol.,* 88**,** 71-6.

Schuemann, J., McNamara, A. L., Ramos-Mendez, J., et al. 2019, TOPAS-nBio: An Extension to the TOPAS Simulation Toolkit for Cellular and Sub-cellular Radiobiology. *Radiat. Res.,* 191**,** 125-138.

Sefl, M., Incerti, S., Papamichael, G., et al. 2015, Calculation of cellular S-values using Geant4-DNA: The effect of cell geometry. *Appl. Radiat. Isotopes,* 104**,** 113-23.

Semenenko, V., Turner, J. & Borak, T. 2003, NOREC, a Monte Carlo code for simulating electron tracks in liquid water. *Radiat. Environ. Bioph.,* 42**,** 213-217.

Seymour, C. B. & Mothersill, C. 2004, Radiation-induced bystander effects-implications for cancer. *Nat. Rev. Cancer,* 4**,** 158-164.

Tang, N. 2019. *Évaluation, à partir de modélisations nanodosimétriques, de l'influence de la compaction de la chromatine sur les effets radio-induits précoces et extension aux effets tardifs (réparation des dommages à l'ADN et mort cellulaire).* Ph.D. thesis, Université de Bordeaux, France.

Terrissol, M. & Beaudre, A. 1990, Simulation of space and time evolution of radiolytic species induced by electrons in water. *Radiat. Prot. Dosim.,* 31**,** 175-177.

Tomita, H., Kai, M., Kusama, T., et al. 1997, Monte Carlo simulation of physicochemical processes of liquid water radiolysis. *Radiat. Environ. Bioph.,* 36**,** 105-116.

Turner, J. E. 2007. Chemical and Biological Effects of Radiation. *Atoms, Radiation, and Radiation Protection.* Wiley.

Turner, J. E., Magee, J. L., Wright, H. A., et al. 1983, Physical and chemical development of electron tracks in liquid water. *Radiat. Res.,* 96**,** 437-449.

Uehara, S., Nikjoo, H. & Goodhead, D. T. 1993, Cross-sections for water vapour for the Monte Carlo electron track structure code from 10 eV to the MeV region. *Phys. Med. Biol.,* 37**,** 1841-1858.

UNSCEAR 1994. Sources and effects of ionizing radiation. *UNSCEAR Report 1994.* New York.

UNSCEAR 2000. Sources and effects of ionizing radiation. *UNSCEAR Report 2000.* New York.

Wiklund, K., Fernandez-Varea, J. M. & Lind, B. K. 2011, A Monte Carlo program for the

analysis of low-energy electron tracks in liquid water. *Phys. Med. Biol.,* 56**,** 1985-2003.

Wilson, W. E., Miller, J. H., Lynch, D. J., et al. 2004, Analysis of low-energy electron track structure in liquid water. *Radiat. Res.,* 161**,** 591-596.

Wilson, W. E. & Nikjoo, H. 1999, A Monte Carlo code for positive ion track simulation. *Radiat. Environ. Bioph.,* 38**,** 97-104.

Wolff, S. 1998, The adaptive response in radiobiology: evolving insights and implications. *Environ. Health Persp.,* 106**,** 277-283.

Zaider, M., Brenner, D. J. & Wilson, W. E. 1983, The applications of track calculations to radiobiology I. Monte Carlo simulation of proton tracks. *Radiat. Res.,* 95**,** 231-247.

Chapter 2

Physical stage

Table of contents

2.1. Development of a new electron elastic scattering cross-section model for Geant4-DNA using ELSEPA for liquid-phase water……...31
2.1.1. Description of potentials and selected options……..……...33
2.1.2. Verification of validity by comparing calculated cross-section with experiments and other cross-section models…….……..38
2.1.3. Results for determination of optimal options and parameters……..……40
2.1.4. Plausibility of the electron elastic scattering models……...46
2.2. The impact of the ELSEPA electron elastic scattering model on Geant4-DNA simulations……...48
2.2.1. Determination of optimal angle and energy binning in ELSEPA……48
2.2.2. Geant4-DNA examples……...49
2.2.3. Results for the implementation of the elastic scattering cross-section into Geant4-DNA……...55
2.2.4. Results for the Geant4-DNA simulations……59
2.3. Conclusions……66
References……68

In order to investigate the induction of biological damage from ionizing radiation at the sub-cellular scale, MCTS simulation codes have been developed for several decades (Nikjoo et al., 1997, Uehara et al., 1999, Semenenko et al., 2003, Wilson et al., 2004, Nikjoo et al., 2006, Tung et al., 2007, Incerti et al., 2010a, Friedland et al., 2011, Alloni et al., 2012). These codes usually approximate the biological medium as liquid water, which composes more than 60% of human body (Mitchell et al., 1945). Since most physical damages are caused by secondary electrons, many sets of electron cross-section models for liquid water have been developed so far, for example, see a selection in Dingfelder et al. (1998), Emfietzoglou et al. (2005), Champion et al. (2009), Bordage et al. (2016), Garcia-Molina et al. (2017). In particular, it was reported that the interactions of low energy electrons below 100 eV should be carefully modeled for the prediction of damages to the DNA molecule, which is induced mainly through ionization (Nikjoo et al., 2016). At such low energy, elastic scattering also plays a key role; even if this process is not associated with significant energy loss, it allows to accurately describe the spatial distribution of electrons.

To accurately calculate the elastic scattering cross-section of electrons in liquid water, three main approaches have been proposed: the Born collision model (Mott and Massey, 1965), the non-relativistic (Schrödinger) partial wave model (Schiff, 1968), the relativistic (Dirac) partial wave analysis (Vanderpoorten, 1975), and other methods such as the Schwinger multichannel method with pseudopotentials at even lower energies (Varella et al., 1999). It has been reported that the Dirac partial wave method is today the most accurate to calculate such cross-section (Staszewska et al., 1984). Unfortunately, the validation of these calculations, especially elastic scattering in liquid water for low energy electrons, is still not possible due to the scarcity of reliable scattering cross-section data (IAEA, 1995, Incerti et al., 2010b).

However, at least over an incident energy of 60 eV, the calculated differential cross-sections (DCSs) between liquid water and vapour water show good agreement in the entire angle range with discrepancies of up to 1.4 times at 0 deg (Aouchiche et al., 2008). For this reason, the plausibility of calculated DCSs is typically verified by comparison with experimental data in the vapour-phase of water.

In this work, we propose to improve the default electron elastic scattering model of Geant4-DNA, initially developed by Champion et al. (2009), which presents several limitations (Champion, 2003):

- This model does not include relativistic corrections;
- At low incident energies, especially below 60 eV, the DCS present too pronounced minima at intermediate scattering angles comparing with experimental data;
- At small scattering angle below 20 deg, the DCS appears underestimated compared with experimental data;

For this, we performed the calculations using the ELSEPA (ELastic Scattering of Electrons and Positrons by neutral Atoms) code, developed by Salvat et al. (2005). This code uses a partial wave approach including relativistic corrections (Dirac partial wave analysis) to calculate electron elastic cross-sections. We used the new version of ELSEPA based on the recently published corresponding paper and kindly provided by the authors of Bote et al. (2009). The advantage of ELSEPA is that one can easily change calculation parameters and models of interaction potentials because ELSEPA is an open-source program written in Fortran 77, and it is possible to calculate the DCSs from a few eV up to 1 GeV in a variety of materials.

In contrast with the existing elastic cross-section models available in Geant4-DNA, we first propose to optimize the phenomenological optical parameters included in the correlation-polarizability potential and in the absorption potential based on the experimental data in vapour-phase water. In parallel, ICRU 77 recommendations are calculated only with default values of optical parameters (ICRU, 2007). After that, the Muffin-tin approximation, typically employed to predict interaction in solid phase material, is employed to calculate the elastic cross-section in liquid-phase water in section 2.1. We compare the cross-sections generated by ELSEPA for the liquid-phase water in the energy range (10 eV - 1 MeV) with the various cross-section models already available in Geant4-DNA and also with reference experimental data measured in the vapour-phase of water.

In section 2.2, the new electron elastic-cross-section developed in the previous section is implemented into Geant4-DNA. For that, the effects of angle and energy bins are evaluated for the simulation efficiency without under-sampling issues. In order to verify the impact of this new model on MCTS simulation, Geant4-DNA application examples are utilised and the results are compared with reference data. In this study, we use 5 examples: *range*, *mean-free-path*, *TestEm12* (for dose-point-kernels), *microyz* (for microdosimetric distributions), and *clustering* (for a rapid estimation of direct DNA damage). Simulations are run with the existing Geant4-DNA physics constructors and with the same constructors where the elastic model has been replaced by the newly implemented model calculated with ELSEPA.

2.1. Development of a new electron elastic scattering cross-section model for Geant4-DNA using ELSEPA for liquid-phase water

In order to accurately calculate cross-sections between electron and matter, the interaction potential model should be first described. The effective interaction between a projectile at distance r from the target is assessed by a summation of potentials:

$$V(r) = V_{st}(r) + V_{ex}(r) + V_{cp}(r) - iW_{abs}(r) \quad (2.1)$$

where $V_{st}(r)$ is the electrostatic interaction potential, $V_{ex}(r)$ is the exchange potential, $V_{cp}(r)$ is the correlation-polarizability potential, and $W_{abs}(r)$ is the magnitude of the imaginary absorption potential. In ELSEPA, users can select the models for the interaction potentials and can use various options listed in Table 2.1. The details of each model will be described below.

Table 2.1: Selectable options (first row), corresponding names (second row) and available models (other rows) available in ELSEPA. To propose two new elastic models, the options in red and blue are used as first choice (or "default" choice) and as second choice (or "alternative" choice), respectively.

Nuclear charge distribution	**Electron distribution**	**Exchange potential**	**Correlation-polarizability potential**	**Absorption potential**	**Phase**	**Coherence**
Point nucleus (P)	Thomas-Fermi-Molière (TFM)	No exchange potential	No correlation-polarizability potential	No absorption potential	Free atom approximation	Additivity (incoherence)
Uniform (U)	Thomas-Fermi-Dirac (TFD)	Furness-McCarthy (FM)	Buckingham (B)	Local density approximation (LDA)	Muffin-tin model	Independent-atom (coherence)
Fermi (F)	Dirac-Hartree-Fock-Slater (DHFS)	Thomas-Fermi (TF)	Local density approximation (LDA)	-	-	-
Helm's uniform-uniform (U_u)	Relativistic Dirac-Fock (DF)	Riley-Truhlar (RT)	-	-	-	-

2.1.1. Description of potentials and selected options

There are several approaches describing the interaction potentials, and it is well-known that the electrostatic and exchange potentials agree well with experimental data in the high energy range above the electron energy of 5 keV. However, the accuracy of the static field and static exchange approximations gets worse below this energy (Salvat, 2003). For correction of the disagreement, the optical models including correlation-polarizability and absorption potentials were proposed, and ELSEPA provides the best empirical parameters for the optical model which were validated with available experimental data in the vapour-phase. However, the optical parameters are basically phenomenological, and in our case, it is mandatory to optimize these parameters for the liquid-phase. We describe in this section not only the potentials and options used in this work but also the optimization of the optical model.

Electrostatic potential

The electrostatic interaction potential between the projectile and a target atom, $V_{st}(r)$, can be expressed as:

$$V_{st}(r) = Z_0 e\varphi(r) = Z_0 e[\varphi_n(r) + \varphi_e(r)] \tag{2.2}$$

where $Z_0 e$ is the charge of the projectile and $\varphi_n(r)$ and $\varphi_e(r)$ are the electrostatic potentials of the target atom originating from the nucleus and from the electrons, respectively. Especially for high energy electrons, the elastic cross-section is affected by the distribution of atomic electrons. These potentials contributing from the nucleus $\varphi_n(r)$ or from the electrons $\varphi_e(r)$ were described by Salvat et al. (2005):

$$\varphi_n(r) = \frac{1}{r}\int_0^r \rho_n(s)4\pi s^2 ds + \int_r^\infty \rho_n(s)4\pi s ds \tag{2.3}$$

$$\varphi_e(r) = -\left(\frac{1}{r}\int_0^r \rho_e(s)4\pi s^2 ds + \int_r^\infty \rho_e(s)4\pi s ds\right) \tag{2.4}$$

where $\rho_{n/e}$ represents the spatial density of protons in the nucleus and orbital electrons, respectively.

The model of charge distribution within the nucleus is also important to calculate the electrostatic potential. The cross-sections calculated with the point model (P) or the uniformly charged sphere model (U) simplifying the distribution of protons might differ substantially from the results of high energy electron elastic scattering experiments (Salvat and Mayol, 1993). Alternatively, the proton density $\rho_n(r)$ can be calculated numerically based on the Helm's uniform-uniform distribution (U_u) or Fermi distribution. The proton density $\rho_n(r)$ can be calculated numerically using the Fermi distribution:

$$\rho_{n,F}(r) = \frac{\rho_0}{exp\{(r - R_n)/z\} + 1} \tag{2.5}$$

where the nuclear radius R_n is $1.07\times10^{-13}\ A^{1/3}$ cm with the atomic mass A in the unit of *g/mol*, the constant ρ_0 for normalization is the double value of the proton density at $r = R_n$, and the skin thickness z is 0.546×10^{-13} cm. We selected the Fermi distribution which is the default option of the nuclear charge distribution in this study because the influence of the proton density is only visible for projectile energies above about 10 MeV.

The simplest theoretical method to obtain approximate atomic electron densities is the Thomas-Fermi (TF) model which considers the electron cloud as a locally homogeneous electron gas bound by the screened Coulomb field of the nucleus, assumed to be a point charge (Thomas, 1954).

$$\rho_{e,TF}(r) = \frac{Z}{4\pi r}\sum_{i=1}^{3} p_i \exp(-q_i r) \tag{2.6}$$

where p_i and q_i are the parameters determined by fitting a variety of TF screening functions. Several analytical approximations have been proposed such as the Thomas-Fermi-Molière (TFM), the Dirac-Hartree-Fock-Slater (DHFS), and the Thomas-Fermi-Dirac (TFD) approximations (Moliere, 1947, Slater, 1951, Bonham and Strand, 1963). However, these approximations are based on the homogeneous electron gas bound by the screened Coulomb field of the nucleus, assumed to be a point charge. Moreover, relativistic treatment is not considered. We selected instead the relativistic Dirac-Fock (DF) model in this study because it is reported to be the most realistic approximation of electron density (ρ_e) model available for free atoms (Desclaux, 1973). ELSEPA calculates the electrostatic potential originating from the orbital electrons using a database of ρ_e generated by a DF program (Desclaux, 1975, Desclaux, 1977).

Exchange potential

The exchange potential indicates that the electron projectile can exchange localization with an atomic electron. Thomas-Fermi (TF) is a generalization of Slater's potential (Slater, 1951) often used in non-relativistic bound-state calculations (Thomas, 1954). This model assumes a simple local exchange potential to describe the atomic orbitals, and approximates the scattering function in the exchange term by a plane wave with a suitable local wavenumber. An alternative exchange potential was initially proposed by Furness and McCarthy (1973) following the expansion method of Perey and Buck, and corrected by Riley and Truhlar (1975). In this study, we choose the Furness-McCarthy (FM) potential which is the default exchange potential in ELSEPA; it includes the possibility of rearrangement collisions, in which the projectile exchanges places with an atomic electron. The exchange potential of the FM model $V_{ex,FM}(r)$ is given by:

$$V_{ex,FM}(r) = \frac{1}{2}(E - V_{st}(r)) - \frac{1}{2}[(E - V_{st}(r))^2 + 4\pi a_0 e^4 \rho_e(r)]^{1/2} \tag{2.7}$$

where a_o is the Bohr radius of 5.291772×10^{-9} cm, E is the energy of the electron in Hartree atomic unit, and $\rho_e(r)$ is the electron density.

Correlation-polarizability potential

The correlation-polarizability potential has an influence on the DCSs at low angle, and the effect is larger especially below 500 eV (Salvat, 2003). ELSEPA provides two possibilities: the Buckingham potential and the local density approximation (LDA). The Buckingham potential is a correlation-polarizability potential based on a phenomenological formula deriving from a measured polarizability of the atom (Buckingham, 1938). This approach is realistic at large distances, however when the projectile is close to the atom, it is not suitable. Because of the above reason, we choose the LDA potential which is a combination of the correlation-polarizability potential model for the long-range and an independent correlation potential for the short-range trajectories (O'Connell and Lane, 1983). The LDA combination can be described by:

$$V_{cp,LDA}^{-}(r) \equiv \begin{cases} \max\{V_{co}^{-}(r), V_{cp}(r)\}, & if\ r < r_{cp} \\ V_{cp}(r), & if\ r \geq r_{cp} \end{cases} \tag{2.8}$$

where $V_{co}^{-}(r)$ and $V_{cp}(r)$ are the short-range correlation potential and long-range correlation-polarizability potential of the electron for a trajectory of range r. r_{cp} is the outer radius at which $V_{co}^{-}(r)$ and $V_{cp}(r)$ cross.

There are two models for the correlation potential $V_{co}^{-}(r)$:

- In the correlation potential suggested by Perdew and Zunger (1981), $V_{co,PZ}^{-}(r)$ is described by:

$$V_{co,PZ}^{-}(r) = \begin{cases} -\dfrac{e^2}{a_0}(0.0311\ln r_s - 0.0584 + 0.001333 r_s \ln r_s - 0.084 r_s), & r_s < 1 \\ -0.1423\dfrac{e^2}{a_0}\dfrac{1 + (7/6)1.0529 r_s^{1/2} + (4/3)0.3334 r_s}{\left(1 + 1.0529 r_s^{1/2} + 0.3334 r_s\right)^2}, & r_s \geq 1 \end{cases} \tag{2.9}$$

- In the correlation potential suggested by Padial and Norcross (1984), $V_{co,PN}^{-}(r)$ is:

$$V_{co,PN}^{-}(r) = \begin{cases} -\dfrac{e^2}{a_0}(0.0311\ln r_s - 0.0584 + 0.006 r_s \ln r_s - 0.015 r_s), & r_s \leq 0.7 \\ -\dfrac{e^2}{a_0}(0.07356 + 0.02224\ln r_s), & 0.7 \leq r_s \leq 10 \\ -\dfrac{e^2}{a_0}(0.584 r_s^{-1} + 1.988 r_s^{-\frac{3}{2}} - 2.450 r_s^{-2} - 0.733 r_s^{-\frac{5}{2}}), & r_s \geq 10 \end{cases} \tag{2.10}$$

with the radius of the sphere which contains one electron of the gas r_s defined as:

$$r_s = \left(3/4\pi\rho(r)\right)^{1/3} \tag{2.11}$$

Regarding the correlation-polarizability potential $V_{cp}(r)$ in equation (2.8), we compared the two models of Buckingham (1938), Lindhard (1954):

- The correlation-polarizability potential of Lindhard $V_{cp,L}(r)$ is based on the Lindhard's high-energy formula. It is provided as default option of the LDA potential for the solid-phase materials in the last version of ELSEPA (Bote et al., 2009) and it can be described by:

$$V_{cp,L}(r) = \frac{\pi m_e e^2}{4\hbar k}\omega_p \tag{2.12}$$

with the linear momentum $\hbar k$ and the plasmon energy ω_p.

- The correlation-polarizability potential of Buckingham $V_{cp,B}(r)$ is:

$$V_{cp,B}(r) = \frac{\alpha_p e^2}{2(r^2 + d^2)^2} \tag{2.13}$$

where α_p represents the polarizability of the target atom or molecule. In the case of a molecule, the effective molecular polarizability is formulated as:

$$\alpha_{\mathrm{p,eff}}(i) = \alpha_{\mathrm{p}}^{(\mathrm{mol})} \alpha_{\mathrm{p}}(i) \left[\sum_i \alpha_{\mathrm{p}}(i) \right]^{-1} \tag{2.14}$$

We used the polarizability of 1.457 Å^3 proposed by ICRU (2007). The d in equation (2.13) indicates a phenomenological cut-off parameter:

$$d^4 = \frac{1}{2} \alpha_p a_0 Z^{-1/3} b_{pol}^2 \tag{2.15}$$

In equation (2.15), the adjustable energy-dependent parameter b_{pol} is suggested by default for noble gases (Salvat, 2003):

$$b_{pol}^2 = \max\left[(E - 50\ eV)/(16\ eV), 1\right] \tag{2.16}$$

However, the default expression of b_{pol} is too low to accurately calculate the DCSs for liquid water. In this study, we qualitatively optimize b_{pol} by comparing with experimental data. It is reported that b_{pol} primarily effects low angles. The DCSs at low angles are the largest when b_{pol} is 1, and they decrease with the increase of b_{pol} (Salvat, 2003).

Absorption potential

To take into account the energy loss by the inelastic scattering, the inelastic absorption potential must contain a negative imaginary part in equation (2.1). The absorption potential has an influence at intermediate and high scattering angles. The absorption strength is large when the energy of electron is high because the inelastic scattering is more frequent at high energy. The absorption potential W_{abs} with relativistic consideration can be described by:

$$W_{abs} = \sqrt{\frac{2(E_L + m_e c^2)^2}{m_e c^2 (E_L + 2m_e c^2)}} A_{abs} \hbar \frac{1}{2} (2E_L/m_e)^{1/2} \times \rho \sigma_{bc}(E_L, \rho, \Delta) \tag{2.17}$$

where $m_{\mathrm{e}}c^2$ and $\hbar$ are the rest mass energy of an electron and Planck constant, respectively. The local kinetic energy E_{L} is the kinetic energy except the electrostatic and exchange potential, and the $\sigma_{bc}(E_L, \rho, \Delta)$ is the cross-section for binary collisions of the projectile with the local free-electron gas calculated from the non-relativistic Born approximation, and we recall that the quantity $(2/\hbar)W_{abs}$ represents the absorption probability per unit time (Salvat, 2003).

In ELSEPA, there are two empirical parameters, the absorption strength A_{abs} and the energy gap Δ. The absorption strength A_{abs} is suggested by Salvat (2003), Bote et al. (2009), as OP-I (2.00) or OP-II (0.75), and has been validated by comparisons with experiments in noble gases. However, it is not suitable for liquid nor vapour water. The energy gap Δ is the minimum energy at which the energy loss by inelastic scattering starts. For electron, Δ is set equal to the first inelastic threshold

ε_1 of the target. In the case of liquid water, we used 8.22 eV which is the threshold of the first excitation level in Geant4-DNA's Emfietzoglou model (Emfietzoglou and Moscovitch, 2002).

Phase influence

ELSEPA provides a potential model in the free atom approximation for the vapour-state phase and the Muffin-tin approximation adapted to the solid-state phase (Yates, 1968, Czyżewski et al., 1990) (see 6th column of Table 2.1). In contrast of the free atom approximation using equation (2.1), the Muffin-tin potential assumes that the electrons of the neutral atom are confined within a sphere of a certain radius R_{mt} centered on the nucleus to approximate solid-state. This assumption implies that at distances $r > R_{mt}$ the electrostatic interaction potential $V_{st}(r)$ of a bound atom is smaller (in magnitude) than that of a free atom:

$$V_{mt}(r) = \begin{cases} V_{st,mt}(r) + V_{ex}(r) + V_{cp}(r) - iW_{abs}(r), & r \le R_{mt} \\ V_{st,mt}(R_{mt}) + V_{ex}(R_{mt}) + V_{cp}(r) - iW_{abs}(r), & r > R_{mt} \end{cases} \quad (2.18)$$

The main difference between Muffin-tin and free atom approximation is that:

$$V_{st,mt}(r) = V_{st}(r) + V_{st}(2R_{mt} - r) \quad (2.19)$$

In this work, we propose to use the Muffin-tin potential to mimic the liquid-phase of water. We choose the radius of the Muffin-tin model R_{mt} as half the average distance between oxygen atoms in liquid water, 1.405 Å (Liljequist et al., 2012), instead of the water molecule radius, 0.991 Å. It must be noted that the Muffin-tin approximation basically vanishes the absorption potential $iW_{abs}(r)$ outside the Muffin-tin sphere because it is valid for solid-phase material which has a lack of free electrons mainly causing inelastic scattering (Salvat et al., 2005). However, we used the absorption potential outside the Muffin-tin sphere with the assumption that there are sufficiently free electrons in the liquid-phase of water such as solvated electrons. In addition, the free atom approximation is employed as an alternative option for the evaluation of the effects of water phase by comparing results obtained with the free atom approach with those of the Muffin-tin approach.

Coherence between individual atoms in condensed matter

In biological medium such as liquid water, the most difficult issues to calculate the cross-sections are the effects of the condensed phase, which originate from the long-range polarizability of the medium by the incident charged-particle field (Dingfelder, 2006). In radiation transport codes, the DCSs for electron elastic scattering in molecules are usually calculated as a summation of the DCSs of the individual atoms that make up the molecule, assuming incoherence, called "additivity approximation". However, this additivity approximation neglects the chemical binding and aggregation effects which refer to the influence of relative positions of individual atoms on multiple-wave scattering (Zecca et al., 1999). On the other hand, the "independent-atom approximation" considers the intramolecular multiple scatterings (Zeitler and Olsen, 1967). In this approximation, the DCS of a molecule is obtained by adding coherently the projectile waves

scattered by the different atoms in the molecule. We chose the independent-atom approximation which coherently considers the aggregation effects.

2.1.2. Verification of validity by comparing calculated cross-section with experiments and other cross-section models.

For the validation of the calculated cross-sections, we compared our values with cross-sections currently available in Geant4-DNA for simulations in liquid water and with vapour water experiments published in the literature. Table 2.2 lists the different models of electron elastic scattering already available in Geant4-DNA, as well as the ELSEPA model presented in this work.

Table 2.2: List of Geant4-DNA physics models for the simulation of elastic scattering of electrons in liquid water.

Geant4-DNA physics constructor	Model	Low energy limit	High energy limit	Model type
Option "default", 2, 3	Champion	9 eV	1 MeV	Interpolated
Option 4, 5 alternative #1	Screened Rutherford	9 eV	1 MeV	Analytical
Option 4, 5 alternative #2	Uehara screened Rutherford	9 eV	1 MeV	Analytical
Option 6	CPA100	11 eV	10 keV	Interpolated
This work	ELSEPA	10 eV	1 MeV	Interpolated

The model proposed by Champion et al. (2009) was developed in a partial-wave framework using experimental measurements of electron density auto-correlation functions to calculate cross-section for the liquid water phase. However, the inelastic absorption potential is neglected, as well as relativistic corrections.

Screened Rutherford (SR) and Uehara Screened Rutherford (USR) are analytical models based on the screened Rutherford theory above 200 eV (Grosswendt and Waibel, 1978, Uehara et al., 1993). The DCSs and TCS are calculated by following expressions of screening function:

$$\left(\frac{d\sigma}{d\Omega}\right)_{SR} = \frac{Z(Z+1)r_e^2}{(1-cos\theta+2\eta)^2}\left[\frac{E+m_0c^2}{E(E+2m_0c^2)}\right]^2 \tag{2.20}$$

$$\sigma_t = \frac{\pi Z(Z+1)r_e^2}{\eta(1+\eta)}\left[\frac{E+m_0c^2}{E(E+2m_0c^2)}\right]^2 \tag{2.21}$$

where Z is the atomic number, r_e is the radius of the electron, E is the kinetic energy of electron, and m_0c^2 is the rest mass energy of the electron. The screening parameter η is given by:

$$\eta = \eta_c \times \frac{1.7 \times 10^{-5} Z^{2/3} \times m_0 c^2}{E[E/(m_0 c^2) + 2]} \tag{2.22}$$

with the energy dependence η_c varied by models and energies:

$$\eta_c = \begin{cases} 1.64 - 0.0825 \times \ln E & SR \\ 1.198 & USR, E < 50\ keV \\ 1.13 + 3.76(Z/137)^2 m_0 c^2/(2E) & USR, E \geq 50\ keV \end{cases} \tag{2.23}$$

Below 200 eV, the fitting formulae proposed by Brenner and Zaider (1983) are employed. However, the formulae from Brenner and Zaider below 200 eV provide only DCSs with arbitrary unit. To calculate the absolute DCSs for the comparison with our results, an independent calibration was performed. The DCSs from Brenner and Zaider $\left(\frac{d\sigma}{d\Omega}(E,\theta)\right)_{BZ}$ are calibrated by the formula for total cross-section of SR, $\sigma_{Ruth}(E)$:

$$\left(\frac{d\sigma}{d\Omega}(E,\theta)\right)_{cal,BZ} (cm^2/sr) = \left(\frac{d\sigma}{d\Omega}(E,\theta)\right)_{BZ} (A.U./sr) \frac{\sigma_{Ruth}(E)\ (cm^2)}{\int \left(\frac{d\sigma}{d\Omega}(E,\theta)\right)_{BZ} d\Omega\ (A.U.)}\ (E < 200\ eV) \tag{2.24}$$

The total cross-section σ_t (TCS) can be calculated by integration of the DCS, for example using a trapezoidal integration:

$$\sigma_t = \int \frac{d\sigma}{d\Omega} d\Omega\ (cm^2) = 2\pi \int_0^{\pi} \frac{d\sigma}{d\Omega} sin\theta d\theta \approx \pi \sum_{i=2}^{n} [\left(\frac{d\sigma}{d\Omega}\right)_i sin\theta_i + \left(\frac{d\sigma}{d\Omega}\right)_{i-1} sin\theta_{i-1}] \times (\theta_i - \theta_{i-1}) \tag{2.25}$$

Other quantities of interest are the so-called transport cross-sections, $\sigma_{l^{th}}$, defined by:

$$\sigma_{l^{th}} \equiv \int [1 - P_l(cos\theta)] \frac{d\sigma}{d\Omega} d\Omega \tag{2.26}$$

where $P_l(cos\theta)$ is the Legendre polynomial of degree *l*. Especially, the first transport cross-section, called, momentum-transfer cross-section, (MCS), $\sigma_{1^{st}}$, plays a fundamental role in studies of electron transport based on the diffusion approximation. The first transport cross-section $\sigma_{1^{st}}$ can be calculated with the formula:

$$\sigma_{1^{st}} = \int [1 - cos\theta] \frac{d\sigma}{d\Omega} d\Omega \approx \pi \sum_{i=2}^{n} [\left(\frac{d\sigma}{d\Omega}\right)_i (1 - cos\theta_i) sin\theta_i + \left(\frac{d\sigma}{d\Omega}\right)_{i-1} (1 - cos\theta_{i-1}) sin\theta_{i-1}] \times (\theta_i - \theta_{i-1}) \tag{2.27}$$

At last, in order to verify the validity of the ELSEPA cross-sections, sets of experimental data in the gaseous phase referenced in Hilgner et al. (1969), Danjo and Nishimura (1985), Katase et al. (1986), Sueoka et al. (1986), Szmytkowski (1987), Zecca et al. (1987), Shyn and Cho (1987), Saglam and Aktekin (1991), Johnstone and Newell (1991), Shyn and Grafe (1992), Cho et al. (2004), Khakoo et al. (2013) were compared with the calculated cross-sections.

2.1.3. Results for determination of optimal options and parameters

We choose the Fermi nucleus distribution (F), the Dirac-Fock electron distribution (DF), the Furness-McCarthy exchange potential (FM), the LDA correlation-polarizability potential (LDA), and the LDA absorption potential (A), a combination denoted as (F-DF-FM-LDA-A). However, the parameters of optical-models are phenomenological and were validated with noble gases. For the optimization of the optical parameters in the correlation-polarizability and absorption potentials for liquid-phase water, we evaluated the evolution of the DCSs according to the optical parameters. In this study, we used the Perdew-Zunger correlation potential, the Buckingham correlation-polarizability potential, and 2.5 for the absorption strength as default option.

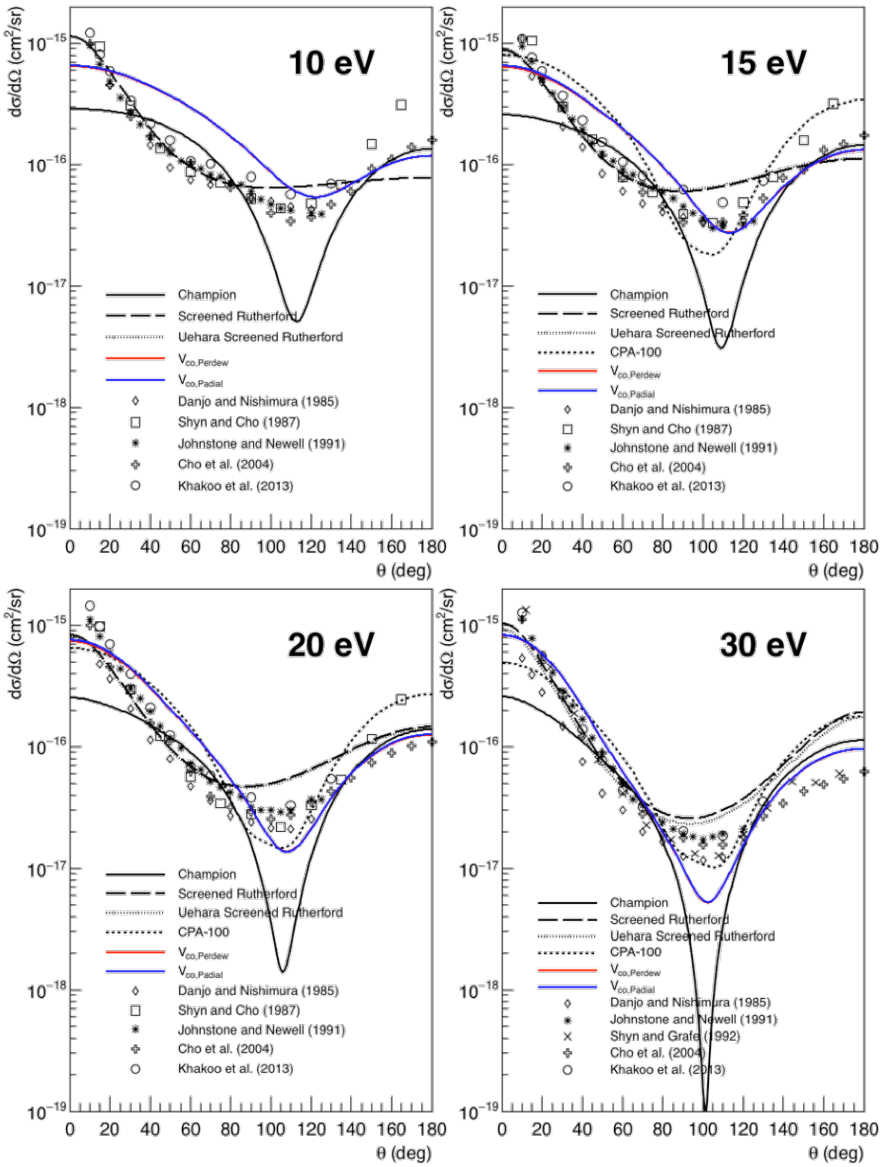


Figure 2.1: DCSs of F-DF-FM-LDA option calculated with the different correlation potentials based on the Perdew-Zunger model (red) or on the Padial-Norcross model (blue). The correlation-polarizability potential V_{cp} is taken as Buckingham.

At first, for the **correlation potential**, we evaluated the effect of the correlation potential ($V_{co}^{-}(r)$) based on the **Perdew-Zunger model** or on the **Padial-Norcross model**, as shown in Figure 2.1 (the correlation-polarizability potential V_{cp} is taken as Buckingham). The differences are negligible, however, the results with the Padial-Norcross model are slightly closer to the experimental data even if ELSEPA provides the Perdew-Zunger model as default. Indeed, at 0 deg, the DCSs of the Padial-Norcross model is higher than Perdew-Zunger model by maximum 2.93% at 20 eV, without differences at intermediate and high scattering angle.

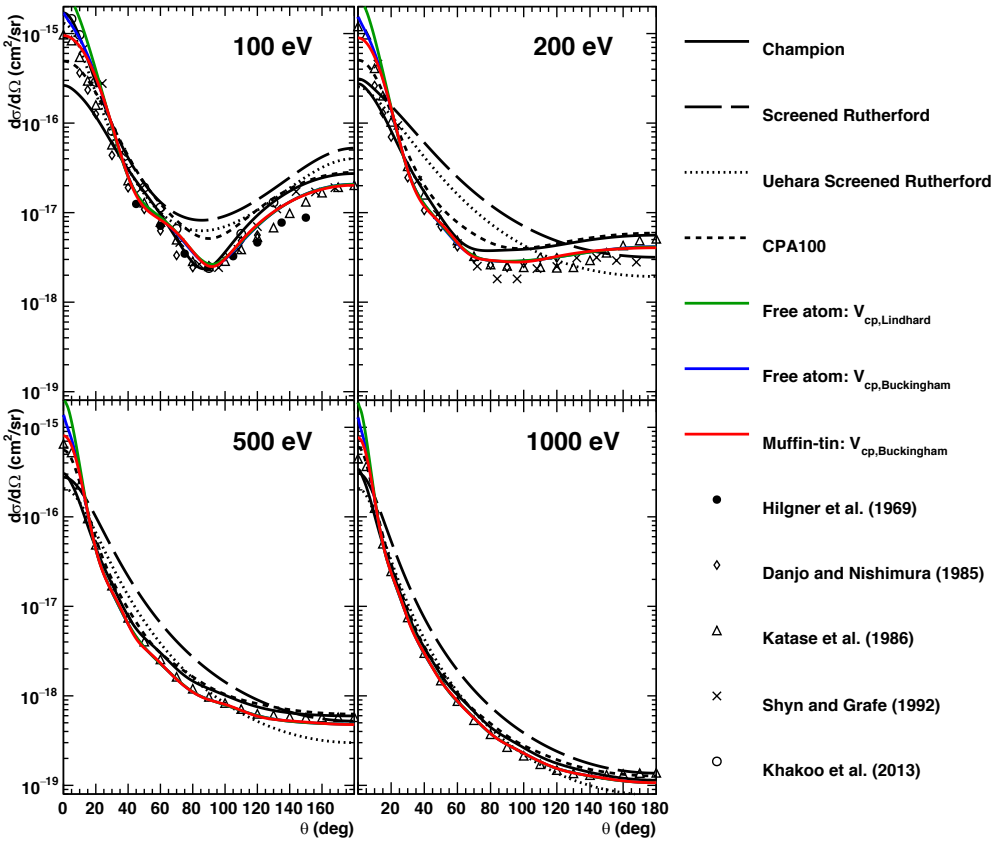


Figure 2.2: DCSs of F-DF-FM-LDA option calculated with the correlation-polarizability potential using the Lindhard model (green), Buckingham model (blue), and the Buckingham model with Muffin-tin approximation (red) from Shin et al. (2018). The correlation potential $V_{co}^{-}(r)$ is taken as Perdew-Zunger.

Second, we have two possibilities: to use the **Lindhard** or the **Buckingham** model for the **correlation-polarizability potential**. Figure 2.2 shows the effect of such models (the correlation potential $V_{co}^{-}(r)$ is taken as Perdew-Zunger). The DCSs with the Lindhard model for the correlation-polarizability potential are much higher than not only the DCSs with Buckingham model but also experiments especially at 0 deg up to 61% at 500 eV, and this tendency is not suitable for our assumption that the DCSs of liquid-phase water is smaller than the DCSs of vapour-phase water at low scattering angle like red curves in the Figure 2.2.

With such results, we decide to select the **Perdew-Zunger model** for the **correlation potential** $\boldsymbol{V_{co}^{-}(r)}$ and the **Buckingham model** for the **correlation-polarizability potential** $\boldsymbol{V_{cp}}$, as components of the LDA combination potential.

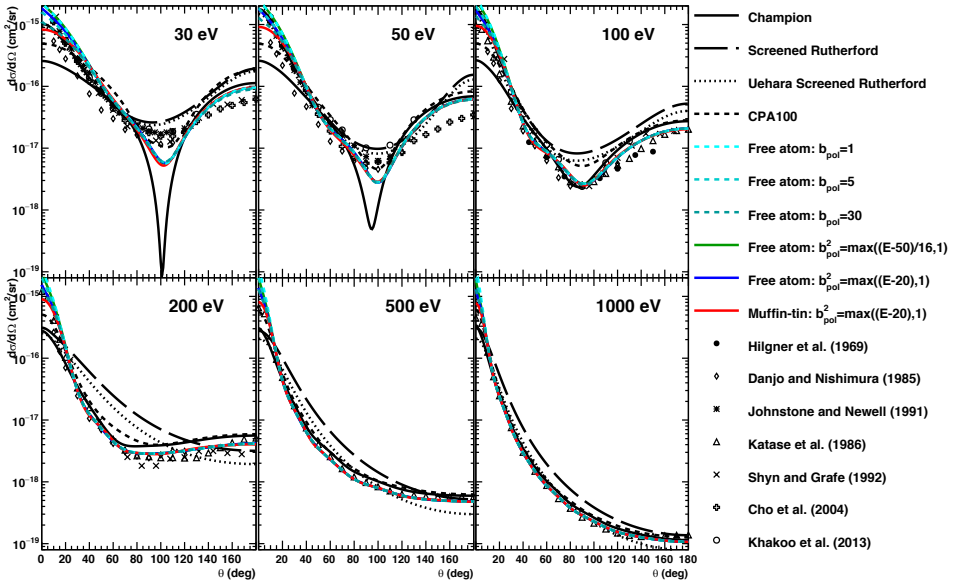


Figure 2.3: DCSs according to the different b_{pol} values in the energy range of 30-1000 eV from Shin et al. (2018). The cyan to green dashed lines indicate the DCSs with b_{pol} of 1 to 30, and the green curves show the DCSs with the default expression of b_{pol} provided in ELSEPA. The blue and red lines represent the DCSs of the proposed expression (2.28) with free atom approximation and Muffin-tin approximation, respectively.

In addition, for the **Buckingham potential**, we have the possibility to change the adjustable energy-dependency of the parameter b_{pol}. Figure 2.3 shows the influence of b_{pol} in the energy range of 30-1000 eV. We can show that the DCSs at low scattering angle is decreased with high b_{pol}, and the tendency rapidly saturates. We qualitatively found the optimal values of b_{pol} according to the electron energy E using the results in Figure 2.3, and we propose to use the following expression:

$$b_{pol}^2 = \max[(E - 20\ eV), 1] \tag{2.28}$$

The green and blue curves show the comparison between the DCSs with the proposed expression of b_{pol} in equation (2.28) and the DCSs with default b_{pol} expression provided by ELSEPA in equation (2.16). b_{pol} is 1 for both of expressions below an electron energy of 21 eV. Above 21 eV, the DCSs with the default energy-dependency parameter are higher than those with the suggested one by 15% to 28% at 0 deg. These results support that the suggested energy-dependency expression is closer to the experimental data. A remarkable thing is that the DCSs with suggested b_{pol} expression and Muffin-tin approximation (red curves) are smaller than those of the free atom approximation (blue curves) by about 63% at 0 deg. This is in agreement with the prediction of Aouchiche et al. (2008) that the DCSs at small angles are larger for the vapour-phase than for the liquid-phase.

In summary of these results, the correlation-polarizability potential has some influence up to 60% at 0 deg, however, it is negligible at the intermediate and high scattering angles.

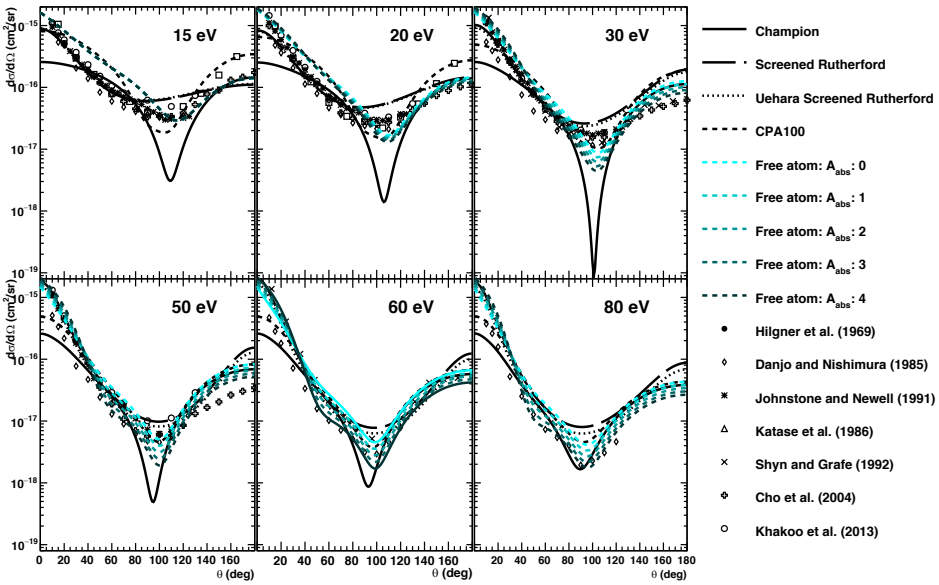


Figure 2.4: DCSs according to different A_{abs} values in the energy range of 15-80 eV from Shin et al. (2018). The light cyan dashed line indicates the lowest absorption strength, and the dark cyan dashed line indicates the highest absorption strength.

At last, in the case of **inelastic absorption potential**, we assumed that the absorption strength A_{abs} varies with energy because the possibility of inelastic scattering depends on electron energy as shown in Figure 2.4. There is no effect of absorption strength below 15 eV due to the threshold of the inelastic excitation. Above 15 eV, the high absorption strength increases the DCSs at low scattering angle, however, at intermediate and high scattering angle above 32.5 degree, the DCSs decrease with high absorption strength. The DCSs with an absorption strength of 0 are closer to the experiments at the lowest energies, however, the absorption strength should be getting larger at higher energies in order to match with the experiments.

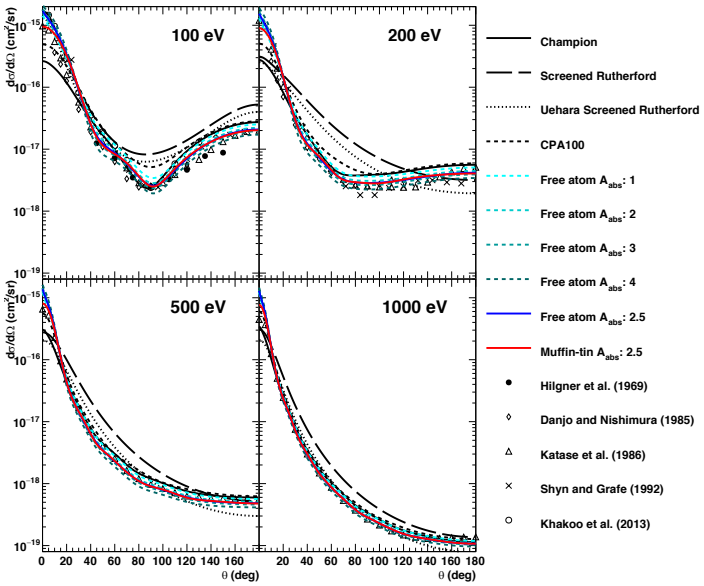


Figure 2.5: DCSs according to different A_{abs} values in the energy range of 100-1000 eV from Shin et al. (2018). The light cyan dashed line indicates the lowest absorption strength, and the dark cyan dashed line indicates the highest absorption strength. The blue and red lines represent the DCSs of the optimal inelastic absorption strength with free atom approximation and Muffin-tin approximation, respectively.

Figure 2.5 shows that a value of 2.5 gives DCSs in reasonable agreement with experimental data especially at intermediate and large angles above 32.5 deg. We thus decide to use an absorption strength of 2.5 above 100 eV, and a reduced absorption strength below this energy. To minimize the discontinuity of the DCSs, we propose the following formula of the absorption strength as function of the electron energy E:

$$A_{abs}(E) = \min(E \times 2.5/100, 2.5) \tag{2.29}$$

The absorption strength according to the electron energy $A_{abs}(E)$ is linearly increased from electron energy of 0 to 100 eV, and then the value is fixed to 2.5 above 100 eV.

In summary of this optimization study, we used both the Muffin-tin (default option) and free atom approximation (alternative option) using correlation-polarizability potential of LDA with a static polarizability of 1.457 Å^3, the Perdew-Zunger model for the correlation potential, the Buckingham potential for the correlation-polarizability potential, and a suggested energy dependence for b_{pol} given by equation (2.28). The inelastic absorption potential is considered with an energy gap Δ of 8.22 eV. The absorption strength A_{abs} of 2.5 is selected above 100 eV, otherwise it varies as a function of energy based on suggested expression in equation (2.29).

2.1.4. Plausibility of the electron elastic scattering models

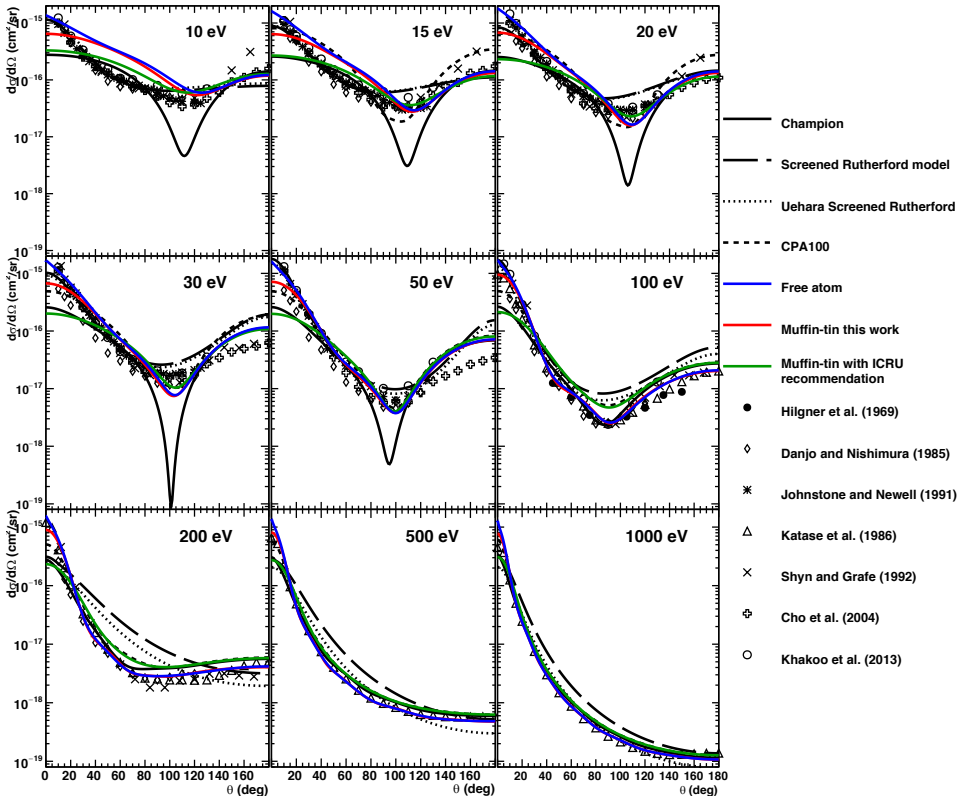


Figure 2.6: DCSs of Muffin-tin approximation (red lines), and free atom approximation (blue lines) with F-DF-FM-LDA-A configuration and optimized optical parameters. The DCSs with ICRU recommendation (green lines) and the experiment data measured in vapour-phase water are indicated as well.

The results of DCSs with selected configuration (F-DF-FM-LDA-A) and optimized optical parameters (correlation-polarizability model, adjustable energy dependence of b_{pol}, absorption potential) show good agreement with experiments as shown in Figure 2.6. However, we can observe that the disagreements below 20 eV which differ from experimental data at low and intermediate scattering angle are not only present in the ELSEPA model but also in the Champion model. We can thus assume that the Dirac partial-wave approximation has a limited accuracy on the electron cross-section below 20 eV, because the exchange potential cannot accurately predict the interaction with the electron below 1 Hartree (27.21 eV) (Bransden et al., 1976). Moreover, the DCSs at 0 deg are overestimated considering the relationship between vapour and liquid water data above 500 eV.

Globally, the DCSs calculated in this study are closer to the measurements than the model of SR, USR, and Champion in the entire energy range. In particular, the cross-sections are much improved at low energy and intermediate angle. The DCSs of the Muffin-tin approach are lower than those of the free-atom approach at low scattering angle by about 60% in agreement with the expected difference between vapour- and liquid-phases.

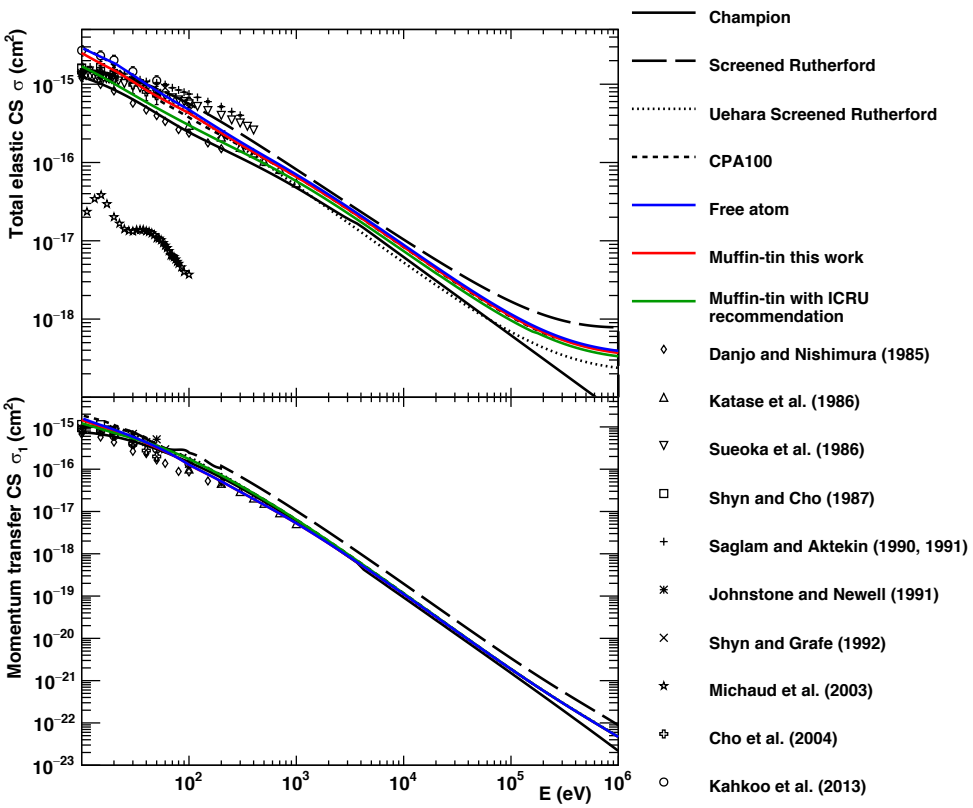


Figure 2.7 TCSs (upper) and MCSs (lower) for electron elastic scattering in liquid water plotted from 10 eV up to 1 MeV as a function of electron incident energy from Shin et al. (2018). The red and blue curves are the cross-section data developed in this study (F-DF-FM-LDA-A using independent-atom approximation with Muffin-tin and free atom approximation, respectively). The results with ICRU recommendation (green lines) and the experiment data are indicated as well.

Figure 2.7 shows the comparison of TCSs and MCSs calculated by ELSEPA, with current Geant4-DNA models and experimental data in the vapour-phase. The MCSs of the SR and the USR model below 200 eV are slightly distorted because of the independent calibration. The TCS and MCS calculated by this study are closer to experiments especially below 50 eV. On the entire energy range, the TCS and MCS of ELSEPA are perfectly following the tendency of the USR model and experiments.

2.2. The impact of the ELSEPA electron elastic scattering model on Geant4-DNA simulations

In this section, the electron elastic scattering model developed in the previous section is implemented into the default physics constructor of Geant4-DNA. Table 2.3 lists the different models used for the simulation of electron interactions in the physics constructors available in Geant4-DNA. The low energy limit of option 2, 4, and 6 when they include the new ELSEPA elastic model were determined from the low energy limit of the new ELSEPA elastic model: the energy ranges covered by the three options including the new elastic model become: 11 eV-1 MeV, 10 eV-10 keV, and 11 eV-256 keV, for option 2, 4 and 6 respectively.

Table 2.3: Elastic and inelastic models employed in Geant4-DNA physics constructors and energy limits of applicability.

	Elastic	**Inelastic**	
		Excitation	**Ionization**
G4EmDNAPhysics _option2*	Champion model (Champion, 2003) (7.4 eV – 1 MeV)	Emfietzoglou dielectric model (9 eV – 1 MeV)	Emfietzoglou dielectric model (11 eV – 1 MeV)
G4EmDNAPhysics _option4*	Uehara Screened Rutherford model (Uehara et al., 1993) (9 eV – 10 keV)	Emfietzoglou-Kyriakou dielectric model (8 eV – 10 keV)	Emfietzoglou-Kyriakou dielectric model (10 eV – 10 keV)
G4EmDNAPhysics _option6* (Bordage et al., 2016)	CPA100 (11 eV – 256 keV)	CPA100 (11 eV – 256 keV)	CPA100 (11 eV – 256 keV)

* Abbreviated later as “option 2”, “option 4”, and “option 6” (or “default”).
* The physics constructors including the ELSEPA elastic model are defined as “option 2 ELSEPA” …

In this section, the validity of the implementation is verified at first. And then, the influence of the new elastic scattering model is evaluated using the examples available in Geant4-DNA.

2.2.1. Determination of optimal angle and energy binning in ELSEPA

ELSEPA provides DCS values with 606 exponential bins in angle as a default. The angle bin at low angle is denser than that at high angle because the DCSs at 0 deg are getting sharper, especially at high energy. However, having 606 angle bins could be a burden for Monte Carlo simulations, loosing much computing time for interpolation. In this study, we qualitatively determined the threshold of the angle bin without losing information especially in the energy range of 10 eV to 1 MeV by comparing the evolution of DCSs according to the initial electron energy, and reduced the angle bins.

Determination of the number of energy bins is indeed one of the important issues to accurately calculate cross-sections because rough energy bins cause loss of information in energy regions which show rapid variation of DCSs. We calculated the DCSs using 1,000 bins on the energy range between 10 eV and 1 MeV. Each energy bin increases as an exponential function described as:

$$E_{i+1} = E_i \times \left(\frac{1\ MeV}{10\ eV}\right)^{\frac{1}{1000-1}}, E_0 = 10\ eV \tag{2.30}$$

In this study, the validity of the energy bin was evaluated using the normalized root-mean-square error (NRMSE) which indicates the difference between two distributions of DCSs with different numbers of energy bins (Herman, 1975):

$$NRMSE(E)\ (\%) = \frac{\sqrt{\frac{\sum_{i=1}^{n}\left(\left(\frac{d\sigma}{d\Omega}\right)_E(i) - \left(\frac{d\sigma}{d\Omega}\right)_{E+\Delta E}(i)\right)^2}{n}}}{\left(\overline{\frac{d\sigma}{d\Omega}}\right)_E} \times 100\ (\%) \tag{2.31}$$

where $\left(\frac{d\sigma}{d\Omega}\right)_E(i)$ indicates the DCS of specific angle *i* and energy *E*. *n* is the number of energy bins. ΔE is the energy gap between *E* and the next energy.

2.2.2. Geant4-DNA examples

We used various Geant4-DNA examples to verify the influence of the new elastic model. The results with the default constructors were compared with the results obtained with these constructors using the new elastic model and with results of previous works. The number of incident electrons was set to 10^6, and the processes of vibrational excitation (***G4DNAVibExcitation***) and molecular attachment (***G4DNAAttachment***) were adopted for "option 2", except in the case of "microyz" example for the fair comparison to previous published works.

The "mfp" example

The "mfp" example available in Geant4-DNA simulates mean-free-path (MFP) values. The MFP indicates the mean distance travelled by the particle between two successive interactions, and it is usually used for the comparison of track structure codes. In this example, user can select to simulate inelastic MFP and total MFP. We calculated both of the inelastic and total MFP; the inelastic MFP is useful for the comparison with reference data (this is only useful for the default Geant4-DNA constructors because the elastic scattering process doesn't influence the inelastic MFP), and the total MFPs for the comparison between default options and options using the new elastic model in the energy range of 10 eV-50 keV (up to 10 keV for option 4, and higher than 11 eV for option 6).

The "range" example

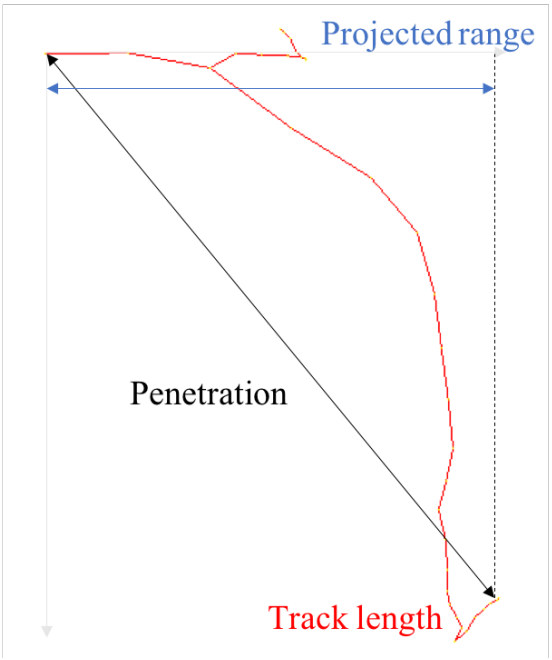


Figure 2.8: Schematic illustration about the quantities used for range simulation.

The "range" example calculates three quantities: track length, penetration, and projected range. The track length is the cumulated path length of the trajectory the particle follows in traveling from its initial position to its final position. The penetration represents the distance between the initial position and the final position of the particle, and the projected range is the projection along the incident direction (axial penetration) as illustrated in Figure 2.8. We calculated the ranges with different options in the energy range of 10 eV-50 keV. In order to verify the simulations, the results of range and MFP simulations were compared with references listed in Table 2.4.

Table 2.4: References used for the range and MFP simulations. Each column indicates: reference paper, calculation method or model, phase of target medium, simulated quantity, and energy range, respectively.

References	Methods and models	Target phase	Quantity	Energy range
Pages et al. (1972)	Bethe formula (Rohrlich and Carlson, 1954) for the energy loss by collision with the Sternheimer theoretical density-effect correction (Sternheimer, 1966) Cross-section generated by the Koch-Motz formula (Koch and Motz, 1959) for the energy loss by radiative collision	No mention about phase	Track length	10 keV – 100 MeV
Ashley (1988)	Analytical model based on optical data (Ashley, 1983) and dielectric response function (Penn, 1987)	Liquid water	Inelastic MFP	40 eV – 10 keV
Watt (1996)	Continuous slowing down approximation (CSDA) range (ICRU, 1984)[a] Modified Bethe theory given in ICRU 37 (ICRU, 1984) > 10 keV Deduction based on the theoretical results of Ashley (1988) and empirical evaluations (Ziegler, 1980, Iskef et al., 1983) < 10 keV	Liquid water	Track length	15 eV – 30 MeV
Dingfelder et al. (1999)	Analytical model based on optical data and dielectric response function (Penn, 1987)	Liquid water	Inelastic MFP	40 eV – 10 keV
Akkerman and Akkerman (1999)	CSDA range (ICRU, 1984) with and without exchange Cross-section based on binary encounter approximation for inelastic (Gryziński, 1965) and partial wave for elastic collision (Akkerman and Chernov, 1978, Akkerman et al., 1992)	No mention about phase	Track length[b] Inelastic MFP	50 eV – 10 keV

Meesungnoen et al. (2002)	MC track-structure code (TRACPRO) (Cobut et al., 1998) Elastic and inelastic cross-sections based on empirical data for amorphous water with a factor of 2 to account for differences between solid and liquid-phase (Michaud and Sanche, 1987) Vibrational excitations, dissociative attachment, and electronic excitation are taken into account (Michaud et al., 2003)	Liquid water	Penetration	0.2 eV – 150 keV
Pimblott and Siebbeles (2002)	MC track-structure code (MOCA) (Paretzke et al., 1986, Pimblott et al., 1996) and CSDA range (ICRU, 1984) Elastic cross-section based on partial wave method (Fernández-Varea et al., 1996, Salvat, 1998, Pimblott et al., 2000) Inelastic cross-sections based on experimental dipole oscillator strength (Nikogosyan et al., 1983, Bartels and Crowell, 2000)	Liquid water	Track length[c] Penetration Projected range Inelastic MFP	70 eV – 100 keV
Wilson et al. (2004)	MC track-structure code (PITS) (Wilson and Nikjoo, 1999) Inelastic cross-section based on the Dingfelder-GSF model for liquid water (Dingfelder et al., 1998) and elastic cross-section based on the experiments and NIST data for vapour water (Jablonski et al., 2010) using the screened Rutherford theory	Liquid water	Track length Penetration	20 eV – 100 keV
Uehara and Nikjoo (2006)	MC track-structure code (KURBUC) (Uehara et al., 1993) Inelastic cross-section Water vapour cross-sections for ionization and excitation (Uehara et al., 1999) are compiled from different sources and elastic cross-section based on the Rutherford formula for vapour water with a screening parameter (Uehara et al., 1993) Vibrational excitation and multi-step thermalization process are taken into account (Terrissol and Beaudre, 1990)	Vapour water	Penetration	0.1 eV – 100 keV
Plante and Cucinotta (2009)	MC track-structure code (RETRACKS) (Plante and Cucinotta, 2009) Ionization cross-section of Rudd's model (Rudd, 1990) below 50 keV and Seltzer's equation (Seltzer and Berger, 1986) above 50 keV Excitation model of Kaplan and Sukhonosov (1991), Kutcher and Green (1976) below 100 eV and above, respectively Elastic cross-section based on the experimental data (Michaud et al., 2003) and DCSs proposed by Brenner and Zaider below 200 eV (Brenner and Zaider, 1983), and Rutherford cross-section above 200 eV (Uehara et al., 1993) Vibrational excitation (Michaud et al., 2003), dissociative attachment (Rowntree et al., 1991), and bremsstrahlung process (Seltzer and Berger, 1986) are taken into account	Liquid water	Penetration	0.1 eV - 10 MeV

Wiklund et al. (2011)	MC track-structure code (GNU Scientific Library) (Galassi et al., 2002) Inelastic cross-section based on Born approximation for liquid water (Dingfelder et al., 1998) and elastic cross-section generated by ELSEPA (P-DF-FM-Buckingham) (Salvat et al., 2005) with density scaling to account for the liquid-phase Cross-sections for dissociative attachment are included	Liquid water	Track length Penetration Projected range	10 eV – 10 keV
ICRU (2014)	CSDA range (ICRU, 1984) Cross-sections based on ESTAR (Berger et al., 1998) for elastic and inelastic scattering	Liquid water	Track length	1 keV – 1 GeV
Emfietzoglou et al. (2017)	In-house developed MC track-structure code (Emfietzoglou et al., 2003) Emfietzoglou-Cucinotta-Nikjoo (ECN) model for inelastic scattering (Emfietzoglou et al., 2005) NIST elastic cross-section (Berger et al., 1998)	Liquid water	Track length Penetration Inelastic MFP	10 eV – 10 keV

[a] Integrating the inverse of the total stopping power.
[b] CSDA range with exchange process was used in this study.
[c] Track length from MC method.

The "TestEm12" example

This example calculates dose point kernels (DPKs) in order to evaluate the energy deposition distributions from point isotropic sources (Bordes et al., 2017). This quantity is the most useful in the fields of radiation protection and medical applications. The DPKs represent the absorbed dose as a function of the distance from the center of the target (absorbed radial dose). We calculated DPKs for several incident energies (50, 100, 200, 500 eV, 1, 2, 5, and 10 keV) using different physics constructors (options). The radius of the target depends on the incident electron energy from 12 nm to 3.5 um, and the number of shells is set to 50. In particular, the comparison method suggested by Maigne et al. (2011) was used for comparison of the unscaled DPKs between the default options and options with the new elastic model. Relative differences between default options and options with the ELSEPA model at a distance *r* are determined by:

$$\Delta(r) = \frac{\delta E_{def}(r) - \delta E_{EL}(r)}{\max(\delta E_{def}, \delta E_{EL})} \times 100\ (\%) \tag{2.32}$$

where $\delta E_{def}(r)$ and $\delta E_{EL}(r)$ represent energy deposition in a shell of radius *r*, calculated with the default option and with the option using new elastic model, respectively. This formula reduces statistical fluctuations due to the low energy deposition in the tail of distributions.

The "microyz" example

Microdosimetry is a formalism used to predict the biological effects caused by ionizing radiation in biological cells. It quantitatively evaluates the stochastic aspects of energy deposition in irradiated media (Rossi and Zaider, 1996). The "microyz" example calculates the characteristics of particle energy transfers through their lineal energies in nanometric-size sites using the microdosimetry formalism proposed by Kellerer and Chmelevsky (1975) as well-described in ICRU report 36 (ICRU, 1983). The lineal energy *y* defined as:

$$y = \frac{\varepsilon}{\bar{l}} \tag{2.33}$$

where ε indicates the energy delivered in a small volume for each energy deposition event and $\bar{l}$ is the mean chord length of the target volume (e.g. for a spherical target, $\bar{l}$ is equal to 2/3 of the diameter).

The frequency-mean lineal energy represents the expected value of the lineal energies per single event taking also into account all the secondary electrons that are produced by the primary collisions and all their interactions in the medium:

$$\bar{y}_F = \int y f(y) dy \tag{2.34}$$

where $f(y)$ represents the PDF of lineal energy $f(y)$.

We performed the simulations with different options as with the other examples using a spherical target diameter of 2 and 10 nm, in the energy range of 50 eV-10 keV. Moreover, we applied the simulation parameters of previous study (Kyriakou et al., 2017) such as a tracking cut value of 11 eV for every physics option, without considering vibration excitation and molecular

attachment processes, and considering atomic de-excitation (emission of fluorescence photons and Auger electrons).

The "clustering" example

The "clustering" algorithm adapted from the Density-Based Spatial Clustering of Applications with Noise (DBSCAN) algorithm (Ester et al., 1996) calculates the identification of clusters from energy depositions. This Geant4-DNA example is useful to simulate the energy deposition patterns in the irradiated medium and can be used to estimate roughly the tendency of direct DNA single strand breaks (SSB) and double strand breaks (DSB) (Francis et al., 2011, Francis et al., 2012), without the need for a detailed geometry of biological targets. We used the calculation parameters of previous studies (Francis et al., 2011). For example, the maximum diameter of a cluster is set to 3.3 nm and the probability to create a strand break increases linearly with the energy deposit between 5-37.5 eV. Incident electrons of energy 500 eV, 1 keV, and 10 keV were used with appropriate target size ($1 \times 1 \times 0.5$ um^3) similar to a small cell size. We calculated the yields of SSB, DSB, and the ratio between them with the default options available in Geant4-DNA and with the options including the new elastic model.

2.2.3. Results for the implementation of the elastic scattering cross-section into Geant4-DNA

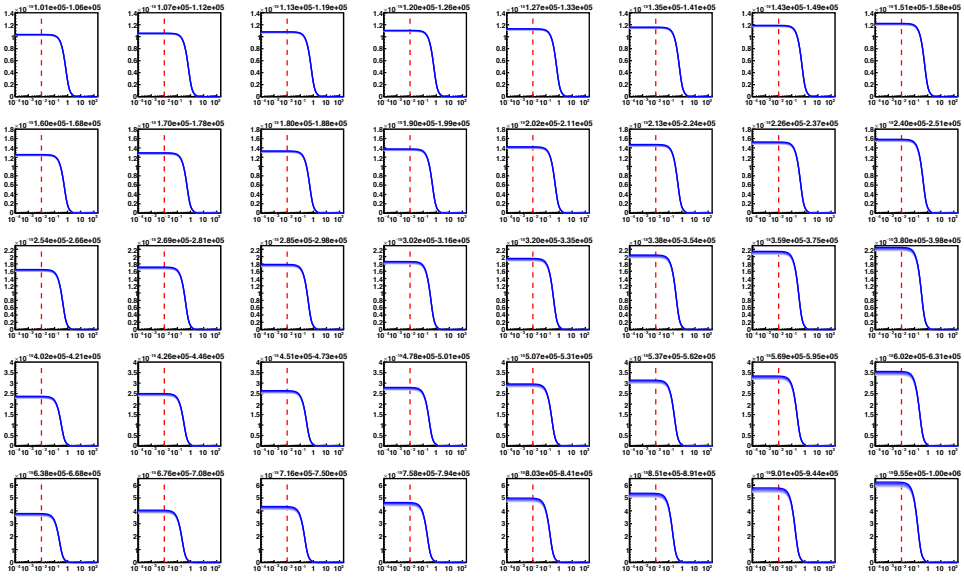


Figure 2.9: DCSs in the energy range (10^5 to 10^6 eV), presented in 40 partial energy ranges. The light blue and dark blue lines correspond respectively to the highest and the lowest energy for each partial energy range. The x-axis indicates scattering angle in logarithmic scale (deg), and the y-axis shows the DCS in linear scale (cm^2). The red dashed line indicates the scattering angle of 0.01 deg.

Figure 2.9 shows the DCSs at high energies, above 100 keV, which show the most rapid variations at low scattering angles, especially from 0.01 deg and above. The plateau region is getting

shorter with increase of the electron energy, nevertheless, the DCSs below 0.01 deg are still flat. We neglect the DCSs below 0.01 deg, this allows us to reduce the number of angle bins from 601 to 534 (11.15%).

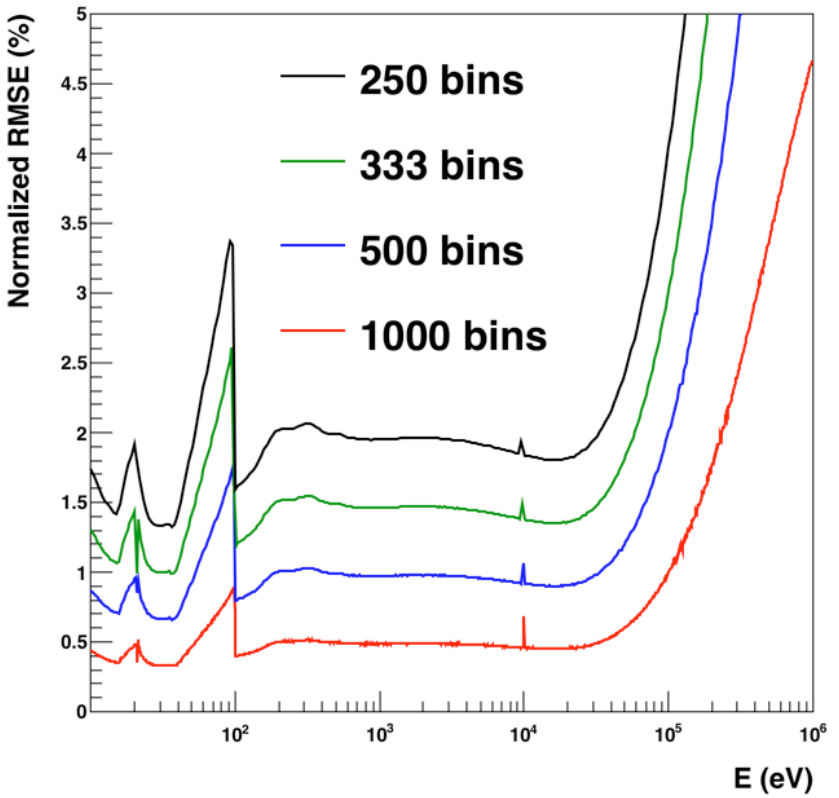


Figure 2.10: NRMSE according to energy, with energy bins of 250 (black), 333 (green), 500 (blue), and 1000 (red).

Even though the width of energy bin is exponentially increased, the NRMSEs decrease at low energy and then increase at high energy, as shown in Figure 2.10. The NRMSE with 1000 energy bins is less than 5% in the entire region. However, we observed three undesired peaks in the NRMS plot. The peaks at 20 and 100 eV are caused by the modified energy dependency b_{pol} in equation (2.28) increasing from 20 eV and by the energy-dependent absorption potential in equation (2.29), respectively. The other peak around 10 keV is due to the fact that ELSEPA automatically turns off the correlation-polarizability potential above 10 keV. The NRMSs of these peaks with 1000 energy bins are 0.49% for 20 eV, 0.89% for 100 eV, and 0.69% for 10 keV, respectively. With these results, we can decide that the number of 1,000 energy bins is a reasonable choice.

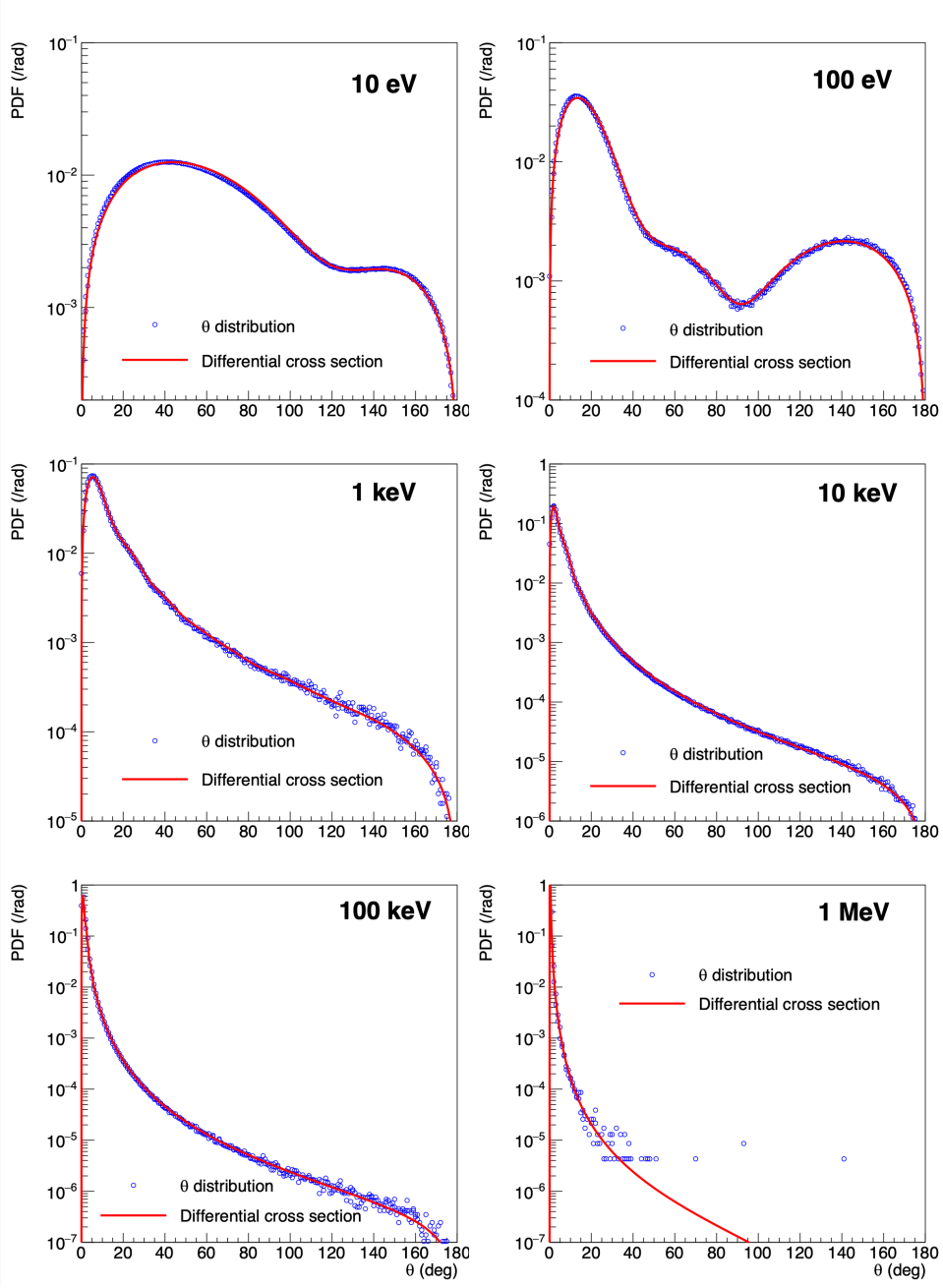


Figure 2.11: Comparison of DCSs and simulated scattering angles in the energy range of 10 eV to 1 MeV. The red lines and the blue circles represent the PDFs of the DCSs and the simulated theta distributions, respectively.

In order to verify the implementation of the new elastic model, we compared the differential cross-sections (DCSs) with the θ distributions obtained by Geant4-DNA simulations at several kinetic energies: 10, 100 eV, 1, 10, 100 keV, and 1 MeV. For this, we extracted two unit vectors which describe the direction of a particle's trajectory before and after the single elastic scattering. The scattering angle θ can be described as:

$$\theta = \arccos(\vec{A}\cdot\vec{B}) \tag{2.35}$$

We recorded the number of interactions as a function of scattering angle θ and normalized it to calculate the probability density function (PDF). The θ distributions are perfectly matching the DCS data as shown in Figure 2.11. In the case of 1 MeV, statistical errors look significant at large scattering angle, however, the distribution still has similar tendency as the DCS. These results demonstrate that the simulations with the new elastic model work exactly as expected.

Table 2.5: Calculation time of range simulations according to each option at several incident energies for 10^6 electrons.

(Unit: CPU-hr)	**Option 2**			**Option 4**			**Option 6**		
Energy	11 eV	1 keV	10 keV	11 eV	1 keV	10 keV	11 eV	1 keV	10 keV
Default (A)	0.48	0.75	3.69	0.18	2.30	41.6	0.11	0.86	4.43
ELSEPA (B)	1.04	1.17	4.76	0.28	2.64	42.7	0.91	0.96	4.72
Ratio (B/A)	2.17	1.56	1.29	1.56	1.15	1.03	8.27	1.12	1.07

* The table denotes only the CPU time spent in user-mode code (user time). The CPU time spent in system calls within the kernel (system time) contributes less than 20 seconds for all options and energies.

The simulation time of each simulation with various options is shown in Table 2.5. The differences between options including the ELSEPA elastic scattering model derive mainly from the different inelastic scattering models. For example, the calculation time using option 4 with ELSEPA at 10 keV was significantly increased about 10 times more than that at 9 keV. This phenomenon derives from the computational cost increase for inelastic cross-sections (excitation and ionization) at around 10 keV. With the option 5, which is the accelerated version of option 4 (using cumulated DCSs for ionization), we can reduce the calculation time at 10 keV to 32.89 CPU-hr; however, option 4 is still employed to compare with ELSEPA results due to the accuracy of physics models (non-cumulated cross-sections). The calculation time of simulations with ELSEPA cross-section is slower than the time of default option by about 1-2 times, and this tendency is getting more significant with small electron energy due to the increased probability of the elastic scattering.

2.2.4. Results for the Geant4-DNA simulations

MFP simulation

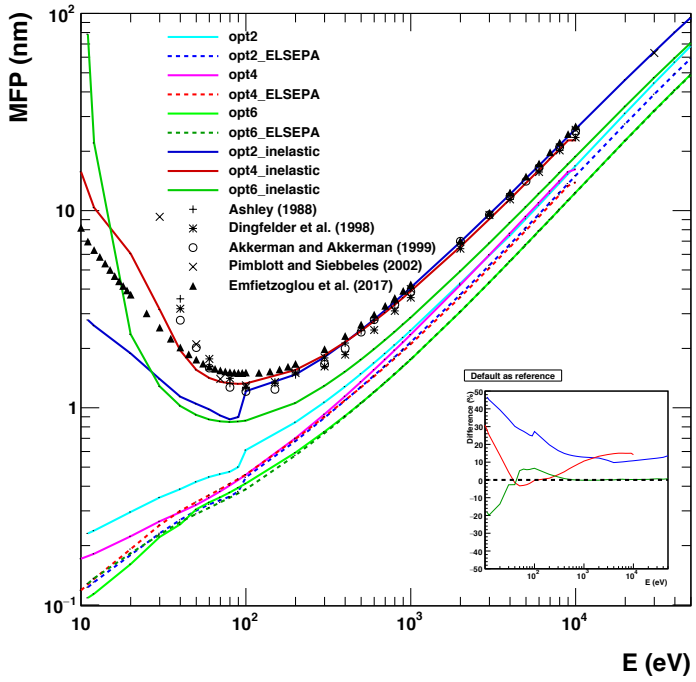


Figure 2.12: Results for the MFP simulations with option 2 (cyan solid), option 4 (magenta solid), option 6 (light green solid), and results of option 2 (blue dashed), option 4 (red dashed), option 6 (green dashed) with the ELSEPA electron elastic model. The markers are reference data given in Table 2.4. The inset shows the relative differences between default options and the options with the ELSEPA elastic model.

Figure 2.12 shows the comparison between the default options and the options with ELSEPA elastic scattering on MFP simulations. The inelastic MFP of option 4 is the closest to the reference data. In addition, the fall-off of option 2 at 100 eV is observed due to the energy limit of the vibrational excitation model (G4DNASancheExcitationModel) in both inelastic and total MFP. The total MFP λ_t is proportional to the inverse of the TCS and analytically can be described as:

$$\lambda_t = (n\sigma_t)^{-1} \tag{2.36}$$

where n is the number of target atoms per unit volume and σ_t is the TCS. In other words, the MFP may change according to the difference between TCS of default options and that including the ELSEPA elastic model.

As shown in the inset of the Figure 2.12, the MFP of option 2 with ELSEPA is smaller than the MFP of default option 2 because the TCS of the Champion elastic model used in option 2 is smaller than that of the ELSEPA elastic model in the entire energy region (Figure 2.7 upper). However, the tendency of option 6 is reversed at 45 eV based on the tendency of the total elastic

cross-section. We can assume that the option with shorter total MFP corresponds to more elastic scatterings compared to the option with longer MFP.

Range simulations

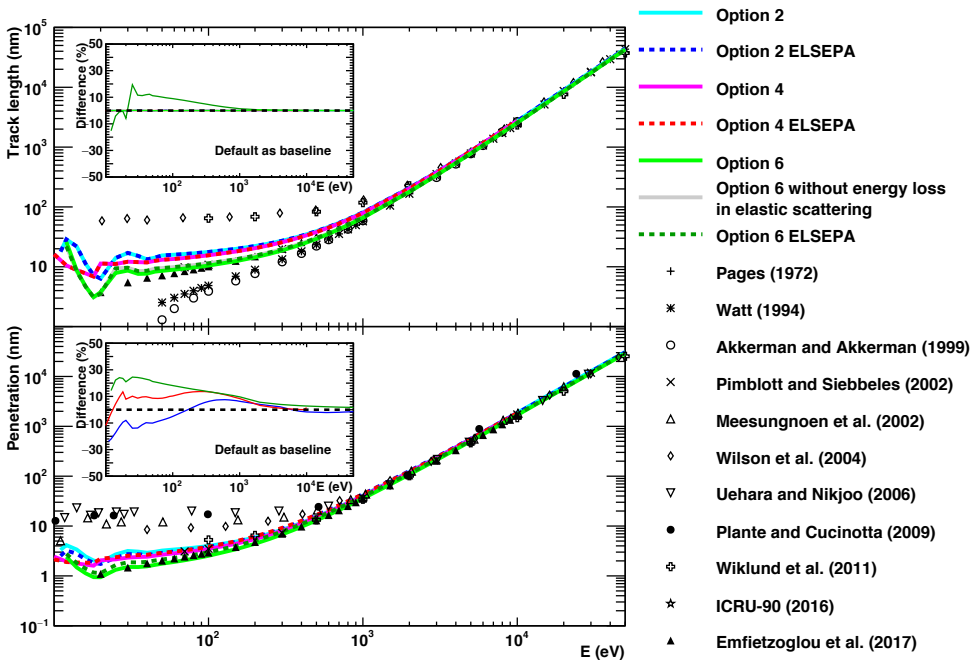


Figure 2.13: Results for the track length (upper) and penetration (lower) simulations of option 2 (cyan solid), option 4 (magenta solid), option 6 (light green solid), option 6 without energy loss in elastic scatterings (grey solid) and results of option 2 (blue dashed), option 4 (red dashed), option 6 (green dashed) with the ELSEPA electron elastic model. The markers are the reference data listed in Table 2.4. The inset shows the relative differences between default options and the options with the ELSEPA elastic model.

Figure 2.13 shows the results for track length and penetration simulations and references. The track length is only affected by energy losses mainly with inelastic scattering except for option 6. The elastic model of CPA100 employed in option 6 considers indeed small energy loss in elastic scattering with the following formula (Edel, 2006):

$$T_i - T_{i+1} = T_i(1 - cos\theta) \times 1.214e^{-4} \tag{2.37}$$

with the kinetic energy at the i^{th} step T_i and the scattering angle θ.

The track lengths of Wilson et al. (2004), Wiklund et al. (2011) which use the same inelastic model (Dingfelder-GSF model for liquid water based on Born approximation) show similar tendency. Moreover, the results of track length between default options and options with ELSEPA are the same except for option 6 because of the energy loss during elastic scattering as described above. Without this energy loss, the result of option 6 with ELSEPA cross-section also shows the same values as option 6. However, this is only for the comparison to validate the simulation, in the other case, option 6 with energy loss is used because it is more realistic. Interestingly, the track lengths with CSDA approach look shorter than the other approaches such as the MC method

because the low energy limit considered in the CSDA approach is larger than the tracking cut used by the MC method (Watt (1996): 50 eV, Akkerman and Akkerman (1999): 20 eV, and ICRU (2014): 10 keV).

With the same logic as for MFP, the relationships between the default options and the options with the ELSEPA model are related to the TCS of elastic scattering. The physics model with less elastic scattering travels longer, but the penetration range of option 6 is always shorter than that with the ELSEPA model due to the energy loss caused by the elastic scatterings. However, the relative differences between the default options and the options with ELSEPA are less than 20%, and the differences are getting smaller with the increase of initial energy because the contribution of the elastic scattering is small at high energy.

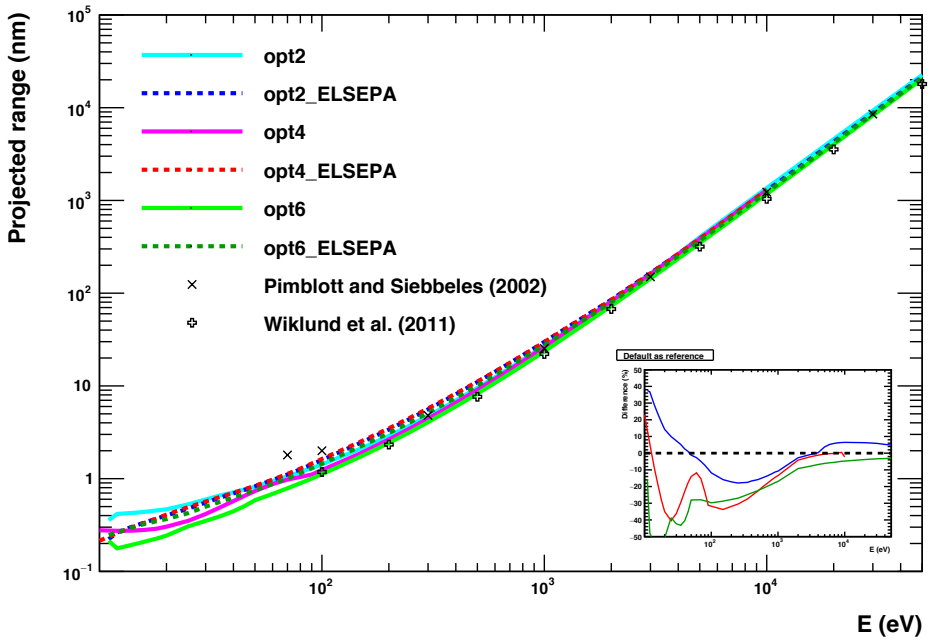


Figure 2.14: Results for projected range simulations of option 2 (cyan solid), option 4 (magenta solid), option 6 (light green solid), and results of option 2 (blue dashed), option 4 (red dashed), option 6 (green dashed) with the ELSEPA electron elastic cross-section. The markers are the reference data listed in Table 2.4. The inset shows the relative differences between default options and the options with the ELSEPA elastic model.

The results of projected range show similar tendency as the results of penetration range: less elastic scattering events cause a more forward directed trajectory. The projected range of option 4 and 6 with ELSEPA is longer than that of default option 4 and 6, and option 2 with the ELSEPA results are longer than that of default option 2 in the energy range of 250 eV-3 keV.

DPK simulation

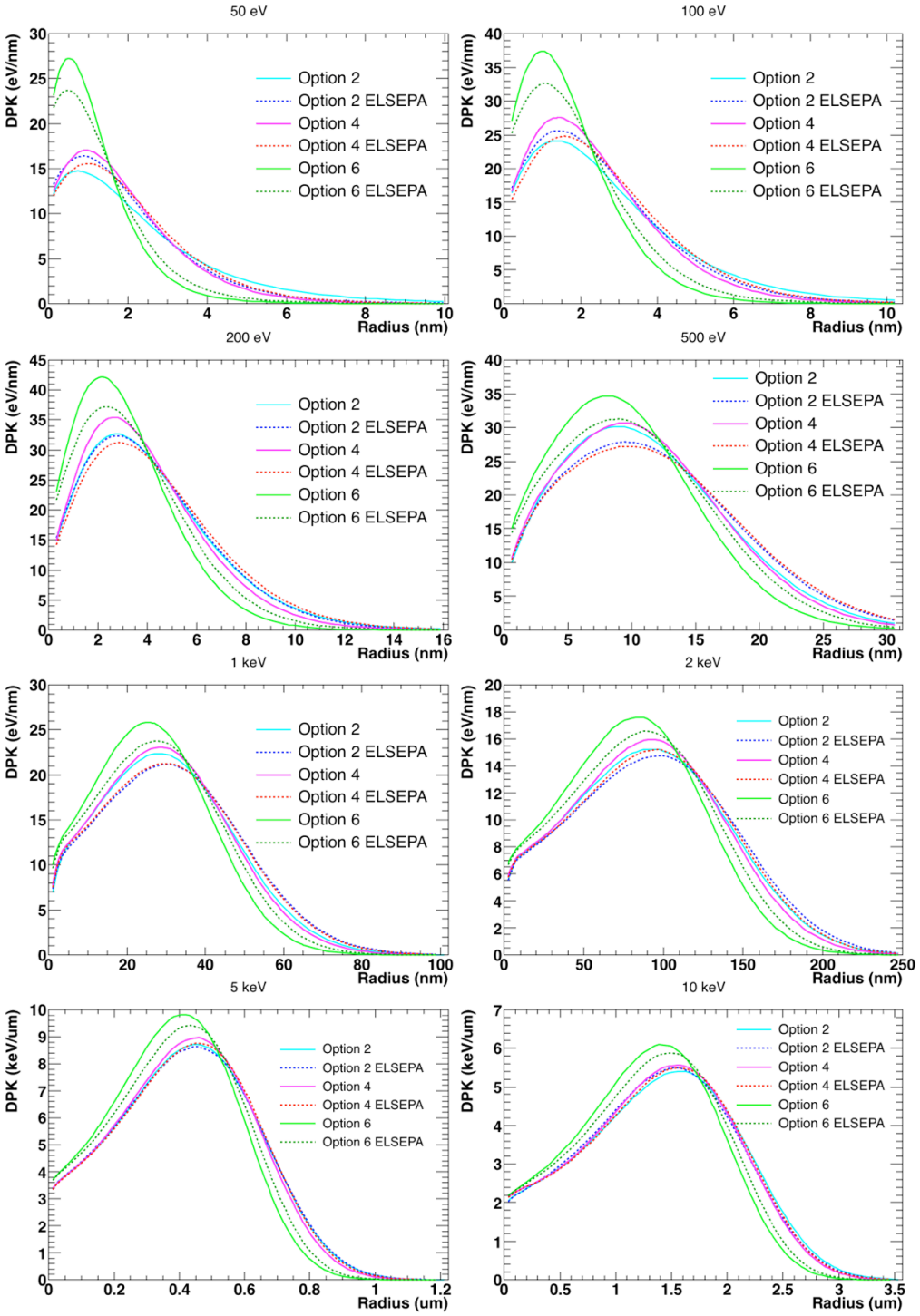


Figure 2.15: DPKs in the energy range of 50 eV-10 keV for default option 2 (cyan solid), option 4 (magenta solid), option 6 (light green solid), and option 2 (blue dashed), option 4 (red dashed), option 6 (green dashed) with the ELSEPA elastic cross-section from Shin et al. (2018).

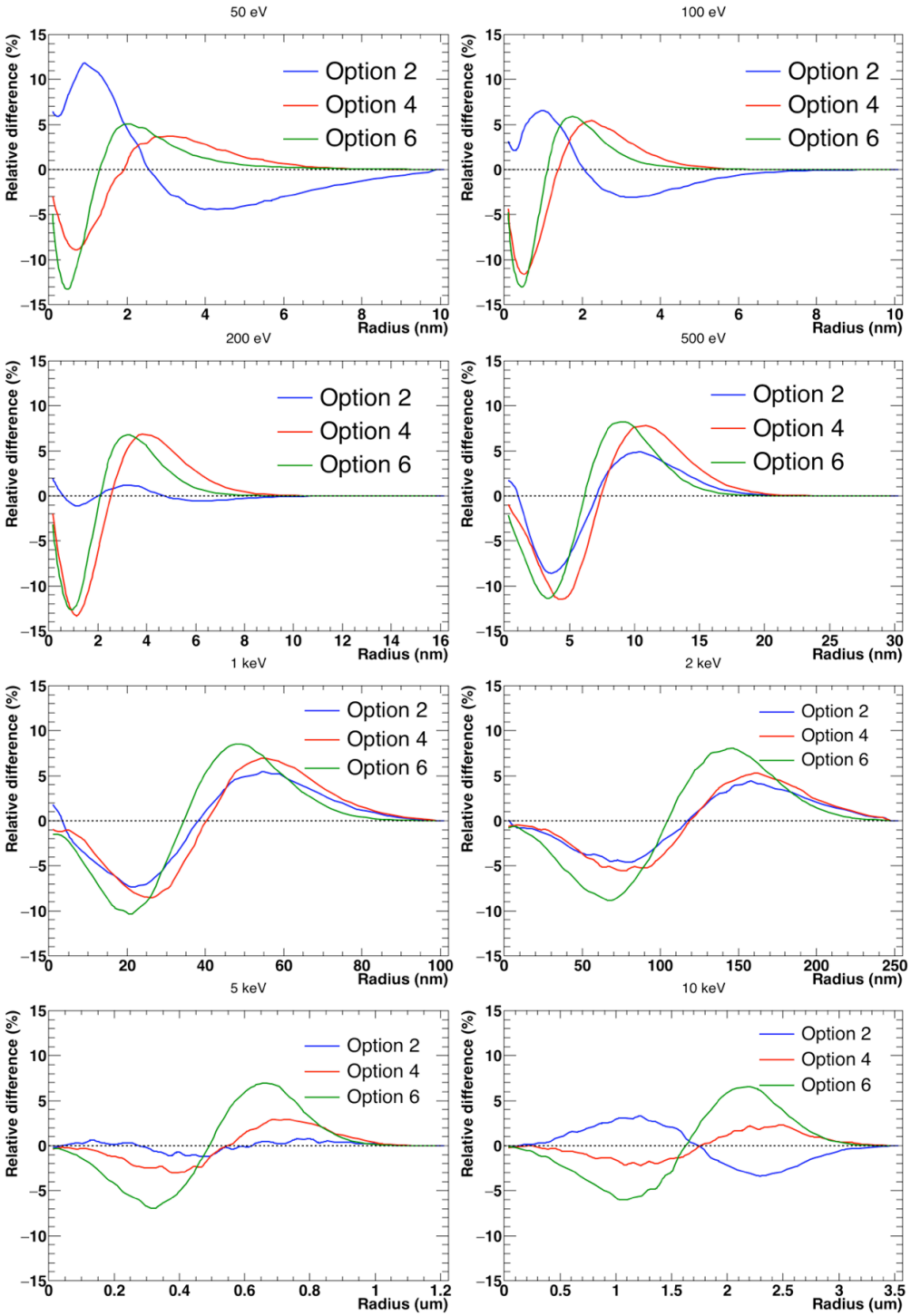


Figure 2.16: Relative differences between default options and options with the ELSEPA elastic model in the energy range of 50 eV-10 keV from Shin et al. (2018).

With the results of Figure 2.15 and 2.16, we can assume that the diffusivity (or spatial "spread") of DPKs is affected by three factors:

1) More elastic scatterings cause less diffusive distribution.

- The TCS of the Champion elastic model in option 2 is always lower than that of the ELSEPA

model over the entire energy range. However, the diffusivity of option 2 DPKs depends on the incident energy;

- The TCS of the Uehara Screened Rutherford elastic model in option 4 is similar or lower than that the ELSEPA model above 50 eV. The electrons transported with option 4 and the ELSEPA model are more diffusive than those transported with the default option 4 over the entire energy range.
- We can thus conclude that the diffusivity of DPK due to the elastic model depends not only on the amount of elastic scattering but also on scattering angle values.

2) Smaller scattering angle makes more diffusive distribution.

- In order to consider both the number of elastic scatterings and the scattering angle values simultaneously, we assume that momentum cross-section (MCS) is in close relationship with diffusivity. The MCS can be described by the integration of $(1 - cos\theta)\frac{d\sigma}{d\Omega}$; in other words, a model with lower MCS causes less amount of elastic scattering and small scattering angle.
- The MCS of ELSEPA model is always smaller than that of the Uehara Screened Rutherford model above 20 eV, and this tendency is compatible with the DPK results that the ELSEPA model makes more diffusive distribution.
- The MCSs of ELSEPA and Champion model cross two times at 80 eV and 4 keV in option 2. However, we can observe such crossing of diffusivity (which corresponds to the minima of the difference between DPKs) at 200 eV and 5 keV.
- The energy shift might be caused by energy losses due to inelastic interactions. We can guess that the contribution of electrons in the 10-80 eV and 10 eV-4 keV energy ranges is more important than the contribution of the electrons in 80-200 eV and 4-5 keV energy ranges to determine the diffusivity.

3) Energy loss in elastic scattering makes less diffusive distribution due to the reduction of the total track length of electrons.

- The MCSs of ELSEPA model and CPA-100 model are similar above 1 keV.
- However, the DPKs of option 6 with ELSEPA elastic model are always more diffusive than those of default option 6 due to the small energy loss of elastic scattering in default option 6.

We can observe that the diffusivity of the elastic scattering model is in close relationship with the MCS, and the small energy loss in elastic scattering leads to less diffusive distribution. The differences are less pronounced at high energy due to the smaller contribution of elastic scattering.

microyz simulation

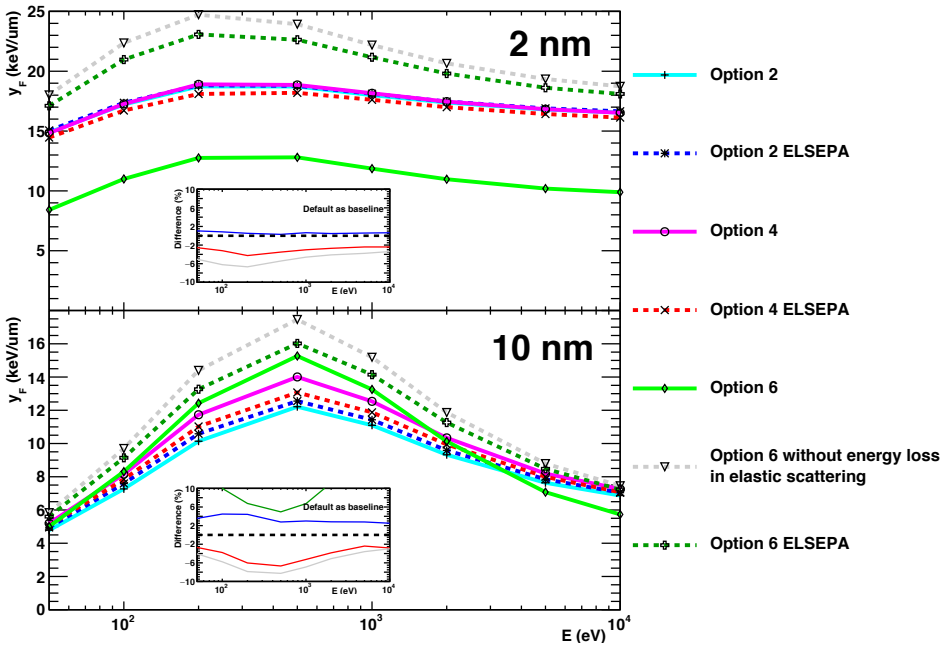


Figure 2.17: Frequency-mean lineal energy y_F as a function of incident energy with option 2 (cyan solid), option 4 (magenta solid), option 6 (light green solid), option 6 without energy loss in elastic scatterings (gray dashed) and results of option 2 (blue dashed), option 4 (red dashed), option 6 (green dashed) with the ELSEPA electron elastic model for a target of diameter 2 nm (upper) and 10 nm (lower). The inset shows the relative differences between default options and the options with the ELSEPA elastic model.

For the sphere of 2 nm in diameter, frequency-mean lineal energies y_F calculated with option 2 and 4 are very similar as shown in Figure 2.17. However, the result of option 6 is almost half of the other options due to the energy loss in elastic scattering. Without the energy loss, the y_F of option 6 is higher than the other options by about 20-40%, and we can assume that the large inelastic cross-section of option 6 is mainly responsible for this difference. The result of option 2 changes less than 1% with the ELSEPA elastic model. In the case of option 4, the ELSEPA elastic model reduces y_F by about 3%. The result of option 6 with the ELSEPA elastic model is significantly different with default option 6, however, it decreases y_F by about 5% in comparison with option 6 without the energy loss. The y_F values for the sphere of 10 nm diameter have similar tendency with that of 2 nm diameter. Option 2 doesn't affect much and options 4 and 6 are decreased with the ELSEPA elastic model. However, the relative differences are less than 5%.

Clustering simulation

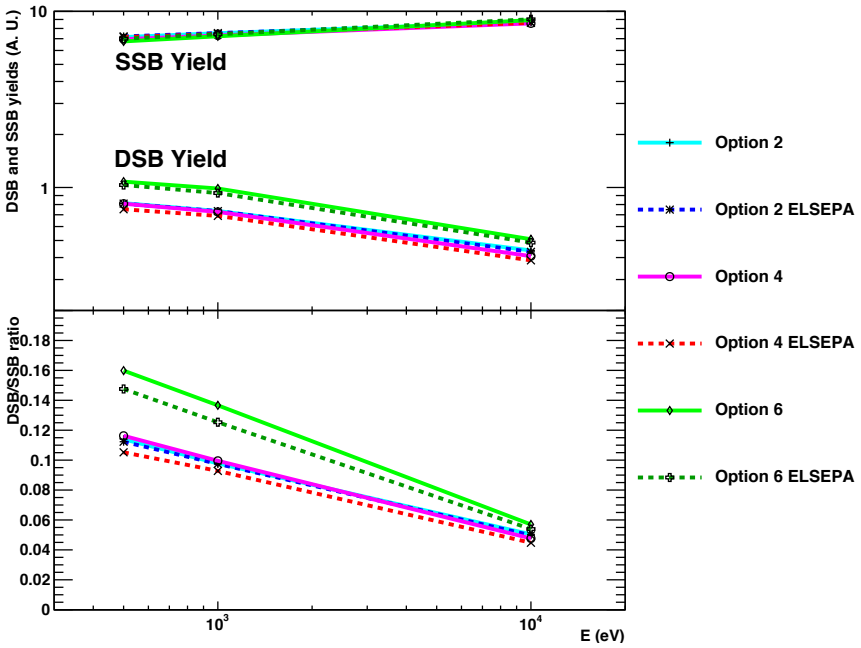


Figure 2.18: Yields of DSB and SSB (upper), and DSB/SSB ratio (lower) calculated by the DBSCAN algorithm as a function of incident energy with the option 2 (cyan solid), option 4 (magenta solid), option 6 (light green solid), and the results of option 2 (blue dashed), option 4 (red dashed), option 6 (green dashed) with the ELSEPA electron elastic cross-section.

Figure 2.18 shows the SSB and DSB yields normalized to energy deposition in arbitrary unit. With the increase of the incident energy, the SSB yield is increased and the DSB yield is decreased because the LET of electrons is getting lower up to an electron energy of 1 MeV. In contrast with the SSB yield results showing less difference (about 1%), the DSB yields of all default options are higher than the options with the ELSEPA model, up to 1.34, 7.05, and 5.78% for option 2, 4, and 6 respectively. Indeed, all default options are less diffusive than the options with the ELSEPA model in the energy range 500 eV-10 keV; in other words, electrons deliver energies in more concentrated area. Consequently, the ratio between DSB and SSB with the ELSEPA elastic scattering model is lower than the ratio with default options.

2.3. Conclusions

We studied two alternative electron elastic cross-section data sets using the ELSEPA software. As a first model for Geant4-DNA, we propose to choose the F-DF-MF-LDA combination using the independent atom approximation with Muffin-tin approximation by default. The free atom approximation will be our second alternative model.

The optical parameters including correlation-polarizability potential and inelastic absorption potential are qualitatively optimized. We used the correlation-polarizability potential of LDA with a 1.457 $Å^3$ static polarizability, the Perdew-Zunger model for the correlation potential, the Buckingham potential for the correlation-polarizability model, and a suggested energy dependence

for b_{pol}. The inelastic absorption potential is considered with an energy gap Δ of 8.22 eV and a value of 2.5 is taken for the absorption strength A_{abs} above 100 eV and varies linearly with energy below 100 eV. The results show significant improvement compared to the other existing Geant4-DNA models such as the SR, the USR, and the Champion. However, there are still limitations such as overestimations at low scattering angle above 500 eV and distortions below 20 eV.

We also investigated the impact of the ELSEPA elastic scattering model, using various Geant4-DNA physics constructors (option 2, 4, and 6). In order to quantitatively evaluate the new model, the examples mfp, range, TestEm12, microyz, and clustering available in Geant4-DNA were used, and the results of default constructors were compared with the results of the same constructors including the new elastic model. In detail, the new elastic model influences mean-free-path and range values, mainly at energies below 100 eV due to the dominant contribution of elastic scattering. We observed that differences are in close relationship with differences in total cross-section values for elastic scattering. Regarding DPKs, the default option 4 is more diffusive than the option 4 with the new elastic model because of the lower number of elastic scatterings with the new model and the smaller scattering angles. The option 6 with new elastic model is always more diffusive than the default option 6 due to the small energy loss in elastic scattering. In the case of option 2, the diffusivity of dose distribution is related to MCS of the elastic scattering model with energy shift caused by the energy losses due to inelastic interactions. The elastic scattering model with smaller MCS, representing less elastic scattering and smaller scattering angles, leads to more diffusive dose distribution. For the microyz example, the frequency mean lineal energy is more influenced by the energy loss in elastic scattering than by the elastic scattering model. The results of option 2 are not influenced, and those of option 4 and 6 with ELSEPA show smaller values in the entire energy. Finally, the effect of the elastic scattering model is negligible for the calculation of SSBs in the clustering example. However, the physics options causing less diffusive distribution make more DSBs resulting from more concentrated deposited energy.

References

Akkerman, A. & Akkerman, E. 1999, Characteristics of electron inelastic interactions in organic compounds and water over the energy range 20–10 000 eV. *J. Appl. Phys.,* 86**,** 5809-5816.

Akkerman, A., Gibrekhterman, A., Breskin, A., et al. 1992, Monte Carlo simulations of secondary electron emission from CsI, induced by 1–10 keV x rays and electrons. *J. Appl. Phys.,* 72**,** 5429-5436.

Akkerman, A. F. & Chernov, G. Y. 1978, Elastic scattering of electrons by atoms in the keV range. *Sov. Phys. Tech. Phys.,* 23**,** 247.

Alloni, D., Campa, A., Friedland, W., et al. 2012, Track structure, radiation quality and initial radiobiological events: considerations based on the PARTRAC code experience. *Int. J. Radiat. Biol.,* 88**,** 77-86.

Aouchiche, H., Champion, C. & Oubaziz, D. 2008, Electron and positron elastic scattering in gaseous and liquid water: A comparative study. *Radiat. Phys. Chem.,* 77**,** 107-114.

Ashley, J. Simple model for calculating stopping powers of condensed organic materials for low-energy electrons. 7th International Congress on Radiation Research, 1983 Amsterdam, The Netherlands. Martinus Nijhoff Publishers, A1-O3.

Ashley, J. 1988, Interaction of low-energy electrons with condensed matter: stopping powers and inelastic mean free paths from optical data. *J. Electron Spectrosc.,* 46**,** 199-214.

Bartels, D. M. & Crowell, R. A. 2000, Photoionization yield vs energy in H2O and D2O. *J. Phys. Chem. A,* 104**,** 3349-3355.

Berger, M. J., Coursey, J., Zucker, M., et al. 1998. *Stopping-power and range tables for electrons, protons, and helium ions*, NIST Physics Laboratory Gaithersburg, MD.

Bonham, R. A. & Strand, T. G. 1963, Analytical expressions for potentials of neutral Thomas—Fermi—Dirac atoms and for the corresponding atomic scattering factors for X rays and electrons. *J. Chem. Phys.,* 39**,** 2200-2204.

Bordage, M. C., Bordes, J., Edel, S., et al. 2016, Implementation of new physics models for low energy electrons in liquid water in Geant4-DNA. *Phys. Med.,* 32**,** 1833-1840.

Bordes, J., Incerti, S., Lampe, N., et al. 2017, Low-energy electron dose-point kernel simulations using new physics models implemented in Geant4-DNA. *Nucl. Instrum. Meth. B,* 398**,** 13-20.

Bote, D., Salvat, F., Jablonski, A., et al. 2009, The effect of inelastic absorption on the elastic scattering of electrons and positrons in amorphous solids. *J. Electron Spectrosc.,* 175**,** 41-54.

Bransden, B., McDowell, M., Noble, C., et al. 1976, Equivalent exchange potentials in electron scattering. *J. Phys. B-At. Mol. Opt.,* 9**,** 1301.

Brenner, D. J. & Zaider, M. 1983, A computationally convenient parameterisation of experimental angular distributions of low energy electrons elastically scattered off water vapour. *Phys. Med. Biol.,* 29**,** 443-447.

Buckingham, R. A. 1938, The classical equation of state of gaseous helium, neon and argon. *P. R. Soc. A,* 168**,** 264-283.

Champion, C. 2003, Theoretical cross sections for electron collisions in water: structure of electron tracks. *Phys. Med. Biol.,* 48**,** 2147-68.

Champion, C., Incerti, S., Aouchiche, H., et al. 2009, A free-parameter theoretical model for describing the electron elastic scattering in water in the Geant4 toolkit. *Radiat. Phys. Chem.,* 78**,** 745-750.

Cho, H., Park, Y. S., Tanaka, H., et al. 2004, Measurements of elastic electron scattering by water vapour extended to backward angles. *J. Phys. B-At. Mol. Opt.,* 37**,** 625-634.

Cobut, V., Frongillo, Y., Patau, J. P., et al. 1998, Monte Carlo simulation of fast electron and proton tracks in liquid water-I. Physical and physicochemical aspects. *Radiat. Phys. Chem.,* 51**,** 229-243.

Czyżewski, Z., MacCallum, D. O. N., Romig, A., et al. 1990, Calculations of Mott scattering cross section. *J. Appl. Phys.,* 68**,** 3066-3072.

Danjo, A. & Nishimura, H. 1985, Elastic scattering of electrons from H2O molecule. *J. Phys. Soc. Jpn.,* 54**,** 1224-1227.

Desclaux, J. P. 1973, Relativistic Dirac-Fock expectation values for atoms with Z=1 to Z=120. *Atom. Data Nucl. Data,* 12**,** 311-406.

Desclaux, J. P. 1975, A multiconfiguration relativistic Dirac-Fock program. *Comput. Phys. Commun.,* 9**,** 31-45.

Desclaux, J. P. 1977, Erratum notice. *Comput. Phys. Commun.,* 13**,** 71.

Dingfelder, M. 2006, Track structure: time evolution from physics to chemistry. *Radiat. Prot. Dosim.,* 122**,** 16-21.

Dingfelder, M., Hantke, D., Inokuti, M., et al. 1998, Electron inelastic-scattering cross sections in liquid water. *Radiat. Phys. Chem.,* 53**,** 1-18.

Dingfelder, M., Hantke, D., Inokuti, M., et al. 1999, Electron inelastic-scattering cross sections in liquid water. *Radiat Phys Chem,* 53**,** 1-18.

Edel, S. 2006. *Modélisation du transport des photons et des électrons dans l'ADN plasmide.* Ph.D. thesis, Toulouse 3.

Emfietzoglou, D., Cucinotta, F. A. & Nikjoo, H. 2005, A complete dielectric response model for liquid water: a solution of the Bethe ridge problem. *Radiat. Res.,* 164**,** 202-211.

Emfietzoglou, D., Karava, K., Papamichael, G., et al. 2003, Monte Carlo simulation of the energy loss of low-energy electrons in liquid water. *Phys. Med. Biol.,* 48**,** 2355.

Emfietzoglou, D. & Moscovitch, M. 2002, Inelastic collision characteristics of electrons in liquid water. *Nucl. Instrum. Meth. B,* 193**,** 71-78.

Emfietzoglou, D., Papamichael, G. & Nikjoo, H. 2017, Monte Carlo electron track structure calculations in liquid water using a new model dielectric response function. *Radiat. Res.,* 188**,** 355-368.

Ester, M., Kriegel, H. P., Sander, J., et al. 1996, A density-based algorithm for discovering clusters in large spatial databases with noise. *Proceeding of the 2nd the International Conference on Knowledge Discovery and Data Mining,* 96**,** 226-231.

Fernández-Varea, J., Liljequist, D., Csillag, S., et al. 1996, Monte Carlo simulation of 0.1–100 keV electron and positron transport in solids using optical data and partial wave methods. *Nucl. Instrum. Meth. B,* 108**,** 35-50.

Francis, Z., Incerti, S., Ivanchenko, V., et al. 2012, Monte Carlo simulation of energy-deposit clustering for ions of the same LET in liquid water. *Phys. Med. Biol.,* 57**,** 209-24.

Francis, Z., Villagrasa, C. & Clairand, I. 2011, Simulation of DNA damage clustering after proton irradiation using an adapted DBSCAN algorithm. *Comput. Meth. Prog. Bio.,* 101**,** 265-70.

Friedland, W., Dingfelder, M., Kundrat, P., et al. 2011, Track structures, DNA targets and radiation effects in the biophysical Monte Carlo simulation code PARTRAC. *Mutat. Res.,* 711**,** 28-40.

Furness, J. B. & McCarthy, I. E. 1973, Semiphenomenological optical model for electron scattering on atoms. *J. Phys. B-At. Mol. Opt.,* 6**,** 2280-2291.

Galassi, M., Davies, J., Theiler, J., et al. 2002, GNU scientific library. *Network Theory Ltd,* 3.

Garcia-Molina, R., Abril, I., Kyriakou, I., et al. 2017, Inelastic scattering and energy loss of swift electron beams in biologically relevant materials. *Surf. Interface Anal.,* 49**,** 11-17.

Grosswendt, B. & Waibel, E. 1978, Transport of low energy electrons in nitrogen and air. *Nucl. Instrum. Meth.,* 155**,** 145-156.

Gryziński, M. 1965, Classical theory of atomic collisions. I. Theory of inelastic collisions. *Phys. Rev.,* 138**,** A336-A358.

Herman, G. T. 1975, A relaxation method for reconstructing objects from noisy X-rays. *Math. Program.,* 8**,** 1-19.

Hilgner, W., Kessler, J. & Steer, E. 1969, Zur Spinpolarisation langsamer Elektronen nach der Streuung an Molekülen II. Wismuttriphenyl und Benzol, Wasser, Tetrachlorkohlenstoff. *Z. Phys. A-Hadron. Nucl.,* 221**,** 324-332.

IAEA 1995. Atomic and Molecular Data for Radiotherapy and Radiation Research. *IAEA-TECDOC-799.* Vienna: International Atomic Energy Agency.

ICRU 1983. Microdosimetry. *ICRU Report 36.* NP ed. Journal of the ICRU: The International Commission on Radiation Units and Measurements.

ICRU 1984. Stopping powers for electrons and positrons. *ICRU Report 37.* NP ed. Journal of the ICRU: The International Commission on Radiation Units and Measurements.

ICRU 2007. Elastic scattering electrons & positrons. *ICRU Report 77.* Journal of the ICRU: The International Commission on Radiation Units and Measurements.

ICRU 2014. Key data for ionizing-radiation dosimetry: measurement standards and applications. *ICRU Report 90.* Journal of the ICRU: The International Commission on Radiation Units and Measurements.

Incerti, S., Baldacchino, G., Bernal, M., et al. 2010a, The Geant4-DNA project. *Int. J. Model. Simul. Sci. Comput.,* 1**,** 157-178.

Incerti, S., Ivanchenko, A., Karamitros, M., et al. 2010b, Comparison of GEANT4 very low energy cross section models with experimental data in water. *Med. Phys.,* 37**,** 4692-4708.

Iskef, H., Cunningham, J. & Watt, D. E. 1983, Projected ranges and effective stopping powers of electrons with energy between 20 eV and 10 keV. *Phys. Med. Biol.,* 28**,** 535.

Jablonski, A., Salvat, F. & Powell, C. J. 2010, NIST electron elastic-scattering cross-section database. *NIST Standard Reference Database,* 64.

Johnstone, W. M. & Newell, W. R. 1991, Absolute vibrationally elastic cross sections for electrons scattered from water molecules between 6 eV and 50 eV. *J. Phys. B-At. Mol. Opt.,* 24**,** 3633.

Kaplan, I. & Sukhonosov, V. Y. 1991, Simulation of the passage of fast electrons and the early stage of water radiolysis by the Monte Carlo method. *Radiat. Res.,* 127**,** 1-10.

Katase, A., Ishibashi, K., Matsumoto, Y., et al. 1986, Elastic scattering of electrons by water molecules over the range 100-1000 eV. *J. Phys. B-At. Mol. Opt.,* 19**,** 2715.

Kellerer, A. M. & Chmelevsky, D. 1975, Concepts of microdosimetry. *Radiat. Environ. Bioph.,* 12**,** 321-335.

Khakoo, M. A., Silva, H., Muse, J., et al. 2013, Erratum: Electron scattering from H2O: Elastic scattering [Phys. Rev. A78, 052710 (2008)]. *Phys. Rev. A,* 87.

Koch, H. W. & Motz, J. W. 1959, Bremsstrahlung cross-section formulas and related data. *Rev. Mod. Phys.,* 31**,** 920-955.

Kutcher, G. & Green, A. 1976, A model for energy deposition in liquid water. *Radiat. Res.,* 67**,** 408-425.

Kyriakou, I., Emfietzoglou, D., Ivanchenko, V., et al. 2017, Microdosimetry of electrons in liquid water using the low-energy models of Geant4. *J. Appl. Phys.,* 122.

Liljequist, D., Liamsuwan, T. & Nikjoo, H. 2012, Elastic scattering cross section models used for Monte Carlo simulation of electron tracks in media of biological and medical interest. *Int. J. Radiat. Biol.,* 88**,** 29-37.

Lindhard, J. 1954, On the properties of a gas of charged particles. *Mat. Fys. Medd. Dan. Vid.,* 28.

Maigne, L., Perrot, Y., Schaart, D. R., et al. 2011, Comparison of GATE/GEANT4 with EGSnrc and MCNP for electron dose calculations at energies between 15 keV and 20 MeV. *Phys. Med. Biol.,* 56**,** 811-27.

Meesungnoen, J., Jay-Gerin, J.-P., Filali-Mouhim, A., et al. 2002, Low-energy electron penetration range in liquid water. *Radiat. Res.,* 158**,** 657-660.

Michaud, M. & Sanche, L. 1987, Total cross sections for slow-electron (1–20 eV) scattering in solidH2O. *Phys. Rev. A,* 36**,** 4672-4683.

Michaud, M., Wen, A. & Sanche, L. 2003, Cross sections for low-energy (1–100 eV) electron elastic and inelastic scattering in amorphous ice. *Radiat. Res.,* 159**,** 3-22.

Mitchell, H. H., Hamilton, T. S., Steggerada, F. R., et al. 1945, The chemical composition of the adult human body and its bearing on the biochemistry of growth. *J. Biol. Chem.,* 158**,** 625-637.

Moliere, G. 1947, Theorie der streuung schneller geladener teilchen i. einzelstreuung am abgeschirmten coulomb-feld. *Zeitschrift für Naturforschung A,* 2**,** 133-145.

Mott, N. F. & Massey, H. S. W. 1965. *The theory of atomic collisions*, Clarendon Press Oxford.

Nikjoo, H., Emfietzoglou, D., Liamsuwan, T., et al. 2016, Radiation track, DNA damage and response-a review. *Rep. Prog. Phys.,* 79**,** 116601.

Nikjoo, H., O'Neill, P., Goodhead, D. T., et al. 1997, Computational modelling of low-energy electron-induced DNA damage by early physical and chemical events. *Int. J. Radiat. Biol.,* 71**,** 467-483.

Nikjoo, H., Uehara, S., Emfietzoglou, D., et al. 2006, Track-structure codes in radiation research. *Radiat. Meas.,* 41**,** 1052-1074.

Nikogosyan, D. N., Oraevsky, A. A. & Rupasov, V. I. 1983, Two-photon ionization and dissociation of liquid water by powerful laser UV radiation. *Chem. Phys.,* 77**,** 131-143.

O'Connell, J. & Lane, N. 1983, Nonadjustable exchange-correlation model for electron scattering from closed-shell atoms and molecules. *Phys. Rev. A,* 27**,** 1893.

Padial, N. T. & Norcross, D. W. 1984, Parameter-free model of the correlation-polarization potential for electron-molecule collisions. *Phys. Rev. A,* 29**,** 1742-1748.

Pages, L., Bertel, E., Joffre, H., et al. 1972, Energy loss, range, and bremsstrahlung yield for 10 keV to 100 MeV electrons in various elements and chemical compounds. *Atom. Data Nucl. Data,* 4**,** 1-127.

Paretzke, H. G., Turner, J. E., Hamm, R. N., et al. 1986, Calculated yields and fluctuations for electron degradation in liquid water and water vapor. *J. Chem. Phys.,* 84**,** 3182-3188.

Penn, D. R. 1987, Electron mean-free-path calculations using a model dielectric function. *Phys. Rev. B,* 35**,** 482-486.

Perdew, J. P. & Zunger, A. 1981, Self-interaction correction to density-functional approximations for many-electron systems. *Phys. Rev. B,* 23**,** 5048-5079.

Pimblott, S. M., LaVerne, J. A., AlbaGarcia, A., et al. 2000, Energy loss by nonrelativistic electrons and positrons in polymers and simple solid hydrocarbons. *J. Phys. Chem. B,* 104**,** 9607-9614.

Pimblott, S. M., LaVerne, J. A. & Mozumder, A. 1996, Monte Carlo simulation of range and energy deposition by electrons in gaseous and liquid water. *J. Phys. Chem. B,* 100**,** 8595-8606.

Pimblott, S. M. & Siebbeles, L. D. A. 2002, Energy loss by non-relativistic electrons and positrons in liquid water. *Nucl. Instrum. Meth. B,* 194**,** 237-250.

Plante, I. & Cucinotta, F. A. 2009, Cross sections for the interactions of 1 eV–100 MeV electrons in liquid water and application to Monte-Carlo simulation of HZE radiation tracks. *New J. Phys.,* 11.

Riley, M. E. & Truhlar, D. G. 1975, Approximations for the exchange potential in electron scattering. *J. Chem. Phys.,* 63**,** 2182-2191.

Rohrlich, F. & Carlson, B. C. 1954, Positron-electron differences in energy loss and multiple scattering. *Phys. Rev.,* 93**,** 38-44.

Rossi, H. H. & Zaider, M. 1996. *Microdosimetry and its Applications*, Springer.

Rowntree, P., Parenteau, L. & Sanche, L. 1991, Electron stimulated desorption via dissociative attachment in amorphous H2O. *J. Chem. Phys.,* 94**,** 8570-8576.

Rudd, M. 1990, Cross sections for production of secondary electrons by charged particles. *Radiat. Prot. Dosim.,* 31**,** 17-22.

Saglam, Z. & Aktekin, N. 1991, Absolute total cross sections for scattering of electrons by H2O in the energy range 4-20 eV. *J. Phys. B-At. Mol. Opt.,* 24**,** 3491.

Salvat, F. 1998, Simulation of electron multiple elastic scattering. *Radiat. Phys. Chem.,* 53**,** 247-256.

Salvat, F. 2003, Optical-model potential for electron and positron elastic scattering by atoms. *Phys. Rev. A,* 68.

Salvat, F., Jablonski, A. & Powell, C. J. 2005, ELSEPA - Dirac partial-wave calculation of elastic scattering of electrons and positrons by atoms, positive ions and molecules. *Comput. Phys. Commun.,* 165**,** 157-190.

Salvat, F. & Mayol, R. 1993, Elastic scattering of electrons and positrons by atoms. Schrödinger and Dirac partial wave analysis. *Comput. Phys. Commun.,* 74**,** 358-374.

Schiff, L. I. 1968. Quantum mechanics. New York: McGraw-Hill.

Seltzer, S. M. & Berger, M. J. 1986, Bremsstrahlung energy spectra from electrons with kinetic energy 1 keV–10 GeV incident on screened nuclei and orbital electrons of neutral atoms with Z= 1–100. *Atom. Data Nucl. Data,* 35**,** 345-418.

Semenenko, V., Turner, J. & Borak, T. 2003, NOREC, a Monte Carlo code for simulating electron tracks in liquid water. *Radiat. Environ. Bioph.,* 42**,** 213-217.

Shin, W. G., Bordage, M. C., Emfietzoglou, D., et al. 2018, Development of a new Geant4-DNA electron elastic scattering model for liquid-phase water using the ELSEPA code. *J. Appl. Phys.,* 124.

Shyn, T. W. & Cho, S. Y. 1987, Vibrationally elastic scattering cross section of water vapor by electron impact. *Phys. Rev. A,* 36**,** 5138-5142.

Shyn, T. W. & Grafe, A. 1992, Angular distribution of electrons elastically scattered from water vapor. *Phys. Rev. A,* 46**,** 4406-4409.

Slater, J. C. 1951, A simplification of the Hartree-Fock method. *Phys. Rev.,* 81**,** 385-390.

Staszewska, G., Schwenke, D. W. & Truhlar, D. G. 1984, Investigation of the shape of the imaginary part of the optical-model potential for electron scattering by rare gases. *Phys. Rev. A,* 29**,** 3078.

Sternheimer, R. M. 1966, Density effect for the ionization loss of charged particles. *Phys. Rev.,* 145**,** 247-250.

Sueoka, O., Mori, S. & Katayama, Y. 1986, Total cross sections for electrons and positrons colliding with H2O molecules. *J. Phys. B-At. Mol. Opt.,* 19**,** L373.

Szmytkowski, C. 1987, Absolute total cross sections for electron-water vapour scattering. *Chem. Phys. Lett.,* 136**,** 363-367.

Terrissol, M. & Beaudre, A. 1990, Simulation of space and time evolution of radiolytic species induced by electrons in water. *Radiat. Prot. Dosim.,* 31**,** 175-177.

Thomas, L. H. 1954, Tables of Statistical Electron Distributions for Atoms with Degree of Ionization Zero to Four and of the Corresponding Electrostatic Potentials. *J. Chem. Phys.,* 22**,** 1758-1767.

Tung, C., Chao, T., Hsieh, H., et al. 2007, Low-energy electron interactions with liquid water and energy depositions in nanometric volumes. *Nucl. Instrum. Meth. B,* 262**,** 231-239.

Uehara, S. & Nikjoo, H. 2006, Monte Carlo simulation of water radiolysis for low-energy charged particles. *Radiat. Res.,* 47**,** 69-81.

Uehara, S., Nikjoo, H. & Goodhead, D. T. 1993, Cross-sections for water vapour for the Monte Carlo electron track structure code from 10 eV to the MeV region. *Phys. Med. Biol.,* 37**,** 1841-1858.

Uehara, S., Nikjoo, H. & Goodhead, D. T. 1999, Comparison and assessment of electron cross sections for Monte Carlo track structure codes. *Radiat. Res.,* 152**,** 202-213.

Vanderpoorten, R. 1975, A local optical potential study of elastic electron-atom scattering. *J. Phys. B-At. Mol. Opt.,* 8**,** 926.

Varella, M. T. d. N., Bettega, M. H. F., Lima, M. A. P., et al. 1999, Low-energy electron scattering by H2O, H2S, H2Se, and H2Te. *J. Chem. Phys.,* 111**,** 6396-6406.

Watt, D. E. 1996. *Quantities for generalized dosimetry of ionizing radiations in liquid water*, CRC Press.

Wiklund, K., Fernandez-Varea, J. M. & Lind, B. K. 2011, A Monte Carlo program for the analysis of low-energy electron tracks in liquid water. *Phys. Med. Biol.,* 56**,** 1985-2003.

Wilson, W. E., Miller, J. H., Lynch, D. J., et al. 2004, Analysis of low-energy electron track structure in liquid water. *Radiat. Res.,* 161**,** 591-596.

Wilson, W. E. & Nikjoo, H. 1999, A Monte Carlo code for positive ion track simulation. *Radiat. Environ. Bioph.,* 38**,** 97-104.

Yates, A. C. 1968, Spin polarization of low-energy electrons scattered elastically from atoms and molecules. *Phys. Rev.,* 176**,** 173.

Zecca, A., Karwasz, G. P., Oss, S., et al. 1987, Total absolute cross sections for electron scattering on H2O at intermediate energies. *J. Phys. B-At. Mol. Opt.,* 20**,** L133-L136.

Zecca, A., Melissa, R., Brusa, R. S., et al. 1999, Additivity rule for electron-molecule cross section calculation: A geometrical approach. *Phys. Lett. A,* 257**,** 75-82.

Zeitler, E. & Olsen, H. 1967, Complex scattering amplitudes in elastic electron scattering. *Phys. Rev.,* 162**,** 1439.

Ziegler, J. 1980. *Handbook of stopping cross-sections for energetic ions in all elements,* Elmsford, NY, Pergamon Press.

Chapter 3

Pre-chemical stage

Table of contents

3.1. Evaluation of the influence of the spatial distribution of molecular species on water radiolysis simulations using Geant4-DNA…….....78
3.1.1. Principles of Geant4-DNA simulation of water radiolysis……..……79
3.1.2. Geant4-DNA elastic scattering models for electrons in liquid water…….81
3.1.3. Geant4-DNA electron thermalization model……..83
3.1.4. Chemistry parameters…….88
3.1.5. The new “chem6” example…….90
3.1.6. Radiochemical yield simulation in water……...93
3.1.7. LET calculations……..…94
3.1.8. Influence of electron elastic scattering models……...…95
3.1.9. Influence of electron thermalization models……...…97
3.1.10. Influence of chemistry parameters……...101
3.2. Evaluation of the impact of the pre-chemical processes…….103
3.2.1. Physical and physico-chemical processes…….……...103
3.2.2. Dissociation channels……...108
3.2.3. Displacement of hot fragments…….……111
3.2.4. Validation study by comparing simulation data with literature data……114
3.2.5. Results for ionization and excitation…….…...117
3.2.6. Results for electron attachment process……...119
3.2.7. Results for electron-hole recombination and dissociation channel……..122
3.2.8. Influence of the pre-chemical processes on water radiolysis simulation...124
3.3. Conclusions……....126
References……..128

Several MCTS tools include the possibility to simulate water radiolysis for the prediction of indirect early DNA damage. For that, a careful study of radiochemical yield simulation in liquid water under irradiation is of strong interest.

It is well-known that the pre-chemical stage which connects the physical and chemical stages of water radiolysis determines the initial radiochemical yields. As of today, it is unfortunately almost impossible to measure femtosecond-scale aspects of the pre-chemical stage. Hence, it is still a common practice, in radiochemical studies, to adjust the pre-chemical stage parameters for calculating G-values (numbers of reactive species divided by the energy deposited in the medium, taken as 100 eV) to match experimental data, due to the lack of a full mechanistic understanding of physico-chemical processes. Such adjustments, however, must be reasonable so that the modeling of the pre-chemical stage could induce beneficial distortions on chemical stage predictions.

In the first section of this chapter, we evaluate the influence of spatial distribution of molecular species on water radiolysis simulations using a new example “chem6” developed in this thesis. Then, in the next section, various methodologies implemented in the different MCTS tools are compared with each other in order to find out an optimized pre-chemistry model for Geant4-DNA.

3.1. Evaluation of the influence of the spatial distribution of molecular species on water radiolysis simulations using Geant4-DNA

The amount of energy loss in the irradiated medium by the incident radiation and all secondary particles directly affects radiochemical yields, as shown by Tran et al. (2016) using Geant4-DNA. In addition, one can intuitively state that the 3D spatial distribution of molecular species also influences results of chemistry simulations. For example, elastic scatterings determine the initial electron 3D concentration which has a close relationship with reaction probabilities of chemical species, even if such scatterings do not lead to any energy loss nor secondary particle creation.

In this section, we thus propose to evaluate the influence on Geant4-DNA radiolysis simulations of the following parameters:

- electron elastic scattering models during the physical stage,
- electron thermalization models during the pre-chemical stage,
- and chemical parameters such as diffusion coefficients and reaction rates during the chemical stage.

For this, we first developed a new Geant4-DNA application dedicated to radiochemical yield simulation, called “chem6”, as an extension of the already existing “chem4” example. In particular, this new application proposes to easily simulate radiochemical yields as a function of time and LET (the latter possibility was not proposed by "chem4"). Then, we calculated

radiochemical yields (G-values) using in particular the new electron elastic scattering cross-section model developed in the previous chapter (Shin et al., 2018), as well as electron thermalization models whose accuracy has been improved, and new chemistry parameters derived from the RE-/RITRACKS code.

3.1.1. Principles of Geant4-DNA simulation of water radiolysis

The MCTS codes that simulate radiolysis, including Geant4-DNA, classically divide the chemistry simulation into three steps: the physical stage (approximately until 1 fs), the pre-chemical stage (until 1 ps), and the chemical stage (1 ps to 1 µs) (Bernal et al., 2015).

a) In the physical stage, the physical interactions between primary particles, secondary particles and medium (water) are calculated. For this, Geant4-DNA uses specific physics constructors, such as the "G4EmDNAPhysics_option2" (see Chapter 2) that gathers all required particles and models describing their physical interactions, down to few electron volts (7.4 to 11 eV), in order to simulate step-by-step (discretely) all physical interactions allowing the evaluation of elementary energy depositions at nanometer scale.

b) Then, in the pre-chemical stage, the water molecules ionized or excited during the physical stage (H_2O^+ or H_2O^*) are dissociated into molecular species based on specified dissociation schemes and branching ratios. Furthermore, the placement of hot fragments generated by these dissociations should also be carefully modelled. Unfortunately, due to the lack of theory or direct validation experiments, Geant4-DNA currently adopts the approach of PARTRAC. The values of PARTRAC have been adjusted based on initial radiochemical yields (implemented in the *G4DNAWaterDissociationDisplacer* class). The hot fragments generated by dissociation processes are placed using a Gaussian distribution and a root-mean-square displacement, noted as r_{rms} as shown in Table 3.1.

Table 3.1: Root-mean-square displacement for the placement of dissociation products from **Bernal et al. (2015)**.

	Hole hopping	**Product 1**	**Product 2**	**Product 3**
$\mathbf{H_3O^+ + {}^{\cdot}OH}$	2 nm (charge transfer)	0.8 nm[a]		-
$\mathbf{H_3O^+ + {}^{\cdot}OH + e_{aq}^-}$ **(auto ionization)**	2 nm (charge transfer)	0.8 nm[a]		Electron thermalization distance
$\mathbf{{}^{\cdot}OH + H^{\cdot}}$	0	-1/18 × 2.4 nm	17/18 × 2.4 nm	-
$\mathbf{H_2 + {}^{\cdot}OH + {}^{\cdot}OH}$	0	-2/18 × 0.8 nm	16/18 × 0.8 nm + 1/2 × 1.1 nm	16/18 × 0.8 nm + 1/2 × 1.1 nm
$\mathbf{H_2 + {}^{\cdot}OH + OH^-}$	0	-2/18 × 0.8 nm	16/18 × 0.8 nm + 1/2 × 1.1 nm	16/18 × 0.8 nm + 1/2 × 1.1 nm

[a] The placement between two products is randomly interchanged by coin tossing (only one product is displaced).

The thermalization distance of electrons generated by auto-ionization process can be determined from an electron thermalization model taking into account ejection energies of electrons. Geant4-DNA is currently using:

- either the "Terrissol" thermalization model for 7 eV electrons in order to calculate the thermalization distance, based on an assumption that at least 10 eV are required to eject the dissociation electron (Beaudre, 1988). This model is used by default in Geant4-DNA.
- or the "Kreipl" thermalization model (as an alternative option) proposed by the paper of Kreipl et al. (2009), but for electrons of 1 nm mean range (as currently coded in Geant4.10.5).

However, a more recent paper proposes to use an energy of 1.7 eV for electrons generated by dissociation processes (Boscolo et al., 2018), based on an empirical study (Han and Bartels, 1990); the thermalization distance for this energy is much shorter than that of the Terrissol model for 7 eV electrons and is much longer than that of the Kreipl model now used in Geant4-DNA.

In addition, electrons that have reached sub-excitation energies (that is, electrons which do not have sufficient kinetic energy to undergo electronic excitation) undergo a thermalization process reaching a 25 meV kinetic energy, where they are assumed to immediately convert into molecular solvated electrons (e^-_{aq}). This thermalization is simulated in a single step by Geant4-DNA (Bernal et al., 2015, Peukert et al., 2019).

c) The algorithm for the simulation of the chemical stage implemented in the most recent version of Geant4-DNA (Geant4.10.5) is well-described in the paper of Karamitros et al. (2014). Briefly, the algorithm consists in a step-by-step approach combining Smoluchowski Brownian diffusion equation (Berg, 1993) describing Brownian motion and Brownian bridge technique. The simulation is divided in successive time steps. The Smoluchowski diffusion equation is used to simulate Brownian transportation of molecular species. Geant4-DNA calculates time steps dynamically (*G4DNAMoleculeEncounterStepper* class) as initially proposed by Michalik et al. (1998). This technique evaluates the probability of a chemical reaction given a selected statistical confidence. For example, the probability $P(d \leq R)$ that the separation distance between two reactants d is smaller than the reaction radius R can be expressed based on the 1-dimensional Smoluchowski diffusion equation at time t as:

$$P(d \leq R) = \mathrm{erf}\left(\frac{d_{95}}{2 \times \sqrt{Dt}}\right) \tag{3.1}$$

where d_{95} is the distance for which 95% of the distance distribution is shorter than d_{95} (P=95% confidence level) with diffusion coefficient D. $\mathrm{erf}(\sqrt{2})$ can be approximated by 0.95, therefore, d_{95} can be obtained from:

$$d_{95} = 2\sqrt{2Dt} \tag{3.2}$$

If the initial separation between the two reactants is d_0 at initial time, then the distance d_Δ at which they can react at time t_Δ is:

$$d_\Delta = d_0 - R = 2\sqrt{2D_A t_\Delta} + 2\sqrt{2D_B t_\Delta} \tag{3.3}$$

Therefore, t_Δ can be expressed for a 95% confidence level as:

$$t_\Delta = \frac{(d_0 - R)^2}{8\left(\sqrt{D_A} + \sqrt{D_B}\right)^2} \tag{3.4}$$

For each time step, the algorithm finds the closest reactant to verify whether the reaction has happened or not. K-d tree algorithm (de Berg et al., 2008) allows to decrease the time complexity from N^2 to $N \times \ln(N)$ using space-partitioning technique for organizing points in a k-dimensional space. Chemical reactions occur when two molecular species are closer than the reaction radius R calculated by the Smoluchowski diffusion equation:

$$R = \frac{k}{4\pi N_A D} \tag{3.5}$$

where N_A is Avogadro's number, k is the effective reaction rate constant considering re-dissociation, and D is the sum of the diffusion coefficients of molecules.

After all chemical reactions have been processed, molecular species are diffused based on their diffusion coefficients and Brownian diffusion equation. The diffusion equation in one dimension with diffusion coefficient D at time step t can be derived using a Gaussian distribution:

$$P(x, t|x_0) = \frac{1}{\sqrt{2\pi}\sigma_x} e^{-\frac{(x-x_0)^2}{2\sigma_x^2}} = \frac{1}{\sqrt{4\pi D t}} e^{\frac{(x-x_0)^2}{4Dt}} \tag{3.6}$$

However, due to the discretization of time steps, we should carefully consider the possible reactions between two time steps. For example, the two reactants are separated by more than reaction radius R at pre-step and post-step point, which are the initial and final times of a time step, respectively. However, it is possible that the separation is less than R during a time step. In order to account for these possible reactions, Brownian bridge technique is implemented in Geant4-DNA (*G4DNASmoluchowskiReactionModel* class). The probability of the possible reaction in the 1-dimentional case with the assumption that the second species is motionless is:

$$p(t_R \leq t \,|d_0, d_f) = \mathrm{e}^{-\frac{d_f d_0}{Dt}} \tag{3.7}$$

where d_0 and d_f are the initial (pre-step) and final (post-step) separation distances from the reaction sphere of the other species, respectively. Geant4-DNA compares a random number to the probability, and if the random number is smaller than this probability, then the reaction occurs.

3.1.2. Geant4-DNA elastic scattering models for electrons in liquid water

Geant4-DNA provides three recommended physics constructors (G4EmDNAPhysics_option2, 4, 6) (Incerti et al., 2018). The default physics constructor (G4EmDNAPhysics_option2) includes a partial wave elastic scattering model (Champion et al., 2009) which presents important limitations such as the lack of relativistic correction, a limited energy range of applicability, and disagreement of DCSs with experimental data at low energy

and intermediate angle (see Chapter 2). In order to improve the modeling of electron elastic scattering in liquid water, we developed a new model using the ELSEPA partial-wave code developed by Salvat et al. (2005), that we recently published (Shin et al., 2018). We remind that the G4EmDNAPhysics_option4 constructor is using the screened Rutherford elastic scattering model with the screening parameter of Uehara et al. (1993), and the G4EmDNAPhysics_option6 constructor adopts the elastic scattering model of CPA100 (Bordage et al., 2016). Another constructor is also available as a prototype in Geant4-DNA: the G4EmDNAPhysics_option8 constructor, which contains the same inelastic models as the G4EmDNAPhysics_option2 constructor and a combination of the CPA100 elastic model in the range (11 eV-256 keV) and the Champion elastic model in the range (256 keV-1 MeV).

In this study, we propose to study the influence of the electron elastic scattering model on the simulation of radiochemical yields using the "chem6" example. The simulations are performed with 1 MeV incident electrons, for which experimental data from the literature exist. Only the G4EmDNAPhysics_option2 and G4EmDNAPhysics_option8 constructors can be used to simulate electrons up to 1 MeV. We summarize in Table 3.2 the processes and models contained in these two constructors, as well as in the G4EmDNAPhysics_option2 constructor in which we have replaced the Champion elastic model by the newly developed ELSEPA-based elastic model.

Table 3.2: Physics constructors, processes, and models used to simulate radiochemical yields for electrons.

Physical process	**G4EmDNAPhysics _option2[a]**	**G4EmDNAPhysics _option2 with ELSEPA elastic cross-section model[a]**	**G4EmDNAPhysics _option8[a]**
Excitation	G4DNABornExcitationModel		
Ionization	G4DNABornIonisationModel		
Vibrational excitation	G4DNASancheExcitationModel		
Dissociative attachment	G4DNAMeltonAttachmentModel		
Elastic scattering	Champion model	ELSEPA model	CPA100 model[b] Champion model[c]
Low energy limit	7.4 eV	10 eV	11 eV
High energy limit	1 MeV	1 MeV	1 MeV

[a] Abbreviated later as “physics option 2”, “physics option 2E”, and “physics option 8”
[b] below 255 keV
[c] above 255 keV

3.1.3. Geant4-DNA electron thermalization model

The simulation of transportation of low energy electrons (typically < 10 eV, so-called "sub-excitation electrons") is a computational burden in MCTS codes due to the small associated energy depositions and large cross-sections (e.g. elastic scattering). In order to reduce the simulation time, MCTS codes usually simulate in one single step the distance travelled by electrons until thermalization (25 meV). A class named *G4DNAOneStepThermalization* is responsible for the placement of such electrons in one step. Once thermalized, these electrons are assumed to transform into solvated electrons. Geant4-DNA currently provides 3 models for electron thermalization proposed by Terrissol and Beaudre (1990), Ritchie et al. (1994), Meesungnoen et al. (2002) called below as Terrissol model, Ritchie model, Meesungnoen model, respectively. These thermalization models use fitting functions of thermalization range based on experimental data or Monte Carlo results as described here below.

The Terrissol model (1990)

The Terrissol model is based on the data of Michaud and Sanche (1987a), Michaud and Sanche (1987b), on sub-excitation electron cross-sections in amorphous-solid water. The paper of Terrissol and Beaudre (1990) suggests probability density functions (PDF) of thermalization distance for electron energies in the range 0.2 - 7 eV. They use two fitting functions (Beaudre, 1988):

- Below 2 eV, a gamma distribution fitting model called "modified exponential 3" (noted as "PE3") $f_{gamma}(r)$ as a function of r is proposed as:

$$f_{gamma}(r) = \frac{r^2}{2b^3} e^{-\frac{r}{b}} \tag{3.8}$$

  where b is an energy-dependent parameter. Based on equation (3.8), the mean value of r and the mean value of r^2 can be derived as:

$$< r >= \int_0^\infty \frac{r^3}{2b^3} e^{-\frac{r}{b}} dr = 3b \tag{3.9}$$

$$< r^2 >= \int_0^\infty \frac{r^4}{2b^3} e^{-\frac{r}{b}} dr = 12b^2 \tag{3.10}$$

  and the corresponding standard deviation (SD) σ is:

$$\sigma = \sqrt{< r^2 > -(< r >)^2} = \sqrt{12b^2 - (3b)^2} = \sqrt{3}b \tag{3.11}$$

- Above 2 eV, a Gaussian fitting function (noted as "G") for thermalization distance $f_{Gaus}(r)$ is proposed:

$$f_{Gaus}(r) = \frac{4r^2}{\sqrt{\pi}b^3} e^{-\frac{r^2}{b^2}} \tag{3.12}$$

  The mean range and mean square range can be derived based on equation (3.12):

$$< r >= \int_0^\infty \frac{4r^3}{\sqrt{\pi}b^3} e^{-\frac{r^2}{b^2}} dr = \frac{2b}{\sqrt{\pi}} \tag{3.13}$$

$$< r^2 >= \int_0^\infty \frac{4r^4}{\sqrt{\pi}b^3} e^{-\frac{r^2}{b^2}} dr = \frac{3}{2}b^2 \tag{3.14}$$

and the SD σ is:

$$\sigma = \sqrt{< r^2 > -(< r >)^2} = \sqrt{\frac{3}{2}b^2 - \left(\frac{2b}{\sqrt{\pi}}\right)^2} = b\sqrt{\frac{3-\frac{8}{\pi}}{2}} \tag{3.15}$$

Table 3.3: Mean and SD values from Terrissol and Beaudre (1990) and *b* values calculated analytically.

Energy (eV)	**Mean range in the paper (Å)**	**SD in the paper (Å)**	***b* value in the paper (Å)**	***b* value calculated from the mean values of column 2 (Å)**	***b* value calculated from the SD values of column 3 (Å)**	**Fitting curve model**
0.2	31.74	17.68	10.6	10.6	10.2	PE3
0.5	41.50	22.30	-	13.8	12.9	PE3
1	56.03	28.49	-	18.7	16.4	PE3
2	92.04	45.35	-	30.7	26.2	PE3
				81.6	95.2	G
3	144.04	70.03	-	127.7	147.1	G
4	204.74	98.05	-	181.4	205.9	G
5	256.10	120.56	-	227.0	253.2	G
6	284.13	132.73	-	251.8	278.7	G
7	307.16	142.60	272.2	272.2	299.5	G

The Terrissol model in the current version of Geant4-DNA has some limitations. First, Geant4-DNA uses the SD values of Terrissol and Beaudre (1990) interpolated linearly and equation (3.15), even though the values for SD of Terrissol and Beaudre (1990) are not consistent with the *b* value also found in that reference as shown in Table 3.3. We note however that the values for the mean found in Terrissol and Beaudre (1990) exactly match with the b values proposed in the reference. Second, only the Gaussian fitting model is taken into account in Geant4-DNA. Moreover, we need to extend this fitting model up to 10 or 11 eV in order to use it in combination with the ELSEPA or CPA100 electron elastic model; unfortunately, Terrissol and Beaudre (1990) only provides data up to 7 eV. Finally, the cross sections for amorphous ice were updated by Michaud et al. (2003).

Thus, in order to re-implement this thermalization algorithm for Geant4-DNA, we first use *b* values based on the mean values table of Terrissol and Beaudre (1990), which are consistent with their data, with equation (3.13) and using a log-log interpolation. In addition, we take the Gamma fitting function into account below 2 eV using the *G4RandGamma* class, provided by Geant4 for the generation of random numbers. Then, the random thermalization range based on Gamma distribution is generated with α value of 3 and λ of $\frac{1}{b}$, assuming a random direction. Above 2 eV, the Gaussian fitting model of Terrissol model follows a similar approach with the paper of Goulet and Jay-Gerin (1989) describing Gaussian fitting function for thermalization distance, $f_{Gaus}(r,\sigma)$:

$$f_{Gaus}(r,\sigma) = G_{1D}(x,\sigma_{1D}) \times G_{1D}(y,\sigma_{1D}) \times G_{1D}(z,\sigma_{1D}) \times 4\pi r^2 = \frac{\sqrt{2}}{\sqrt{\pi}\sigma_{1D}^3} r^2 e^{-\frac{r^2}{2\sigma_{1D}^2}} \quad (3.16)$$

where $G_{1D}(x,\sigma_{1D})$, $G_{1D}(y,\sigma_{1D})$, $G_{1D}(z,\sigma_{1D})$ are 1D Gaussian distributions corresponding to x, y, z directions. The equation (3.12) can be derived similarly as equation (3.16) with the assumption that $b^2 = 2\sigma_{1D}^2$. This means that the distribution of thermalization range can be reproduced from three random Gaussian distributions (using the Geant4 *G4RandGauss* class for random number generation) for the x, y, and z axes and with 1-dimentional sigma σ_{1D}.

In addition, in order to extend the Terrissol model up to 10 eV, the correlation between electron energy E and b values above 2 eV is fitted as a cubic function with the hypothesis that the thermalization range of the Terrissol model is longer than that of the Meesungnoen model with a correction factor of 2 in order to take into account phase effects:

$$b_{Gaus}(E) = -0.91E^3 + 7.9E^2 + 28E + 0.6 \quad (3.17)$$

where $b_{Gaus}(E)$ is expressed in Å, and E is the energy of the electrons in eV.

We suggest to keep the electron thermalization model for amorphous ice that could be of interest for biomolecular experiments (e. g. radiation transport in cosmic ice (Dubochet et al., 1988)). However, in the other hand, we also propose to replace the Terrissol model for amorphous ice with a new thermalization model: the “Meesungnoen amorphous” model, as described below.

The Ritchie model (1994)

Ritchie et al. (1994) approximated that the relationship between electron energy *E* and mean value of "projected range" *x*, which is $1/\sqrt{3}$ smaller than the corresponding range in three dimensions, has a linear correlation, based on experimental data on projected range of thermalized electrons proposed by Rotenberg and Gurevich (1975), Neff et al. (1980), Kreitus et al. (1982), Konovalov et al. (1985), Konovalov et al. (1986). However, MCTS codes such as TOPAS-nBio, previous version of PARTRAC, and Geant4-DNA (Ballarini et al., 2000, Ramos-Mendez et al., 2018) have been using this correlation without the $\sqrt{3}$ correction factor, as:

$$< r >= 1.8 \times E \tag{3.18}$$

$< r >$ in equation (3.18) represents the $< x >$ quantity defined in the Ritchie et al. (1994), so we propose to correct the Ritchie model as:

$$< r >= 1.8 \times \sqrt{3} \times E \tag{3.19}$$

In order to implement the Richie model in Geant4-DNA, we use exactly the same approach as the Gaussian fitting model of the Terrissol model based on equation (3.16). The mean range can be expressed with the integration of 3-dimentional Gaussian distribution multiplying by *r* as:

$$< r >= \int_0^\infty \frac{\sqrt{2}}{\sqrt{\pi}\sigma_{1D}^3} r^3 e^{-\frac{r^2}{2\sigma_{1D}^2}}\, dr = \frac{2^{\frac{3}{2}}\sigma_{1D}}{\sqrt{\pi}} \tag{3.20}$$

Therefore, the 1-dimentional SD σ_{1D} for x, y, and z axis is:

$$\sigma_{1D} =< r >\times \frac{\sqrt{\pi}}{2^{\frac{3}{2}}} \tag{3.21}$$

The Meesungnoen model (2002)

The Meesungnoen model is the most recent model for electron thermalization, and it is used as the default thermalization model in Geant4-DNA. Meesungnoen et al. (2002) performed MC simulations using the more recent cross-sections of Michaud et al. (2003) for amorphous ice scaled to take into account the differences between amorphous and liquid-phase water. Geant4-DNA is currently using a 12th order fitting curve to reproduce the thermalization range of Meesungnoen model, and the model, based on 3-dimentional Gaussian fitting function as in equation (3.16), well reproduces published results (Meesungnoen et al., 2002).

In addition, we also developed an alternative version of this model, using cross sections of Michaud et al. (2003) for amorphous ice (thus, without the correction of phase influence) in order to replace the Terrissol model (applicable to amorphous ice). We named this model as the "Meesungnoen amorphous" model. Compared to the Terrissol model, this alternative model is free from limitations of the Terrissol model including energy limit, fitting errors, and using outdated cross sections.

The Kreipl model (2009)

PARTRAC proposes to use another PDF to describe the thermalization of electrons (Kreipl et al., 2009). The Kreipl model is based on the mean range data of the Meesungnoen model, however it uses different fitting function to describe the PDF of thermalization distance:

$$f(r) = 4re^{-2r} \tag{3.22}$$

The mean value of this gamma distribution with α of 2 and λ of 2 is 1, then we can sample a random point (using the G4RandGamma(2,2) method) and weight the random value by the mean thermalization range of the Meesungnoen model.

3.1.4. Chemistry parameters

The default chemistry list in Geant4-DNA (G4EmDNAChemistry, abbreviated as "chemistry default" later) is based on the PARTRAC MCTS (Friedland et al., 2011). We will also use the term "chemistry constructor", by analogy with physics constructors. It gathers all important chemical parameters for the simulation of radiolysis, such as list of molecular species and their diffusion coefficient, as well as list of chemical reactions and corresponding reaction rates. This chemistry list, which is available since Geant4 version 10.1 (end of 2014), has shown some lack of accuracy in reproducing experimental radiochemical yields (Peukert et al., 2019). In collaboration with the TOPAS-nBio group (Ramos-Mendez et al., 2018), which is member of the Geant4-DNA collaboration since 2016, we proposed a new chemistry list (G4EmDNAChemistry_option1, later abbreviated as "chemistry option 1") taken from the RE-/RITRACKS MCTS (Plante and Devroye, 2017).

Table 3.4: Dissociation schemes and branching ratios for default and option 1 chemistry lists.

Process		Type	Dissociation channels	Probability
Ionization		Dissociative decay	$H_3O^+ + {}^{\cdot}OH$	1
Excitation	A1B1	Dissociative decay	${}^{\cdot}OH + H^{\cdot}$	0.65
		Relaxation	$H_2O + \Delta E$	0.35
	B1A1	Auto-ionization	$H_3O^+ + {}^{\cdot}OH + e^-_{aq}$	0.55
			${}^{\cdot}OH + {}^{\cdot}OH + H_2$	0.15
		Relaxation	$H_2O + \Delta E$	0.3
Rydberg, diffusion bands		Auto-ionization	$H_3O^+ + {}^{\cdot}OH + e^-_{aq}$	0.5
		Relaxation	$H_2O + \Delta E$	0.5
Dissociative attachment (H_2O^-)		Dissociative decay	${}^{\cdot}OH + OH^- + H_2$	1
Electron-hole recombination		Dissociative decay	${}^{\cdot}OH + {}^{\cdot}OH + H_2$	0.15
		Dissociative decay	${}^{\cdot}OH + H^{\cdot}$	0.55
		Relaxation	$H_2O + \Delta E$	0.3

In the pre-chemical stage, the water molecules excited or ionized in the physical stage generate initial molecular species based on dissociation schemes and branching ratios listed in

Table 3.4. We used the same dissociation schemes and branching ratios as PARTRAC, not only in default chemistry, but also in chemistry option 1. These values are based on the data of Cobut et al. (1998) with small modifications in order to match up with the picosecond yields of each species (Ballarini et al., 2000, Kreipl et al., 2009) due to the lack of information about branching ratios.

Table 3.5: Diffusion coefficients of default and option 1 chemistry lists at 25 °C.

Molecular species	**Diffusion coefficient (10^{-9} m^2s^{-1})**	
	G4EmDNAChemistry	G4EmDNAChemistry_option1
e_{aq}^-	4.9	4.9
$^{\cdot}OH$	2.8	2.2
$H^{\cdot}$	7.0	7.0
H_3O^+	9.0	9.46[a]
H_2	5.0	4.8
OH^-	5.0	5.3
H_2O_2	1.4	2.3

[a] denoted as H^+ in the paper of Plante and Devroye (2017)

Table 3.6: Reaction rate constants k_{obs} (10^{10} $M^{-1}s^{-1}$).

Reaction	**Reaction rate constant k_{obs} (10^{10} $M^{-1}s^{-1}$)**	
	G4EmDNAChemistry	G4EmDNAChemistry_option1
$e_{aq}^- + e_{aq}^- \rightarrow H_2 + 2OH^-$	0.5	0.636
$e_{aq}^- + H^{\cdot} \rightarrow H_2 + OH^-$	2.65	2.5
$e_{aq}^- + {}^{\cdot}OH \rightarrow OH^-$	2.95	2.95
$e_{aq}^- + H_3O^+ \rightarrow H^{\cdot}$	2.11	2.11[a]
$e_{aq}^- + H_2O_2 \rightarrow OH^- + {}^{\cdot}OH$	1.41	1.10
${}^{\cdot}OH + {}^{\cdot}OH \rightarrow H_2O_2^-$	0.44	0.550
${}^{\cdot}OH + H^{\cdot} \rightarrow H_2O$	1.44	1.55
$H^{\cdot} + H^{\cdot} \rightarrow H_2$	1.2	0.503
$H_3O^+ + OH^- \rightarrow H_2O$	14.3	11.3[a]

[a] H_3O^+ is denoted as H^+ in the paper of Plante and Devroye (2017)

From 1 picosecond, the chemical stage, in charge of transporting molecular species and performing chemical reactions, starts using the diffusion coefficients and reaction rates shown in Tables 3.5 and 3.6, respectively. The diffusion coefficients and reaction rates of default chemistry are selected from the paper of Buxton et al. (1988), Frongillo et al. (1998) with their own adjustment for the agreement with literature data. It is also described in the papers of PARTRAC (Ballarini et al., 2000, Kreipl et al., 2009). In the case of chemistry option1, the diffusion coefficients and reaction rates are partially the values reported by Frongillo et al. (1998) based on the previous study of Elliot (1994).

In this work, G-values calculated with the different chemistry options are compared with each other in order to verify the influence of the chemistry parameters. Moreover, an alternative constructor for chemistry option 1 in which we replaced the diffusion coefficients with those of

default chemistry is also proposed in order to evaluate the influence of diffusion coefficients values on radiochemical yield simulations.

3.1.5. The new "chem6" example

Geant4-DNA provides many sets of examples (Bernal et al., 2015, Incerti et al., 2018) for physics and chemistry simulations. For the chemistry simulation, the example "chem4" can be used in order to calculate time evolution of radiolytic species. However, "chem4" example does not provide the calculation of radiochemical yields as a function of LET, a requirement for a complete benchmarking of radiolysis simulations. In addition, it is difficult to use. In this study, a new example, called "chem6", is proposed for radiochemical yield simulation versus time and LET, including a full macro control.

Primary killer

In the case of electron simulations, water radiolysis typically uses only small segment of the entire physical track in order to simulate enough number of tracks with reasonable CPU times, and it is reported that the radiochemical yields are not significantly affected by this approximation (Pimblott and LaVerne, 1998). There are several approaches restricting the physical track such as small sensitive volume (Boscolo et al., 2018). The *PrimaryKiller* class has been developed in the Geant4-DNA "chem4" example in order to restrict the energy deposition of primary electrons (Karamitros et al., 2011). In order to restrict such energy deposition, two parameters are used for energy threshold selection, the minimum energy deposition T1 and the maximum energy deposition T2. Primary particles losing their energy more than T1 are killed in the simulation; in addition, if the total energy deposition of a simulation event (that is, the primary particle and all associated secondaries) is larger than T2, the event is aborted (thus it is fully ignored in the simulation). In other words, the total energy deposition of simulated events is always between T1 and T2. The T1 and T2 values should be carefully selected for impact on the accuracy of the chemistry simulation, in this study, we selected the same T1 and T2 values as recently suggested for TOPAS-nBio (Ramos-Mendez et al., 2018).

Time step limit

Geant4-DNA uses dynamic time steps in order to calculate proper time step based on the stochastic calculation. We can also apply static time steps in order to restrict the minimum time step allowing to ignore negligible time steps (in the same spirit as we use tracking cuts during physical stage). The "chem6" example can control static time steps in the *UserTimeStepAction* class, and, for reference, we used in this work exactly the same static time steps as proposed by Kreipl et al. (2009), Karamitros et al. (2011) and shown in Table 3.7.

Table 3.7: Time step limits proposed by Kreipl et al. (2009).

Time interval (ps)	Minimum time step (ps)
Until 10	0.1
10-100	1
100-1,000	3
1,000-10,000	10
Above 10,000	100

LET calculation

In order to evaluate the effects of radiation quality on the number of chemical species, radiobiology studies have tried to link radiation effectiveness to a single physical parameter such as mean unrestricted linear energy transfer (LET) (Belli et al., 2009), mean-free-path (MFP) of primary ionizing particle (Chen and Watt, 1985), mean restricted LET (Blohm and Harder, 1985), and beta (Sauer Jr et al., 1977). The "chem6" example provides a new class for the calculation of LET (new *ScoreLET* class), and users can calculate such physical parameter during the chemistry simulation.

Several approaches for LET have been proposed like dose-average LET (Cortes-Giraldo and Carabe, 2015), track-average LET, restricted LET, and unrestricted LET.

- **Dose-average LET** (LET_d) can be calculated from the energy deposition of i^{th} charged particle (event) ε_i and the tracking step length l_i:

$$LET_d = \sum_{i=1}^{n} \left(\frac{\varepsilon_i}{l_i}\right) \omega_{i,d} \tag{3.22}$$

where $\omega_{i,d}$ is the dose weighting factor for the ith event, expressed as:

$$\omega_{i,d} = \frac{D_i}{\sum_{i=1}^{n} D_i} = \frac{\varepsilon_i/m}{\sum_{i=1}^{n} \varepsilon/m} = \frac{\varepsilon_i}{\sum_{i=1}^{n} \varepsilon_i} \tag{3.23}$$

where D_i is the dose deposition by the i^{th} charged particle of the specified type within the sensitive volume with the mass of m. Therefore, LET_d can be expressed as:

$$LET_d = \frac{\sum_{i=1}^{n} \frac{\varepsilon_i^2}{l_i}}{\sum_{i=1}^{n} \varepsilon_i} \tag{3.24}$$

- **Track-average LET** (or fluence-average LET, LET_t) is the arithmetic mean value of the fluence spectrum of LET. It is essentially very close to the physical meaning of the LET concept defined by ICRU. LET_t is calculated as:

$$LET_t = \sum_{i=1}^{n} \left(\frac{\varepsilon_i}{l_i}\right) \omega_{i,t} \tag{3.25}$$

where $\omega_{i,t}$ is the track-length weighting factor for the i^{th} event:

$$\omega_{i,t} = \frac{l_i}{\sum_{i=1}^{n} l_i} \tag{3.26}$$

Therefore, LET_t can be derived as:

$$LET_t = \frac{\sum_{i=1}^{n} \varepsilon_i}{\sum_{i=1}^{n} l_i} \tag{3.27}$$

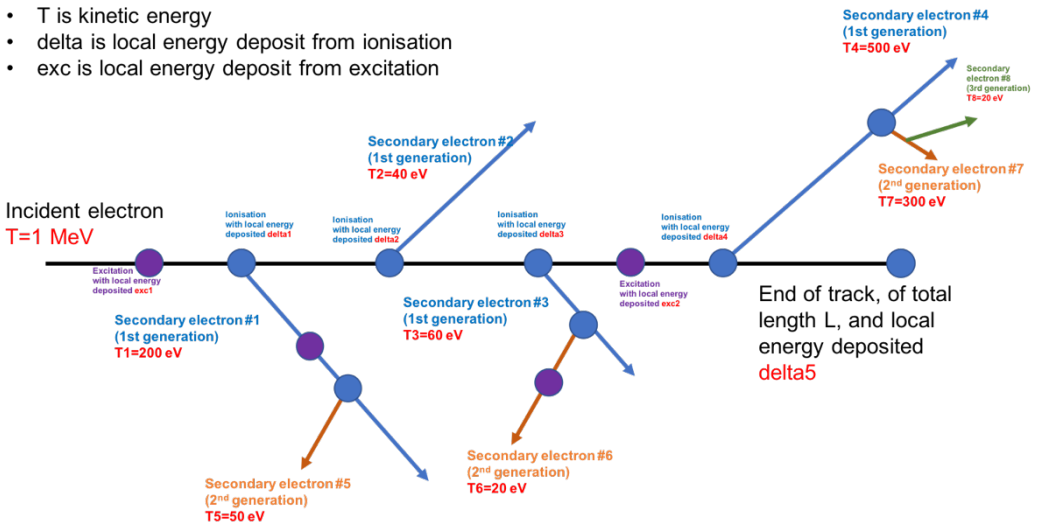


Figure 3.1: A schematic illustration of a trajectory of an incident electron (assumed to be straight for better clarity).

In the ***ScoreLET*** class, the track-average LET is calculated with two methods: restricted and unrestricted LET. The definition of track-average restricted LET is the sum of the energy depositions by primary particle and the kinetic energies of the first generation of secondary particles below than an energy cut-off, divided by the track segment length of the primary particle. Figure 3.1 illustrates the principle of these two calculations, expressed as for 100 eV restricted LET:

$$LET_{100eV} = \frac{exc_1 + exc_2 + delta_1 + delta_2 + delta_3 + delta_4 + T_2 + T_3}{L} \tag{3.28}$$

and for unrestricted LET (which is equal to collision stopping power) as:

$$LET_{\infty} = \frac{exc_1 + exc_2 + delta_1 + delta_2 + delta_3 + delta_4 + T_1 + T_2 + T_3 + T_4}{L} \tag{3.29}$$

In this section, we calculated the track-average restricted LET with a cut-off energy of 100 eV, which is reported to have a close relation with radiation effectiveness (Blohm and Harder, 1985), for radiochemical yields ("G-values") at 1 μs (typically used in radiolysis studies).

User macro commands available in "chem6"

Table 3.8: User macro commands.

Class	Command	Parameters	Description
Detector Construction	/det/setSize	x y z unit	Target size e.g. infinite for electrons, or 5 x 5 x 5 μm^3 for protons
ScoreLET	/scorer/LET/method	LET_res LET_unres	LET calculation method (restricted or unrestricted)
	/scorer/LET/cutoff	cut-off unit	Cut-off energy for restricted LET calculation
ScoreSpecies	/scorer/species/addTimeToRecord	time unit	Add a time point to score G-values
	/scorer/species/nOfTimeBins	number	Time bin is divided into this number
PhysicsList	/phys/addPhysics	dna_opt2 …	Selection of physics constructor
	/process/dna/e-SolvationSubType	Kreipl2009 Meesungnoen2002 Meesungnoen2002_ amorphous Ritchie1994 Terrisol1990	Selection of one step thermalization model

In the new example “chem6”, users can control several simulation conditions using user macro files (that is, text files containing Geant4 macro commands, that can be read by the example) as shown in Table 3.8. We also kept the other user macros already available in "chem4" such as "primaryKiller" and "scheduler" to specify the energy threshold and time parameters, respectively.

3.1.6. Radiochemical yield simulation in water

In this study, the most recent combination of physics constructor G4EmDNAPhysics_option2E, chemistry constructor G4EmDNAChemistry_option1, and the Meesungnoen electron thermalization model is selected as default combination, and the influences of electron elastic scattering models, electron thermalization models, and chemical parameters are evaluated by comparing G-values versus time and G-values versus LET with each other.

Sets of calculation and experimental data presented by Buxton (1972), Wolff et al. (1973), Draganić and Draganić (1975), Burns et al. (1981), Sumiyoshi and Katayama (1982), LaVerne and Pimblott (1991), Belloni et al. (1983), Elliot et al. (1993), Tomita et al. (1997), Pimblott and LaVerne (1997), Bartels et al. (2000), Jay-Gerin and Ferradini (2000), LaVerne (2000), Muroya et al. (2005), Wang et al. (2018) are used for comparison with the simulated G-values versus time. For G-values versus LET, calculation and experimental data obtained by Schwarz et al. (1959),

Naleway et al. (1979), Burns and Sims (1981), Wasselin-Trupin et al. (2002), Ramos-Mendez et al. (2018) are used. It should be noted that most of the experimental data are measured the G-values under solvent concentration arbitrary controlled, whereas the simulations are performed for pure liquid water.

3.1.7. LET calculations

Table 3.9: LET calculations.

Energy (keV)	**T1 (keV)**	**T2 (keV)**	**N of particles (passing T2)**	**LET_{100eV} (keV/um)**	**σ of LET_{100eV} (keV/um)**
2	1.2	1.212	30,000 (8,500)	6.321	1.624
3.5	1.6	1.616	30,000 (9,600)	3.736	0.883
7.5	2.3	2.323	30,000 (11,400)	1.834	0.367
12.5	3.8	3.838	5,000 (2,400)	1.189	0.182
30	6.0	6.06	3,000 (1,600)	0.541	0.070
80	8.0	8.08	3,000 (1,600)	0.244	0.025
999.999[a]	10.0	10.1	3,000 (1,700)	0.084	0.008

[a] Representing 1 MeV

Table 3.9 shows the kinematic simulation conditions and the results of LET calculations. These results are based on the default options (G4EmPhysics_option2E with ELSEPA elastic model, G4EmChemistry_option1, Meesungnoen electron thermalization model), but the other options show almost same result because we only changed the electron elastic scattering model, electron thermalization model, and chemistry options, which are irrelevant with the energy loss of particles determining LET values. The results of restricted-LET calculation (with 100 eV cut-off) are in good agreement at low energy (2 keV) with the values given by the ICRU 16 report (ICRU, 1970) but show some discrepancy at larger energies (30 keV and 1 MeV) . It should be noted that the ICRU 90 report provides a more recent dataset for LET values, however, restricted LET values are unfortunately not provided (ICRU, 2014). We remind that non-restricted LET calculations obtained with Geant4-DNA agree with the ICRU 90 recommendations (Incerti et al., 2018).

3.1.8. Influence of electron elastic scattering models

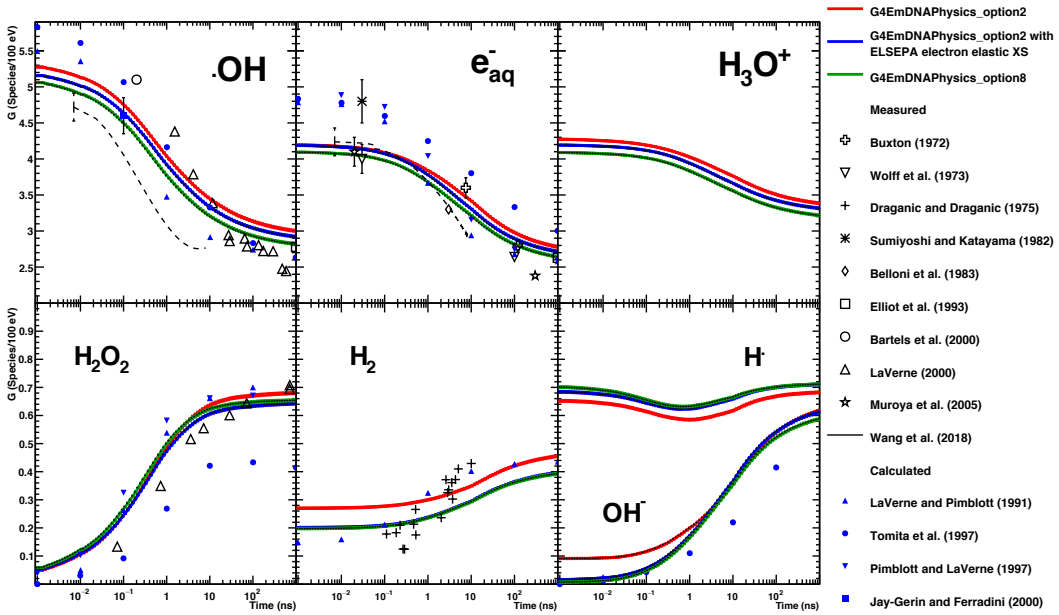


Figure 3.2: Time evolution of G-values according to the different electron elastic scattering models from Shin et al. (2019).

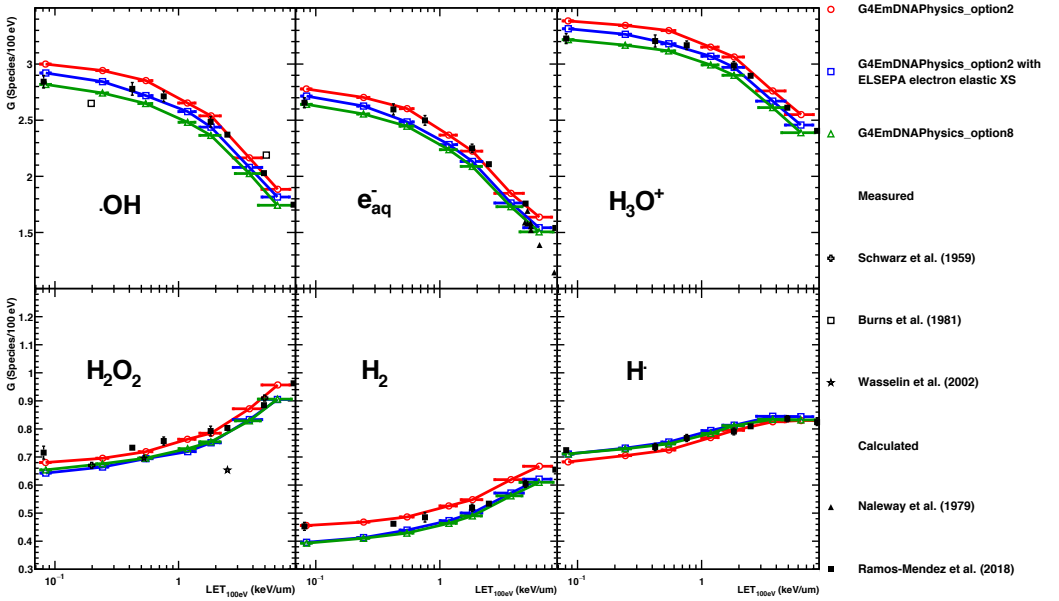


Figure 3.3: G-values versus LET according to the different electron elastic scattering models from Shin et al. (2019).

Figure 3.2 and 3.3 show the G-values versus time and LET according to the different electron elastic scattering models. We remind that the elastic model of Champion used in physics option 2 is the most diffusive model (Shin et al., 2018) and CPA100 used in physics option 8 is the most concentrated model (Shin et al., 2018). In other words, all options lead to same amount of

dissociations, but the Champion model induces the smallest number of chemical reactions, and on the opposite, the CPA100 model induces the largest number.

Indeed, for instance, H_3O^+ and e_{aq}^- are only generated by the dissociation process during the chemical stage simulation, and they vanish through chemical reactions. Therefore, G-values for H_3O^+ and e_{aq}^- simulated with the Champion elastic model are the largest, and those simulated with the CPA100 model show the smallest values. We also observed that the G-values of minor species such as H_2O_2, H_2, and H• radicals with physics option 2 disagree with the other options, even if option 2 and option 8 use the same elastic scattering model from 255 keV and above. It can thus be concluded that these molecular species are dominantly affected by secondary electrons and those low energy elastic cross-section.

The G-values calculated with physics option 2 are quite different with the other curves and experimental data. We can guess that the disagreements arise from the inaccuracy of the Champion elastic cross-section below 50 eV because physics option 8 also uses the same Champion elastic cross-section above 255 keV. The curves for physics options 2E and 8 match well with experimental data, especially with recent data such as Wasselin-Trupin et al. (2002), Muroya et al. (2005), Ramos-Mendez et al. (2018). The differences between physics option 2E and 8 are not huge, however, we don't recommend to use the combination of elastic physics models as implemented in option 8 due to the discontinuity between the two models at 255 keV. In addition, the CPA100 electron elastic model shows larger angular differential cross-section values below 50 eV, especially at intermediate and large scattering angles, while the ELSEPA models shows a better agreement (Shin et al., 2018).

3.1.9. Influence of electron thermalization models

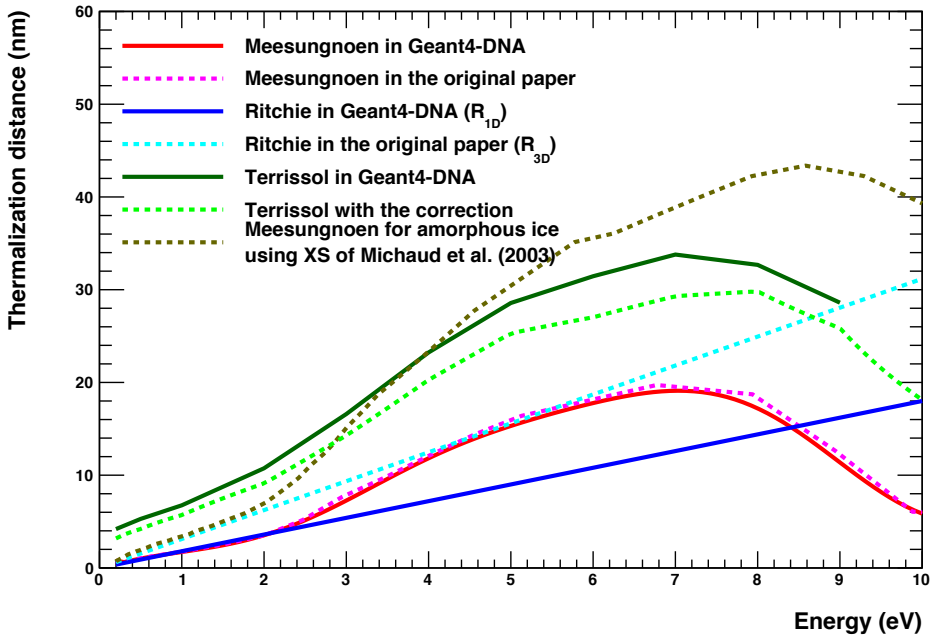


Figure 3.4: Electron thermalization distance as a function of electron energy.

Figure 3.4 shows the thermalization distance as a function of energy calculated by the three thermalization models as described in the original references and as implemented in Geant4-DNA. The thermalization curve for the Meesungnoen model implemented in Geant4-DNA is well matching the original publication data (Meesungnoen et al., 2002). The 12th order fitting model is able to nicely reproduce the range data, and the thermalization algorithm using Gaussian distribution leads to a good agreement. The Ritchie model with the correction ($\sqrt{3}$ factor) converting projected range to 3D range predicts longer thermalization range than the model without the correction. The corrected one seems similar to the curve of Meesungnoen below 5 eV and predicts longer range because the original paper of Ritchie et al. (1994) only provides the fitting curve up to 5 eV. Finally, the thermalization distance of the corrected Terrissol model is still longer than that of the more realistic Meesungnoen model (with includes a factor of 2 in order to consider phase influence between liquid and amorphous-solid water). The “Meesungnoen amorphous" model shows much longer ranges than the Terrissol model above 4 eV, but it shows opposite behavior below 4 eV.

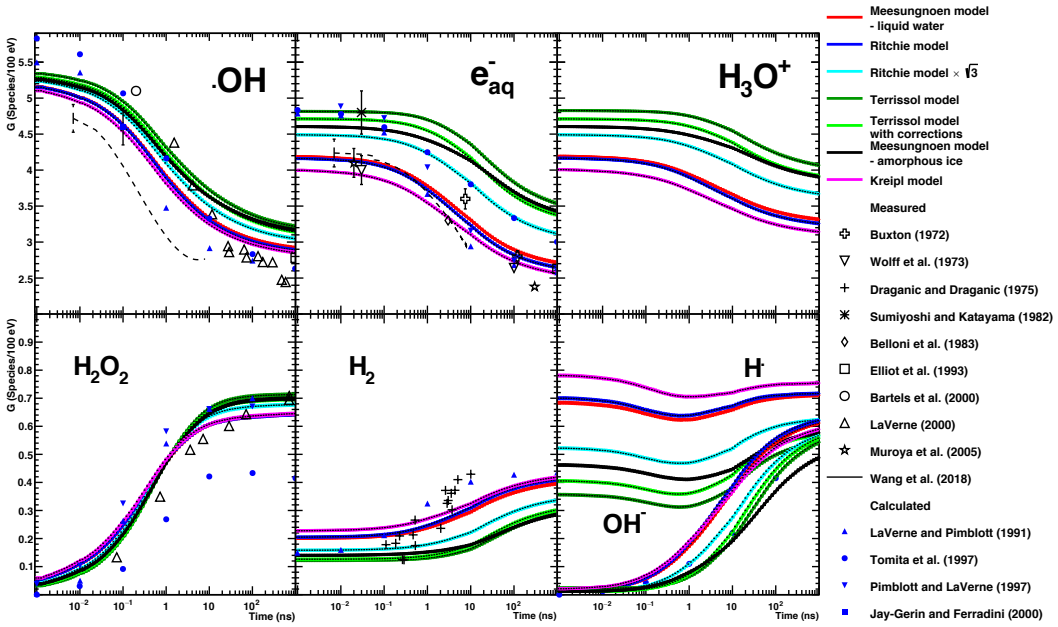


Figure 3.5: Influence of electron thermalization models on G-values versus time.

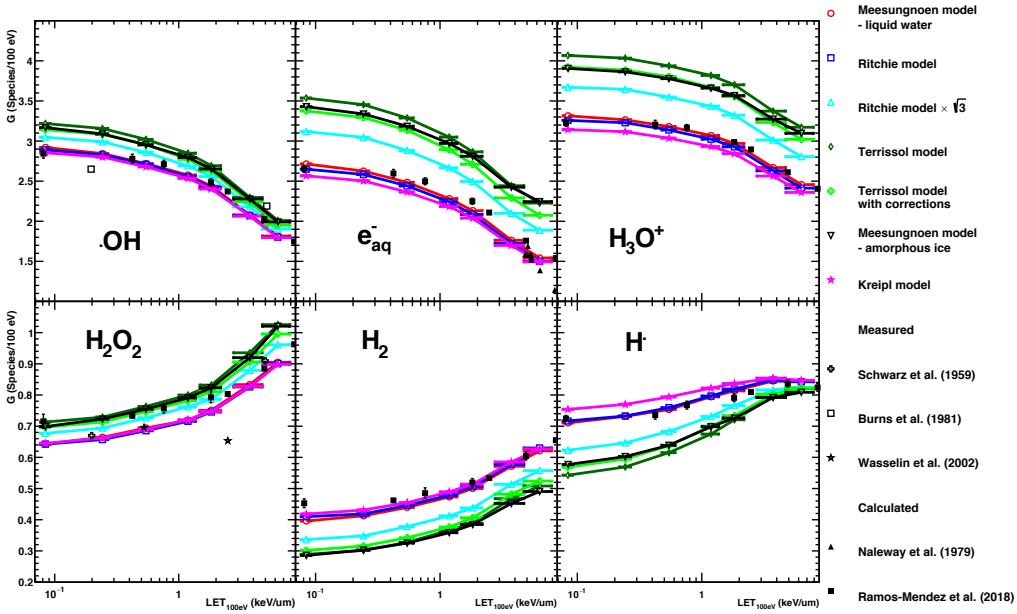


Figure 3.6: Influence of electron thermalization models on G-values versus LET.

Electron thermalization models affect the results on water radiolysis especially for solvated electrons as shown in Figure 3.5 and Figure 3.6. The results obtained with the Meesungnoen model and the Ritchie model without correction show reasonable thermalization range compared to the other experimental data and recent simulation results of Ramos-Mendez et al. (2018).

In the case of Ritchie model, the corrected model corresponds to the original paper (Ritchie et al., 1994). However, the Ritchie model without correction leads to better agreement with experimental data and consistency with the other MTCS codes. The curves obtained with the

uncorrected Ritchie model match with the other data by accident, the thermalization range above 7 eV being not realistic. Thus, we don't highly recommend to use Ritchie model with G4EmDNAPhysics_option2E. The Kreipl model gives mostly similar results to the Meesungnoen model because both these models use same data on mean electron thermalization range. However, in order to be coherent with the original paper of Meesungnoen et al. (2002), we decide to keep the Meesungnoen model as a default model. It should be noted that the uncorrected Terrissol model provides the thermalization range up to 9 eV, and the cut-off value of G4EmDNAPhysics_option2E is 10 eV. In other words, 9-10 eV of electrons don't undergo any elastic scattering. The results of the models for amorphous ice are indeed quite different with those of the Meesungnoen model and the Ritchie model applicable to the liquid phase of water, up to 44.3% especially for e_{aq}^-. The curves of the "Meesungnoen amorphous" model representing the thermalization distance in amorphous ice which is proposed in replacement of the Terrissol model show smaller initial G values than the Terrissol model (already available in Geant4-DNA) by about 4.6% for H_3O^+ and 4.4% for e_{aq}^- for 1 MeV electrons. However, the gaps are getting smaller versus time, for example, 3.9% for H_3O^+ and 3.0% for e_{aq}^- at 1 µs. We can thus conclude that the short thermalization range dominantly affects chemical reactions at early time (Shin et al., 2019).

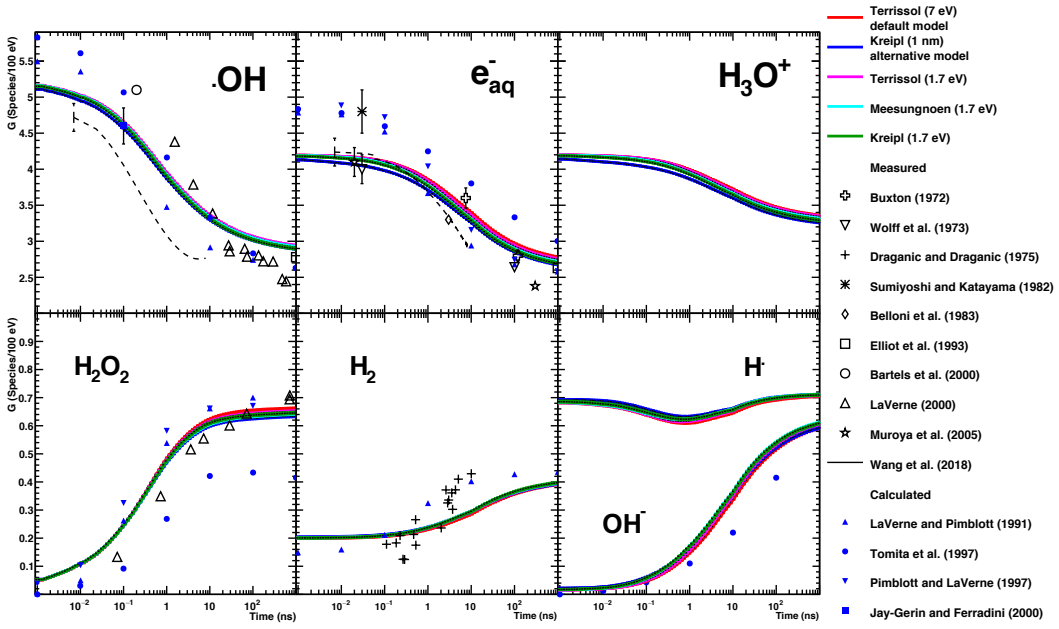


Figure 3.7: Influence of thermalization distance of dissociated electrons by auto-ionization process. The mean thermalization range is: 30.7 nm for the Terrissol model at 7 eV (red curve), 8.2 nm for the Terrissol model at 1.7 eV (magenta curve), and 2.8 nm for the Meesungnoen and Kreipl model at 1.7 eV (cyan and green curves).

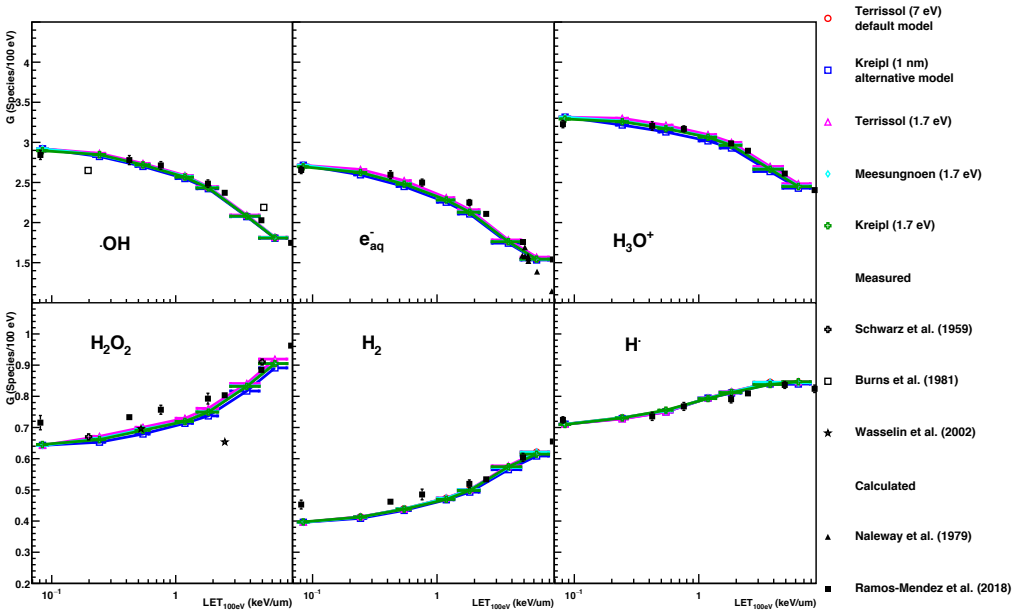


Figure 3.8: Influence of thermalization distance of dissociated electrons by auto-ionization process on G-values versus LET.

In order to evaluate the influence of electron thermalization also on electrons created through the auto-ionization process, we calculated G-values using the default electron thermalization model for sub-excitation electrons (Meesungnoen model) and different models for electrons created by auto-ionization. As shown in Figure 3.7 and 3.8, the influence is not very significant, but small differences are observed, this is expected since the different thermalization models predict slightly different thermalization ranges which influences chemical reactions, and the contribution of sub-excitation electrons is larger than the one of electrons created from auto-ionization.

Generally speaking, it is hard to compare with experimental data due to the lack of influence of these thermalization models on simulation results. However, we could decide to use the energy of 1.7 eV for electrons generated by dissociation process based on the empirical study of Han and Bartels (1990), and the thermalization distance would be coherently calculated using the same thermalization model for both sub-excitation electrons and electrons created by auto-ionization.

3.1.10 Influence of chemistry parameters

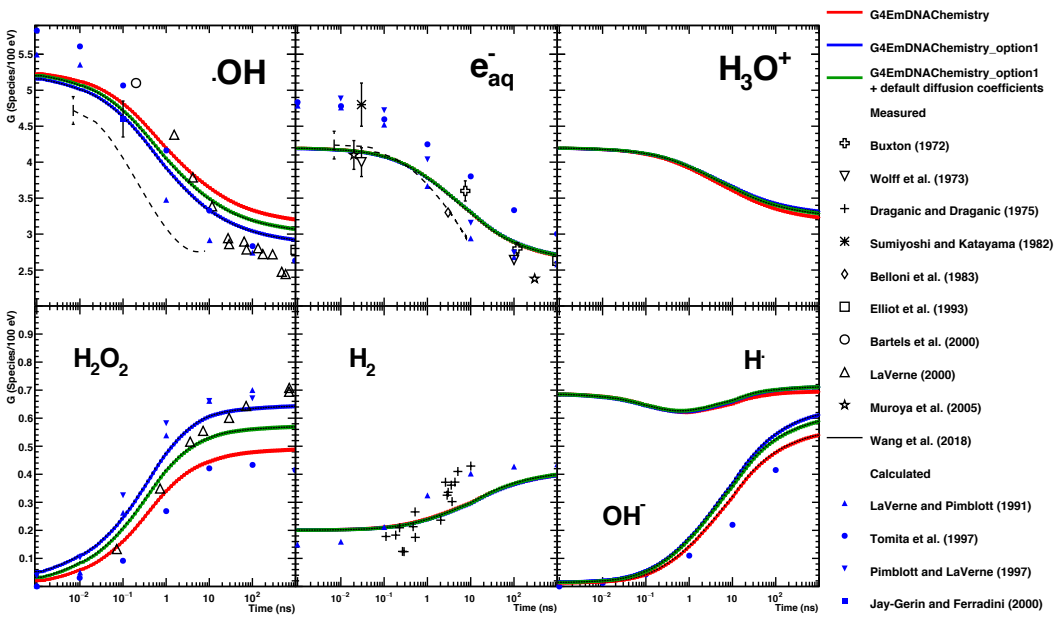


Figure 3.9: Influence of chemistry lists on G-values versus time.

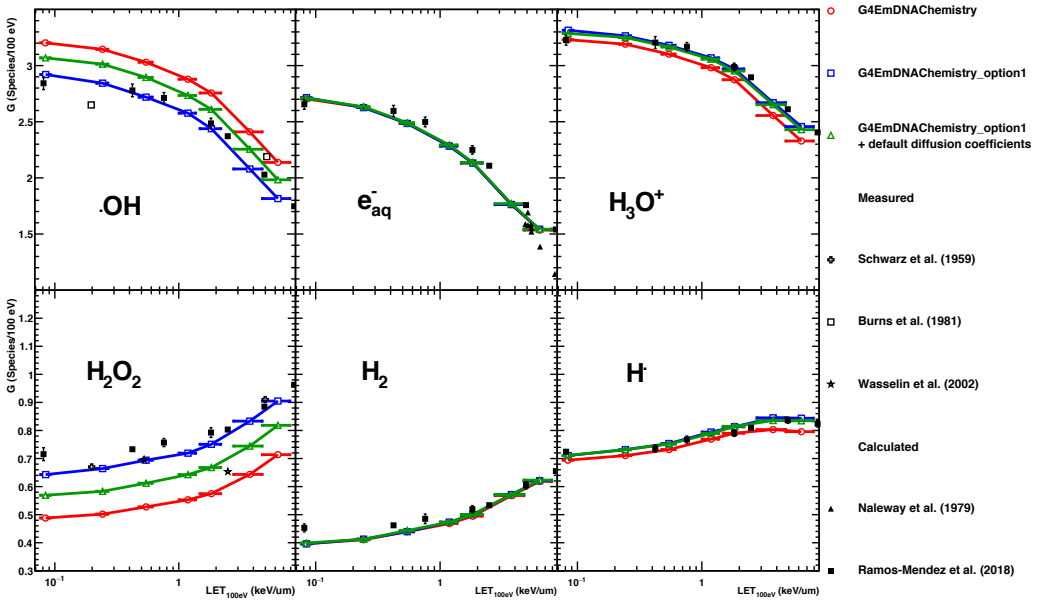


Figure 3.10: Influence of chemistry lists on G-values versus LET.

The influence of *reaction rates* could be evaluated by comparing the G-values of the default chemistry with G-values of chemistry option 1, using default diffusion coefficients (red and green curves in Figure 3.9 and 3.10). The reaction rate of the chemistry option 1 generating •OH radical ($e_{aq}^{-} + H_2O_2 \rightarrow OH^- + {}^{\cdot}OH$) is smaller than for default chemistry option (1.41 for default and

1.10 for option 1), and it is opposite for the reactions involving •OH radical like ($e_{aq}^- + {}^{\cdot}OH \rightarrow OH^-$, ${}^{\cdot}OH + {}^{\cdot}OH \rightarrow H_2O_2$, ${}^{\cdot}OH + H^{\cdot} \rightarrow H_2O$, $e_{aq}^- + H_2O_2 \rightarrow OH^- + {}^{\cdot}OH$). Regarding the influence of *diffusion coefficients*, by comparing the G-values of chemistry option 1 with those of chemistry option 1 + default diffusion coefficients (blue and green curves in Figure 3.9 and 3.10), the results show that •OH radicals and H_2O_2 species which have significantly different diffusion coefficients between default chemistry option and chemistry option 1, are affected by as much as the influence of the reaction rate.

We can thus conclude that not only reaction rates but also diffusion coefficients significantly affect the amount of chemical reactions, and that the chemistry option 1 based on RE-/RITRACKS is more suitable for the current chemistry algorithm implemented in Geant4-DNA in order to reproduce the experimental data of radiochemical yields versus time and LET in water.

3.2. Evaluation of the impact of the pre-chemical processes

In this section, we compared the methodologies employed by MCTS tools such as Geant4-DNA (Bernal et al., 2015), PARTRAC (Kreipl et al., 2009), TRACELE, and TRACPRO (Cobut et al., 1998) (TRACs below) from which RITRACKS has been developed (Plante and Devroye, 2017), TRAX-Chem (Boscolo et al., 2018), and CPA100 (Peudon, 2007) for the modeling of the pre-chemical stage. The models and parameters for each MCTS tool have been collected, and those are re-implemented into Geant4-DNA. The influence of the adjustable settings on the pre-chemical stage is then evaluated based on the initial and final G-values as a function of LET, by comparing with published experimental data and calculated data. It should be noted that the water radiolysis simulations performed in this study is obtained by the independent reaction time (IRT) approach. This approach will be described in Chapter 4.

3.2.1. Physical and physico-chemical processes

Ionization and electronic excitation

In order to evaluate the influence of the pre-chemical stage only, we kept the set of physical models determined in section 3.1 (Shin et al., 2019). Briefly, we considered the G4EmDNAPhysics_option2 constructor, which for electrons uses the Born ionization and excitation model, replacing the default electron elastic scattering model with the newly developed ELSEPA model, calculated using Dirac partial wave analysis (Shin et al., 2018).

The ionization process of water molecule considers five ionization levels ($1b_1$, $3a_1$, $1b_2$, $2a_1$, and oxygen K-shell), but it is reported that the ionization level doesn't affect the initial radiochemical yields (Thomsen et al., 1999). The process ejects an electron from a water molecule, and the H_2O^+ ion is generated. Very rarely less than 1% (Cobut et al., 1998), an Auger electron can also be emitted. In this case, two secondary electrons are generated along with the H_2O^{2+} cation. In the high-LET case, the multi-ionization process affects the number of secondary electrons and the charge of the ionized water molecules (Meesungnoen and Jay-Gerin, 2005); however, the MCTS tool considered in this study does not take this process into account.

The accurate modeling of the excitation process is essential even though the excitation process doesn't generate any secondary particle. The excited water molecule H_2O^* is dissociated into molecular species according to the excitation state. Five excited states (two electronic excitations A^1B_1 and B^1A_1, two Rydberg series Ryd A+B and Ryd C+D, and diffuse bands) are implemented in Geant4-DNA (Dingfelder et al., 1998). Most of MCTS tools classify the excitation levels as A^1B_1 and B^1A_1 states. A^1B_1 state represents the excitation of the 5^{th} electronic layer ($1b_1 \rightarrow 4a_1$), whereas the B^1A_1 state corresponds to the excitation of the 4^{th} layer ($3a_1 \rightarrow 4a_1$), and the other states. TRACs argued that there is no Rydberg band in condensed-phase water

(Cobut et al., 1998). Instead of that, they use plasmon decay, which represents the collective excitation of free electrons in the material, usually metals.

Even though the public version of Geant4-DNA currently takes into account Auger electron emission during the physical stage, a H_2O^+ cation (instead of H_2O^{2+} cation) and two electrons (recoil and Auger electrons) are generated by the interaction. In this study, the H_2O^{2+} cation induced by Auger effects and its chemistry kinetics have been implemented into Geant4-DNA. Then, the influence of this correction for Auger effect on water radiolysis simulation is evaluated, even if the contribution of Auger effect is expected to be relatively small.

Electron attachment

A low energy electron, below 15 eV, can be captured by a neighbor water molecule H_2O as:

$$H_2O + e^- \rightarrow H_2O^- \tag{3.30}$$

Geant4-DNA (Francis et al., 2011) and CPA100 (Martin, 2003) are currently using the electron attachment cross-section proposed by Melton for vapour phase water (Melton, 1972). This paper describes the energy dependence of the electron attachment process, however it is reported that the cross-section of Melton is overestimated compared not only to the liquid-phase of water, but also to the vapour-phase (Rawat et al., 2007). Rowntree et al. (1991) presented the electron attachment cross-section in amorphous ice, unfortunately, this cross-section was provided in arbitrary unit. In order to implement the Rowntree model in MC method, TRACs scaled the cross-section based on the agreement between the calculated initial H_2 yield and that supposed by Schwarz (1969), with an approximate initial G-value of 0.15 molecules / 100 eV (Cobut et al., 1996).

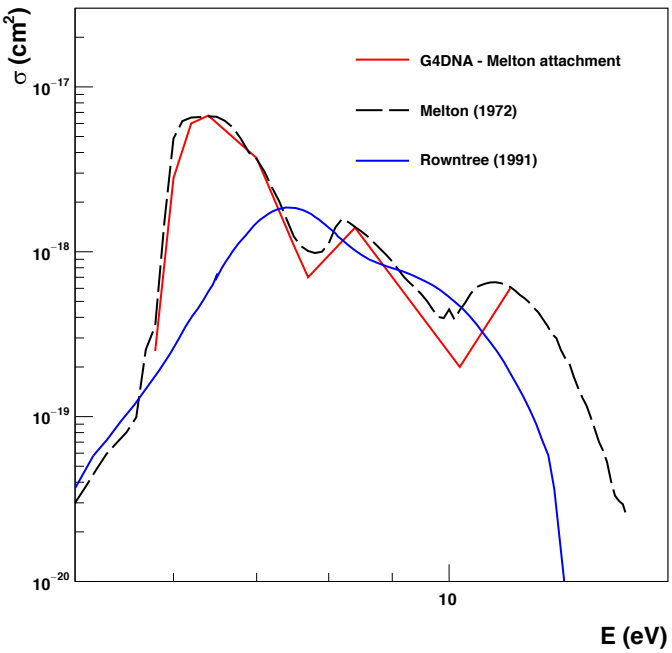


Figure 3.11: Cross-section data for electron attachment process as a function of electron energy. The cross-section of Rowntree (blue line) is scaled in the paper of Cobut et al. (1998).

As shown in Figure 3.11, the two cross-sections for vapour (plain red and dashed black lines) and amorphous ice (blue line) show definitely different peaks, scale, and dissociation channels (further described in section 3.2.2). Unfortunately, predicting the energy dependence of the electron attachment cross-section for liquid water is still challenging due to the lack of reliable experimental data in the liquid phase. Note that the public version of Geant4-DNA (up to Geant4 version 10.6) presents a limitation. It fully ignores the electron attachment process below the tracking cut energy of electrons (7.4 eV or 10 eV), because sub-excitation electrons directly undergo the one step thermalization process. This is shown in Figure 3.12, where the number of occurrences of electron attachment process depends on the tracking cut (while it should not).

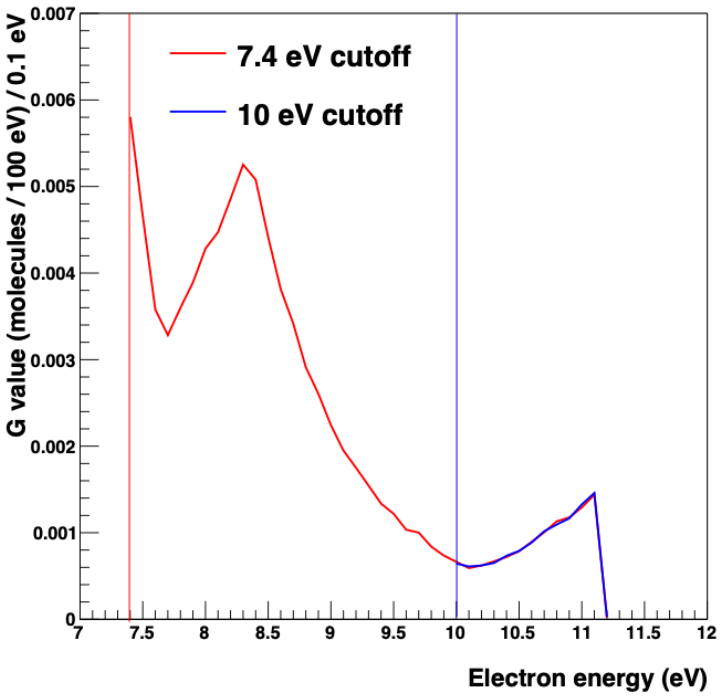


Figure 3.12: Energy spectra of electron attachment process according to the electron cut-off energy (7.4 and 10 eV) in the public version of Geant4-DNA. The vertical axis shows the number of occurrences of the attachment process per 100 eV of energy deposited in the liquid water medium, per energy bin.

In this work, for a more realistic simulation, we thus propose to take into account the electron attachment process below the tracking cut. Technically, the attachment process is sampled in competition with the one step thermalization process based on the corresponding cross-sections. In addition, we propose to determine the position of H_2O^- anions using the mean thermalization distance of electrons when the process is sampled by the one step thermalization method of section 3.1.3:

$$\overline{R_{H_2O^-}(k_{initial}, k_{attachment})} = \overline{R_{e^-_{subexc}}(k_{initial})} - \overline{R_{e^-_{subexc}}(k_{attachment})} \quad (3.31)$$

where $\overline{R(k)}$ is the mean thermalization distance of electrons which have a kinetic energy k. The k_{initial} is the initial kinetic energy of the sub-excitation electrons needed by the one step thermalization algorithm, and $k_{\text{attachment}}$ is the energy deposition induced by the attachment process when it occurs (it is equal to the remaining kinetic energy of the electron), respectively.

We evaluate in section 3.2.5 the influence of this new assumption on the initial G-values, and the optimal cross-section and scaling factor are determined.

Electron-hole recombination

The ionized water molecules H_2O^+ might be dissociated into two reactants of H_3O^+ and $^{\bullet}OH$, however, the water cation can capture a neighbor solvated electron e_{aq}^- by the process named "electron-hole recombination" at femtosecond time scale:

$$H_2O^+ + e_{aq}^- \rightarrow H_2O^* \tag{3.32}$$

This process can frequently happen when the spatial distribution of molecular species (H_2O^+ and e_{aq}^- particularly) becomes denser, and a sparse distribution leads to small numbers of recombinations. Thus, electron-hole recombination is very important for understanding the LET dependency of the initial radiochemical yields in the pre-chemical stage. The recombination process is only mentioned in the papers of CPA100 (Martin, 2003) and TRACs (Cobut et al., 1998), not in the papers of PARTRAC and TRAX-Chem. Geant4-DNA and CPA100 (Martin, 2003) provide this process using an empirical survival fraction given by Lu et al. (1989):

$$P(r) = 1 - exp(-r_c/r) \tag{3.33}$$

where r_c is the Onsager's radius, and r is the separation distance between the H_2O^+ cation and the solvated electron e_{aq}^-.

On the other hand, TRACs (Cobut et al., 1998) mention that the probability of the recombination and the energy dependence are obtained from Monte Carlo simulations. Thus, we can assume that the approach of TRACs follows the rule of the diffusion-controlled reactions just like the chemical stage. We thus supposed a "contact reaction" (see next paragraph) with a reaction radius of 0.775 nm, which is the summation of the water molecule radius (0.275 nm) and the reaction radius of solvated electron (0.5 nm). This value is used for our validation study even though the value apparently underestimates the reaction probability of the recombination process due to not using an effective reaction radius caused by the Coulomb force between H_2O^+ and e_{aq}^-.

In this work, the electron-hole recombination processes based on the empirical survival fraction of equation (3.33) and that based on diffusion-controlled reaction will be compared with each other in order to match at best the measured G-values.

Immediate reactions

TRACs define "contact reaction" if the distance r is smaller than the reaction radius σ because the Green function diffusion equation (GFDE) can't describe it (Frongillo et al., 1998). The contact reaction immediately occurs at the position with the following probability of contact reaction, $P_{contact}$:

$$P_{contact} = \frac{\exp(-r_c/\sigma) - \exp(-r_c/(\sigma + R_s))}{\exp(-r_c/\sigma) - \exp(-r_c/(\sigma + R_s)) - (k_{diff}/k_{act})(1 - \exp(-r_c/\sigma))} \tag{3.34}$$

where σ is the effective reaction radius, k_{diff} and k_{act} are the diffusion-controlled and activation reaction rate constants, respectively. R_S is the limit of the separation distance after unsuccessful

encounter which possibly leads to another re-encounter. This approximate distance between water molecules in TRACs and Geant4-DNA is taken as 0.3 nm (Goulet and Jay-Gerin, 1992).

3.2.2. Dissociation channels

The water molecules, excited or ionized, are dissociated into molecular species. Several dissociation channels and associated probabilities have been proposed based on empirical studies. Table 3.10 shows the dissociation channels and the probabilities for each channel provided by different MCTS tools.

Table 3.10: Dissociation scheme and associated probability for each channel currently available in the MCTS tools.

		Channel	This work	Geant4-DNA (Karamitros et al., 2011)	TRACs (Cobut et al., 1998)	TRAX-Chem (Boscolo et al., 2018)	PARTRAC (Kreipl et al., 2009)	CPA100 (Martin, 2003)
Ionization	H_2O^+	$H_3O^+ + {}^{\bullet}OH$	100	100	100	100	100	100
	Auger effect H_2O^{2+}	$2H_3O^+ + H_2O_2$	100	-	100	-	-	100[a]
Excitation	A^1B_1	$H^{\bullet} + {}^{\bullet}OH$	65	65	65	75	65	50
		H_2O	35	35	35	25	35	50
	B^1A_1	$H_3O^+ + {}^{\bullet}OH + e_{aq}^-$	50	55	50	55	55	50
		$H^{\bullet} + {}^{\bullet}OH$	25.35	-	25.35	-	-	-
		$H_2 + O(^1D)$[b]	3.25 ($H_2 + 2{}^{\bullet}OH$[d])	15 ($H_2 + 2{}^{\bullet}OH$[d])	3.25 ($H_2 + H_2O_2$[e])	30 ($H_2 + H_2O_2$[e])	15 ($H_2 + 2{}^{\bullet}OH$[d])	20 ($H_2 + H_2O_2$[e])
		$2H^{\bullet} + O(^3P)$[c]	3.9	-	3.9	-	-	-
		H_2O	17.5	30	17.5	15	30	30
	Rydberg, Diffuse bands	$H_3O^+ + {}^{\bullet}OH + e_{aq}^-$	50	50	-	57 or 100[f]	50	100
		$H + {}^{\bullet}OH$	-	-	-	20 or 0[f]	-	-
		H_2O	50	50	-	23 or 0[f]	50	-
Electron capture	Attachment	$OH^- + {}^{\bullet}OH + H_2$	100	100	100	-	-	100
	Recombination	$H^{\bullet} + {}^{\bullet}OH$	35.75	55	35.75	-	-	100
		$H_2 + O(^1D)$	13.65 ($H_2 + 2{}^{\bullet}OH$[d])	15 ($H_2 + 2{}^{\bullet}OH$[d])	13.65 ($H_2 + H_2O_2$[e])	-	-	-
		$2H^{\bullet} + O(^3P)$	15.6	-	15.6	-	-	-
		H_2O	35[g]	30	35[g]	-	-	-

[a] It is basically $H_3O^+ + H_2O_2$ in the original CPA100 paper (Peudon, 2007), but $2H_3O^+ + H_2O_2$ is used for a balanced chemical equation.
[b] Oxygen atom in the singlet D state.
[c] Oxygen atom in the triplet P state.
[d] $O(^1D) + H_2O \rightarrow 2{}^{\bullet}OH$ from Burns and Marsh (1981).
[e] $O(^1D) + H_2O \rightarrow H_2O_2$ from Taube (1956).
[f] The excited water molecule is completely (100%) auto-ionized in the case of diffuse bands. This data for Rydberg excitations is from Nikjoo et al. (2006).
[g] Same relaxing probability for B^1A_1 excitation except auto-ionization.

The dissociation scheme for ionized water molecules is clearly reported (Thomsen et al., 1999), that is $^{\bullet}OH$ and H_3O^+, regardless of the ionization level, and all MCTS tools use the same dissociation scheme. Besides, the dissociation channel for H_2O^{2+} generated by Auger effect has been implemented into Geant4-DNA in this work.

In the case of excited water molecules H_2O^*, however, the dissociation channels and the corresponding probabilities for each excited level are usually approximated due to the lack of literature data, especially for B^1A_1 excitation. Beaudre (1988) proposed individual dissociation scheme based on Buxton (1982). Nikjoo et al. (1997), Nikjoo et al. (2006), Peudon (2007), Boscolo et al. (2018) follow the scheme of Beaudre (1988) with their own adjustment in order to match the measured G-values. On the other hand, the methodology determining the dissociation scheme of TRACs is well-described in the paper of Cobut et al. (1998). The dissociation channels of PARTRAC and Geant4-DNA are from the work of Ballarini et al. (2000), which is based on the work of TRACs (Cobut et al., 1998) with small modification. However, the adjustment of B^1A_1 channel of PARTRAC is for the reproduction of the initial experimental yields, without considering the attachment and recombination processes which are ignored by PARTRAC but not by Geant4-DNA. Thus, in this study, we tried to apply the original B^1A_1 dissociation channel proposed by TRACs (Cobut et al., 1998), and we compared the G-values with the initial experimental yields.

For the dissociation of H_2O^- anions generated by electron attachment processes, three dissociation schemes are described in the paper of Melton (1972) as:

$$H_2O^- \rightarrow O^- + H_2 \tag{3.35}$$

$$H_2O^- \rightarrow H^- + {}^{\bullet}OH \tag{3.36}$$

$$H_2O^- \rightarrow H^{\bullet} + OH^- \tag{3.37}$$

The cross-sections of equations (3.35) and (3.36) are dominant according to the paper, and the generated H^- and O^- anions in those schemes instantly react with nearby water molecules at femtosecond scale:

$$O^- + H_2O \rightarrow OH^- + {}^{\bullet}OH \tag{3.38}$$

$$H^- + H_2O \rightarrow OH^- + H_2 \tag{3.39}$$

In both cases of equations (3.35) and (3.36), we thus can conclude that most of H_2O^- anions will be dissociated as:

$$H_2O^- \rightarrow {}^{\bullet}OH + OH^- + H_2 \tag{3.40}$$

Thus, the unified dissociation given by equation (3.40) is used in Geant4-DNA, TRACs, and CPA100 MCTS tools for the electron attachment process.

As shown in the last row of Table 3.10, the dissociation of the excited water molecule induced by electron-hole recombination is considered in CPA100, TRACs, and Geant4-DNA. A simple dissociation channel of A^1B_1 excitation without considering water relaxation is used in

CPA100. However, an empirical dissociation channel has been proposed by Rowe et al. (1988). It is reported that the probability of •OH + H• dissociation channel is 55% without considering relaxation in the paper. However, the public version of Geant4-DNA is using the value including the probability of relaxation. We thus implemented the recombination dissociation probabilities of TRACs into Geant4-DNA in order to keep the consistency with the original paper.

3.2.3. Displacement of hot fragments

Even though the displacement of the daughter molecules doesn't influence the initial G-values of molecular species, such displacement can impact the encounters and the reactions of reactants at early time.

PARTRAC (and Geant4-DNA, following the approach of PARTRAC) calculates the displacement based on the empirically observed RMS distances and momentum conservation (Kreipl et al., 2009), and TRACs and CPA100 also provide several RMS distances in their papers (Cobut et al., 1998, Beaudre, 1988). TRAX-Chem calculates the mean-free-path based on the Einstein-Smoluchowski equation:

$$\lambda_\tau = \sqrt{4\frac{DE_k}{k_B T}\Delta t} \tag{3.41}$$

where Δt is 1 ps which is the start time of the chemical stage, E_k is the empirical translational energy transfer measured by the laser-fluorescence technique, k_B is the Boltzmann constant, T is the temperature, and D is the diffusion coefficient.

Table 3.11: RMS distances of the mother and the daughter molecules generated by the ionization process (and the auto-ionization channel of B^1A_1, Rydberg, and diffuse bands excitation process).

	Geant4-DNA	**TRACs**	**TRAX-Chem**	**CPA100**
Migration (H_2O^+)	2.0 nm	1.375 nm	1.5 nm	-
H_3O^+	0.8 or 0 nm	0.3 nm	1.96 nm	0.56 nm
•OH	0 or 0.8 nm	0 nm	-1.16 nm	0.81 nm
e_{aq}^-	1.7 eV	1.3 eV	1.7 eV	7 eV

Table 3.11 shows the RMS distances of the ionized water molecule and the daughter molecules according to the different MCTS codes. In the case of the RMS distance of the solvated electron, Geant4-DNA (Shin et al., 2019) and TRAX-Chem (Boscolo et al., 2018) are using the electron thermalization distance corresponding to a kinetic energy of 1.7 eV based on empirical data (Han and Bartels, 1990). On the other hand, TRACs mentioned that the B^1A_1 state lies approximately 1.3 eV above the first ionization potential (Cobut et al., 1998), and CPA100 argued that at least 10 eV are required to eject the dissociation electron (Beaudre, 1988).

Table 3.12: RMS distances of the daughter molecules generated by the Auger effect.

	TRACs	CPA100
Migration (H_2O^{2+})	-	-
H_3O^+	1.2 nm	0 nm
H_3O^+	1.2 nm[a] 0.3 nm	0.29 nm
H_2O_2	1.2 nm[a] 0 nm	0.56 nm

[a] RMS distance for the molecule which is in an intermediate state.

Table 3.12 shows the RMS distances of H_2O^{2+} cation induced by Auger effect taken in TRACs and CPA100. There is an intermediate state like:

$$H_2O^{2+} + 3H_2O \rightarrow H_3O^+ + (OH^+ + 2H_2O) \rightarrow H_3O^+ + (H_3O^+ + H_2O_2) \quad (3.42)$$

Thus, the H_3O^+ cation firstly dissociated at the intermediate state follows different RMS distances from the second one.

Table 3.13: RMS distances of the daughter molecules generated by the A^1B_1 dissociation (and same dissociation channel in B^1A_1 dissociation).

	Geant4-DNA	TRACs	TRAX-Chem	CPA100
Migration (H_2O^*)	-	-	-	-
$^{\bullet}OH$	0.13 nm	0 nm	0.23 nm	0.56 nm
$H^{\bullet}$	2.27 nm	0.41 nm	-1.51 nm	0.29 nm

Table 3.13 shows the RMS distances of A^1B_1 excited water molecule and the daughter molecules in the different MCTS codes. According to the rule of momentum conversion, the lighter molecule $H^{\bullet}$ should move far, and the heavy $^{\bullet}OH$ may be positioned closer, except for the RMS distances of CPA100.

Table 3.14: RMS distances of the daughter molecules generated by the B^1A_1 dissociation.

	Geant4-DNA	TRACs	TRAX-Chem	CPA100
Migration (H_2O^*)	-	-	-	-
H_2	-0.71 nm[a]	0.36 nm	1.23 nm	0.56 nm
$2^{\bullet}OH$	0.09 nm[a,b] ± 0.55 nm	-	-	-
H_2O_2	-	0 nm	0 nm	0.29 nm

[a] The RMS distances for intermediate state are opposite in the paper of Kreipl et al. (2009), Bernal et al. (2015).
[b] RMS distance for the molecule which is in intermediate state.

Table 3.14 shows the RMS distances of the dissociation scheme:

$$H_2O^* + H_2O \rightarrow H_2 + O(^1D) + H_2O \rightarrow H_2 + 2^{\cdot}OH (\text{or } H_2O_2) \quad (3.43)$$

The intermediate state for $O(^1D)$ is also considered in Geant4-DNA. It is reported that $O(^1D)$ is immediately dissociated into H_2O_2 as proposed in the paper of Taube (1956). This dissociation channel is taken into account in TRACs, TRAX-Chem, and CPA100. However, PARTRAC and Geant4-DNA use more recent data proposed by Burns and Marsh (1981), that is, two •OH radicals. In the paper of PARTRAC (Kreipl et al., 2009) and Geant4-DNA (Bernal et al., 2015) (including the public version of Geant4-DNA), reversed RMS distances at the intermediate state for the H_2 molecule and for $O(^1D)$ are applied, as -0.09 and 0.71 nm, respectively. We instead use the RMS distances based on the momentum conservation in this work (that is, heavier molecular species travel less distance than light ones).

Table 3.15: RMS distances of the daughter molecules generated by B^1A_1 dissociation.

	TRACs
Migration (H_2O^*)	-
$2H^•$	± 0.8 nm
$O(^3P)$	0 nm

Table 3.15 shows the RMS distances of the simple B^1A_1 dissociation channel in TRACs. The heavy $O(^3P)$ molecule is positioned at the site of H_2O^*, and two hydrogen radicals move to the opposite directions.

Table 3.16: RMS distances of daughter molecules generated by electron attachment.

	Geant4-DNA	**TRACs**	**CPA100**
Migration (H_2O^-)	-		-
H_2	-0.71 nm[a]	0.36 nm[b] 0 nm	0.56 nm
•OH	0.09 nm[a,b] + 0.55 nm	0 nm	0.81 nm
OH^-	0.09 nm[a,b] - 0.55 nm	0.36 nm[b] 0.30 nm	0 nm

[a] The RMS distances for intermediate state are reversed in the public version of Geant4-DNA.
[b] RMS distance for the molecule which is in intermediate state

Table 3.16 shows the RMS distances of the electron attachment process in Geant4-DNA, TRACs, and CPA100. Geant4-DNA currently uses similar displacement approach with B^1A_1 dissociation based on equation (3.35) (Buxton et al., 1988). However, TRACs and CPA100 are using equation (3.36) (Curtis and Walker, 1992) because the contribution is dominant according to Melton cross-section (Melton, 1972), and there is no empirical evidence of the other dissociation channel in condensed-phase water (Cobut et al., 1996). Thus, in this study, we implemented in Geant4-DNA the displacement approach of TRACs.

3.2.4. Validation study by comparing simulation data with literature data

In summary, this study includes the dissociation of H_2O^{2+} generated by Auger effect, the fix on electron attachment process (which was ignored during thermalization process), the modifications of dissociation channel for B^1A_1 excitation and recombination process, and the modifications of displacement for the B^1A_1 excitation channel ($2^{\bullet}OH+H_2$) and the attachment channel ($^{\bullet}OH+OH^-+H_2$).

In order to validate the simulation of the pre-chemical stage, we first calculated the number of water molecules excited and ionized by the irradiation. For electron simulations, the conditions proposed by the previous study such as 1 x 1 x 1 km^3 water phantom, initial energies and corresponding energy thresholds are used (Shin et al., 2019). Proton and alpha particles are generated at the edge of a 5 x 5 x 5 μm^3 water phantom, as similar with typical water radiolysis studies without the primary killer technique.

At first, the initial G-values as a function of LET are evaluated by comparing with those found in the literature. This allows us to identify the optimal models for the simulation of the pre-chemical stage. We finally evaluate the G-values as a function of time and LET.

Table 3.17: Experimental expectations of G-values (molecules / 100 eV) for initial radiolytic yields (at about 1 ps).

Reference	$^{\bullet}OH$	e_{aq}^-	H_3O^+	H_2	$H^{\bullet}$	Experimental condition
Wolff et al. (1973)	-	4.0 ± 0.2	-	-	-	41 MeV e- @ 30 ps
Sumiyoshi and Katayama (1982)	-	4.8 ± 0.3	-	-	-	45 MeV e- @ 30 ps
LaVerne (2000)	5.6 ± 0.3	-	-	-	-	^{60}Co extrapolated
Bartels et al. (2000)	-	4.0 ± 0.2	-	-	-	20 MeV e- extrapolated
Muroya et al. (2005)	-	4.1 ± 0.2	-	-	-	20 MeV e- @ 20 ps
Wang et al. (2018)	4.72	4.40	-	-	-	7.5 MeV e- @ 7 ps

[a] It is mentioned that the G-values of $^{\bullet}OH$ radicals and H_3O^+ are similar with those of e_{aq}^-.

In the case of initial G-values, it is impossible to measure the molecular concentration at 1 ps. Most of the experimental data from very early time used to obtain G-values at 1 ps are shown in Table 3.17.

Table 3.18: Calculated G-values (molecules / 100 eV) for initial radiolytic yields (at about 1 ps).

Reference	$^{\bullet}OH$	e_{aq}^-	H_3O^+	H_2O_2	H_2	$H^{\bullet}$	OH^-	Condition
Bisby et al. (1977)	5.70	4.78	4.78	0	0.15	0.62	0	Analytical calculation
Trumbore et al. (1978)	6.0	4.7	4.7	0	0.25	0.8	-	Analysis from references (Jonah et al., 1976, Jonah and Miller, 1977, Schwarz, 1969)
Naleway et al. (1979)	5.40	4.78	4.78	0	-	0.62	0	1.6 keV e- Analytical calculation
Turner et al. (1983)	8.4	6.3	6.3	-	0.3	2.1	-	5 keV e- MCTS
Green et al. (1990)	6.53	3.89	3.95	-	0.02	4.19	-	22 MeV e- MCTS
Kaplan et al. (1990)	6.82	6.57	4.8	-	0.62	0.84	-	10 keV e- MCTS
LaVerne and Pimblott (1991)	5.50	4.78	4.78	-	0.15	0.42	-	Fitting data Laplace transform technique
Pimblott and LaVerne (1997)	5.37	4.93	4.93	0	0.16	0.45	0	1 MeV e- initial yields
	5.57	4.88	4.90	0.04	0.16	0.44	0.02	1 MeV e- including 1 ps reactions
	5.41	4.97	4.97	0	0.15	0.45	0	10 keV section of 1 MeV e- initial yields
	5.60	4.92	4.94	0.04	0.16	0.44	0.02	10 keV section of 1 MeV e- including 1 ps reactions MCTS
Tomita et al. (1997)	5.89	4.88	-	-	-	0.96	-	1 MeV e- MCTS (TRACELE and RADYIE)
Cobut et al. (1998)	6.05	5.30	5.38	0.06	0.13	0.72	0.08	150 keV e- MCTS (TRACELE) without 1 ps reactions
Jay-Gerin and Ferradini (2000)	4.6	-	-	-	-	-	-	Inverse Laplace transform analyses MC Fitted data @ 100 ps
Kreipl et al. (2009)	5.78	4.83	4.83	-	0.16	0.63	-	Averaged value of ^{60}Co 11 electron energy values (30-750 keV) 21 protons (0.5-300 MeV)

								11 alphas (2-300 MeV) 10 carbons (10-1,000 MeV) MCTS (PARTRAC)
Plante and Devroye (2017)	5.10	4.18	-	0.15	0.34	0.69	-	300 MeV protons MCTS (RITRACKS)
Boscolo et al. (2018)	5.635	4.688	-	0.108	0	-	-	1 MeV e- 3 MeV proton MCTS (TRAX-Chem)
	5.697	4.774	-	0	0	-	-	

Due to the lack of experimental data, the G-values calculated by the other MCTS tools and by analytical methods are collected as shown in Table 3.18. It is clear that the size of the water phantom influences the time evolution of G-values, however, those collected G-values are directly compared with the G-values calculated in this study in order to verify the plausibility of our simulations.

In addition, in this study, the time evolution of G-values (up to 1 us) and G-values as a function of LET are assessed with and without the proposed improvements. The experimental and calculation data used in the section 3.1 are compared with the simulated G-values versus time and LET. For the high LET cases, Anderson and Hart (1961), Pastina and LaVerne (1999) are additionally compared with the calculated results.

3.2.5. Results for ionization and excitation

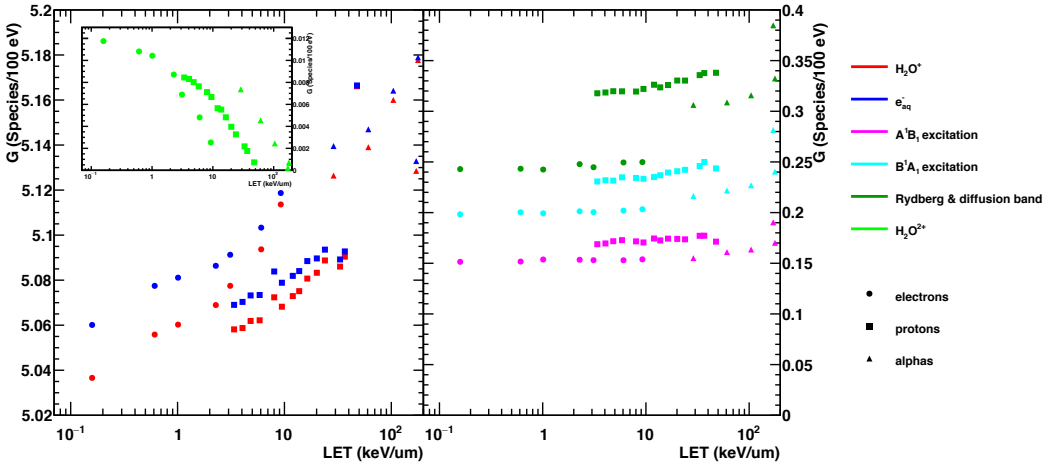


Figure 3.13: The number of ionized and excited water molecules, and solvated electrons as a function of LET. The number of H_2O^{2+} are shown in the inset. Results are presented for 100 eV of energy deposited in the medium.

Figure 3.13 shows the ionization, excitation, and solvated electron yields as a function of LET. It has to be noted that the electron attachment and the electron-hole recombination processes are not considered in these results, thus the number of H_2O^+ and e_{aq}^- would be reduced by those processes.

The number of solvated electrons is always bigger than the number of H_2O^+ cations, as twice as the number of Auger emissions (leading to H_2O^{2+} cations). This is because ionization with Auger emission generates in total two electrons. The contribution of Auger effect decreases as a function of LET.

The observed discontinuities up to 30% between electron-proton-alpha cases, especially for protons, are due to the discontinuities of the physics models. For example, the last point of proton (at 500 keV at 47.66 keV/µm) shows a disagreement because the protons below 500 keV undergo

excitation described by the Miller-Green model and ionization described by the Rudd ionization model, instead of the Born excitation and ionization models above 500 keV. In addition, the last point of alpha (1 MeV, at 171.25 keV/µm) is apparently influenced by the phantom size (5x5x5 μm^3) because the CSDA range of 1 MeV alpha particle is less than 6 µm according to the NIST-ASTAR data (Berger et al., 1998). Thus, the selection of track segment is apparently biased for 1 MeV alpha particles.

In view of the limited magnitude of the observed discrepancies on G dependency versus LET between different particle types, we conclude that the physics models of Geant4-DNA are acceptable to simulate water radiolysis. We underline that in the future, the inelastic scattering model of Emfietzoglou et al. (2012) (currently being extended from 10 keV to 10 MeV) used with our newly developed ELSEPA electron elastic scattering model (Shin et al., 2018) will probably become the default combination of Geant4-DNA physics models for electrons. This new inelastic scattering model may enable to calculate more reliable dependency of the ionizations and the excitations versus LET.

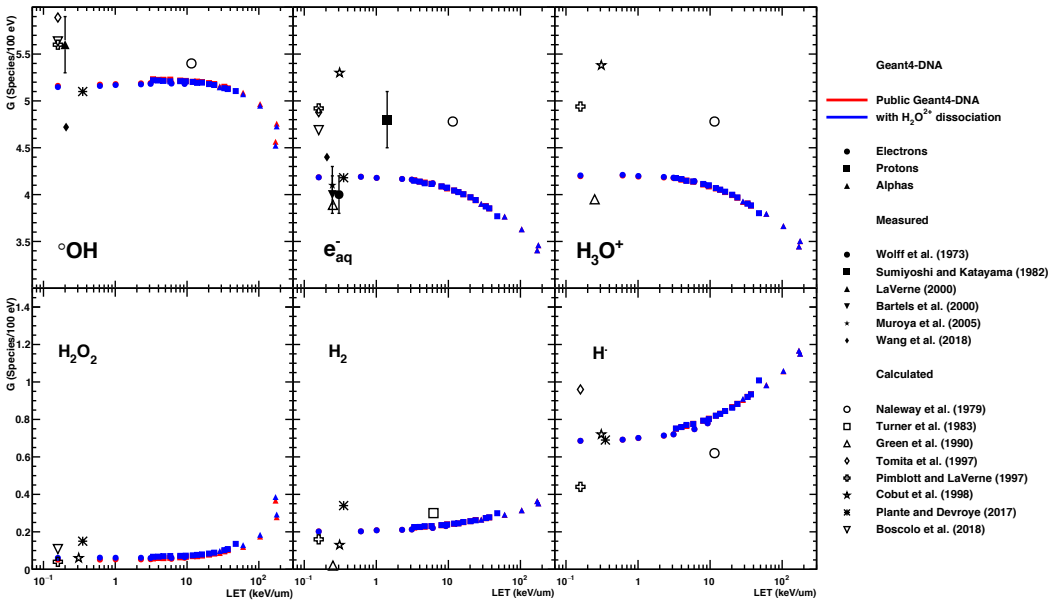


Figure 3.14: Initial G-values (at 1 ps) as a function of LET with and without the dissociation channel of H_2O^{2+} generated by Auger effect.

The influence of Auger effect is not observable on initial G-values due to its small contribution in the physical stage, except for a small increase of G-values of H_2O_2 up to 0.03 molecules/100 eV as shown in Figure 3.14. However, we don't recommend turning off the Auger effect in physical stage because the Auger electrons affect the other pre-chemical processes, such as the number of solvated electrons, electron attachment and recombination processes.

3.2.6. Results for electron attachment process

We propose here to verify the validity of the sampling method described in section 3.2.1 allowing to fix the non-realistic dependence of the electron attachment process with tracking cut. For validation of this methodology, the energy spectra of the new attachment process are compared with those without using the one step thermalization approach (that is, using instead the step-by-step approach). Even though the step-by-step approach of Geant4-DNA below cut-off energy consists only of vibrational excitation and attachment processes without any elastic scattering, the number of attachment processes and the delivered energy can be evaluated.

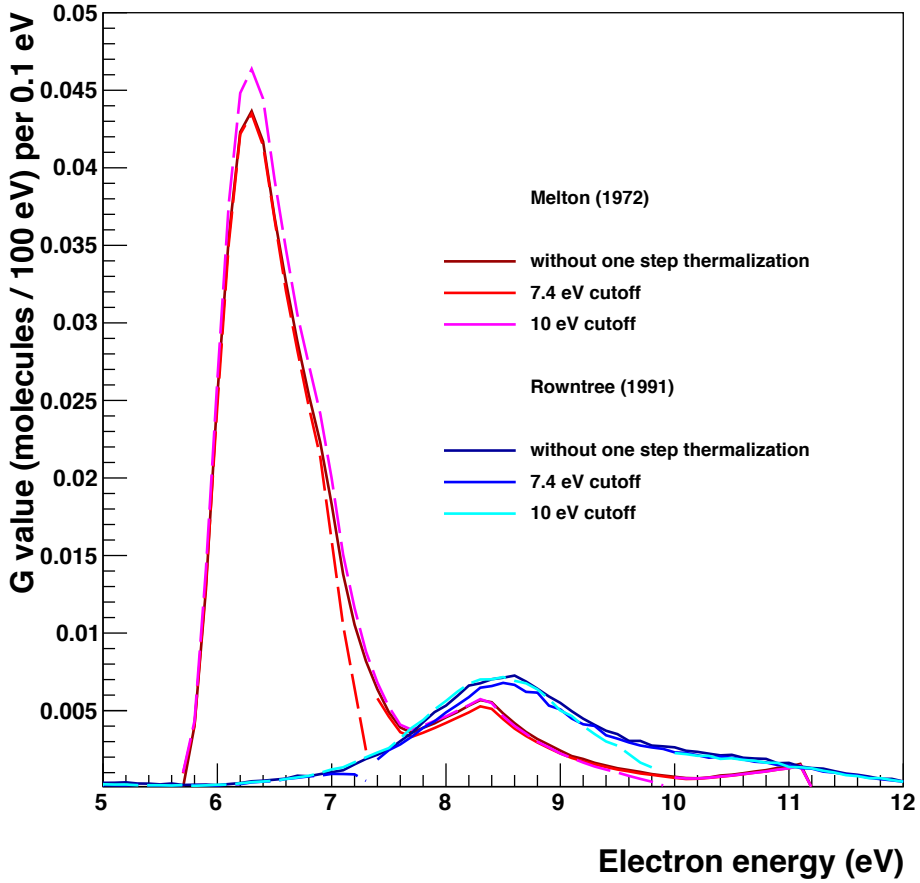


Figure 3.15: Energy spectra of electrons undergoing electron attachment. The solid and dashed lines indicate the energy spectra sampled using step-by-step approach and the one step thermalization process, respectively.

This new sampling algorithm makes the one step thermalization process slower by a factor of 1.1, but this is almost negligible compared to the whole simulation time of the chemical stage. Even though a small discontinuity at the cut-off energy (7.4 or 10 eV) is observed due to the biased sampling during the one step thermalization process, the new sampling algorithm shows a good agreement compared to the energy spectra of the step-by-step approach ("without one step

thermalization" curves) shown in Figure 3.15, for both Melton and Rowntree models. Then, the total G-values for H_2O^- anion with Melton cross-section are 0.500, 0.474, and 0.525 molecules/100 eV for 0, 7.4, 10 eV cut-off energy, respectively. The cut-off dependency is reduced as less than 10% (without the fix, the disagreement between 7.4 and 10 eV cut-off values is 18 times). However, these values are much higher than the value of 0.01 molecules/100 eV obtained without the bug fix on the electron attachment process, also 0.08 molecules/100 eV of Cobut et al. (1998). Even though, it is clearly reported that Melton cross-section overestimates the number of electron attachment process not only in liquid water but also in vapour water (Rawat et al., 2007), a new scale factor for adjusting the G-values should be found out.

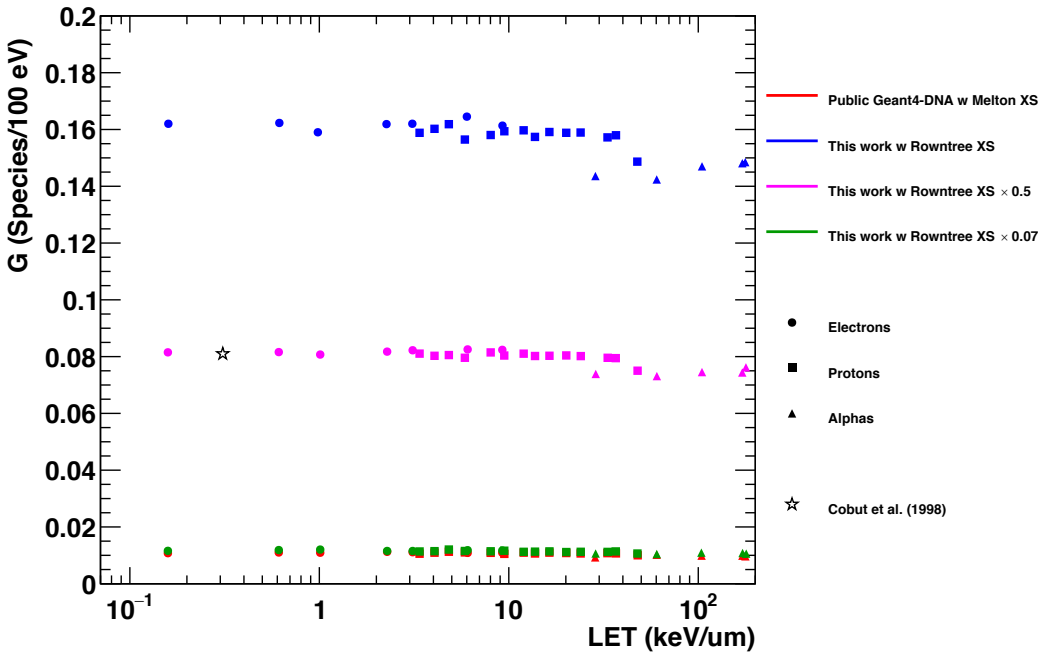


Figure 3.16: G-values of initial H_2O^- anions as a function of LET according to the different attachment cross-section models and scale factors.

The number of occurrences of electron attachment process is apparently not much affected by LET as shown in Figure 3.16. However, this number depends strongly on the selected attachment cross-section model. The public version of Geant4-DNA uses the Melton model (red symbols). One can reproduce such values by applying a 0.07 factor to the Rowntree cross-section model (green symbols). A factor of 0.5 allows to reproduce the calculated value of the paper of Cobut et al. (1998) (open star and pink symbols). Without scaling applied, values are twice larger (blue symbols). As it is difficult to determine what should be the most reliable value of this scaling parameter, we set it to the default value of 0.07 (which leads to similar results as the public version of Geant4-DNA) and we leave the possibility to the user to change this value when experimental data become available.

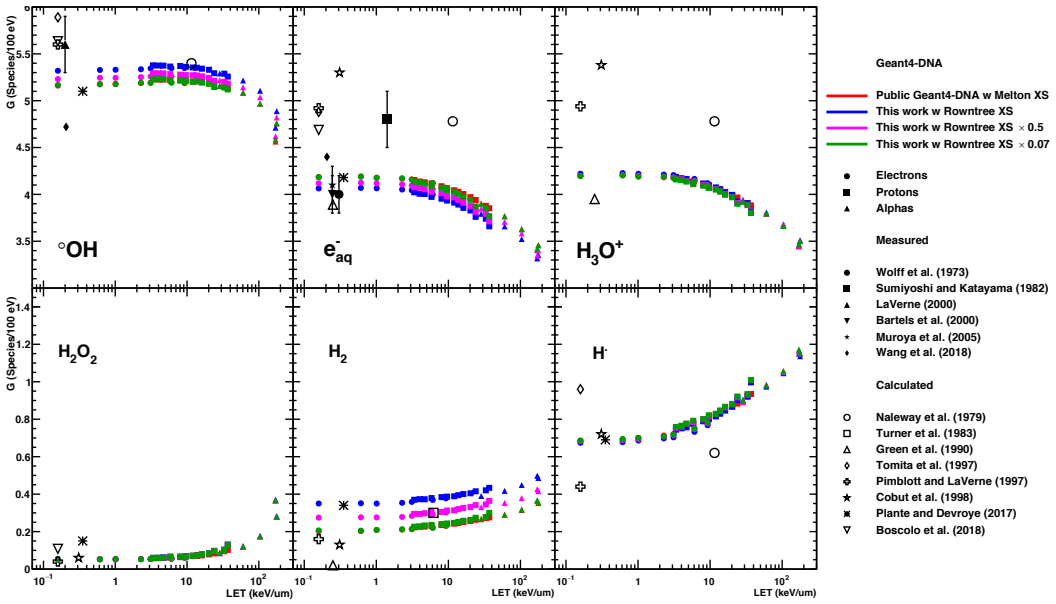


Figure 3.17: Initial G-values (at 1 ps) as a function of LET according to the different electron attachment cross-section models and scale factors.

The electron attachment process directly impacts the initial G-values of solvated electrons as shown in Figure 3.17. In addition, •OH radicals and H_2 molecules generated by the dissociation are impacted as well. It remains difficult to compare the calculated G-values with literature data due to the different measurement environments, solvent concentration and the measurement times. Nevertheless, some interesting agreement with the recent experimental data of Muroya et al. (2005) (closed down triangles, on e_{aq}^-) and Wang et al. (2018) (small stars, on e_{aq}^-) is observed in Figure 3.17.

3.2.7. Results for electron-hole recombination and dissociation channel

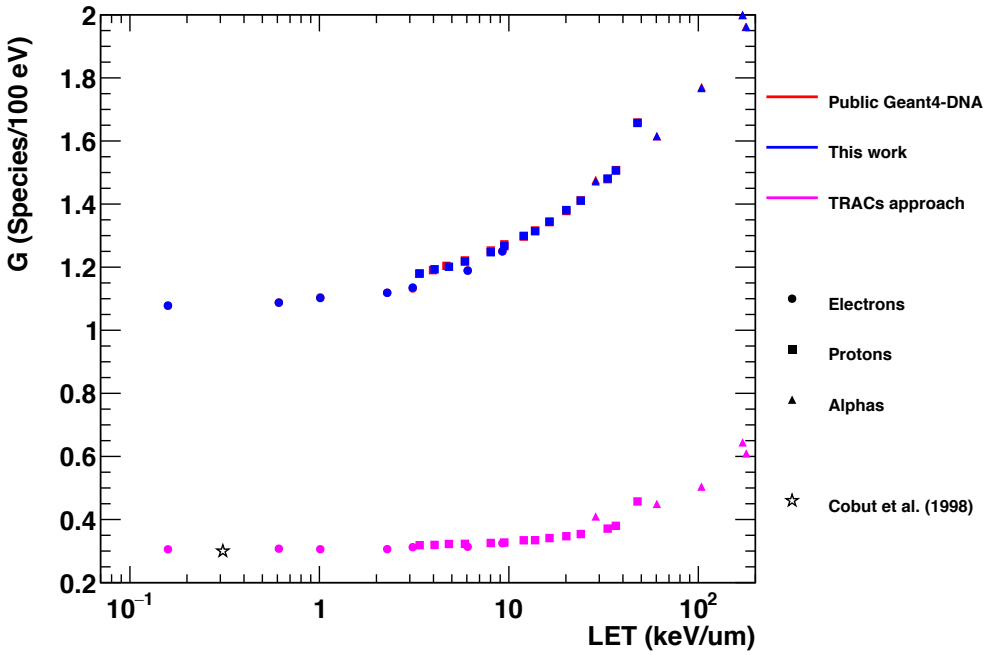


Figure 3.18: The number of occurrences of the electron-hole recombination process as a function of LET, according to the recombination model.

It is known that the number of electron-hole recombinations depends on the LET, due to the distance between solvated electron and water cation, as shown in Figure 3.18. In this study, the G-values of H_2O^* molecules generated by recombination are not affected by the recombination models, which differ only by their dissociation channels.

The approach of TRACs (using contact reaction and a reaction radius of 0.775 nm) allows to reproduce the value of Cobut et al. (1998) while it underestimates the number of recombinations by about 5 times less than the empirical model of Geant4-DNA. Thus, we recommend keeping the method using the empirical model.

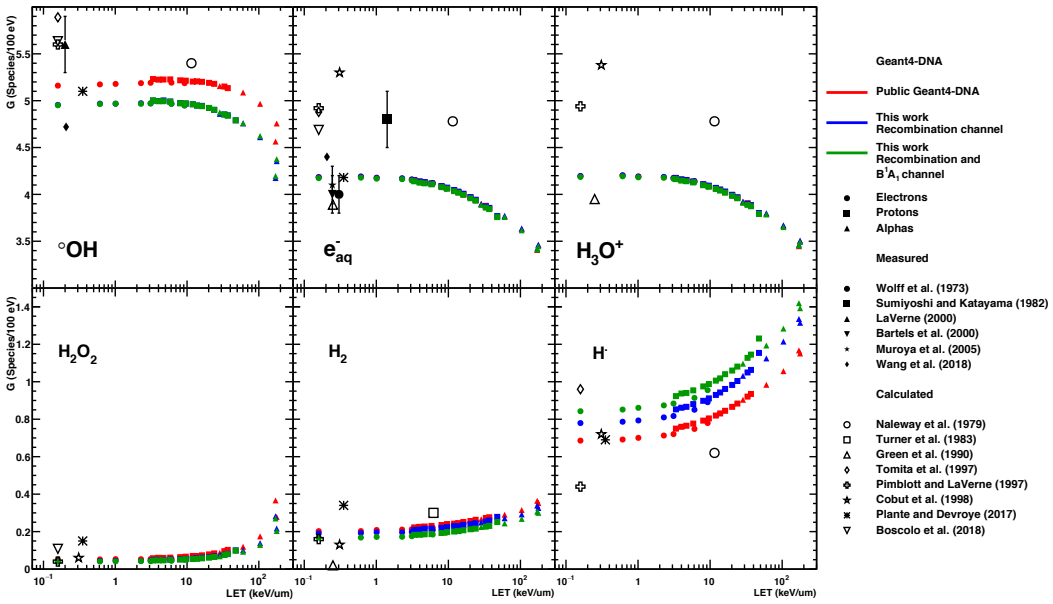


Figure 3.19: Initial G-values (at 1 ps) as a function of LET according to the dissociation channels of electron-hole recombination and B^1A_1 excitation.

Regarding the changes of dissociation channel for electron-hole recombination process, the most influenced species is •OH radicals which changes 85% to 63% in the number of recombination processes as shown by the blue markers of Figure 3.19. In addition, the probability of generating H• radicals increased from 55% to 67%, and the number of H_2 molecules generated by recombination process is slightly decreased from 15% to 13.65%. Moreover, the B^1A_1 dissociation channel of TRACs is evaluated as shown by the green markers of Figure 3.19. The changes of generation probability for H• radical and H_2 molecule induced by B^1A_1 excitation are 0% to 33.15% and 15% to 3.25%, respectively.

Compared to the G-values calculated by public Geant4-DNA (red markers), the G-values of •OH radical are getting closer to the most recent experimental data by Wang et al. (2018) (closed star, on •OH radical). In addition, the initial H_2 molecule yields are decreased to 0.168 molecules/100 eV for low LET electron. That value is similar with the assumption of Cobut et al. (1998) (asterisk) based on the experimental data of Schwarz (1969) (0.15 molecules/100 eV).

3.2.8. Influence of the pre-chemical processes on water radiolysis simulation

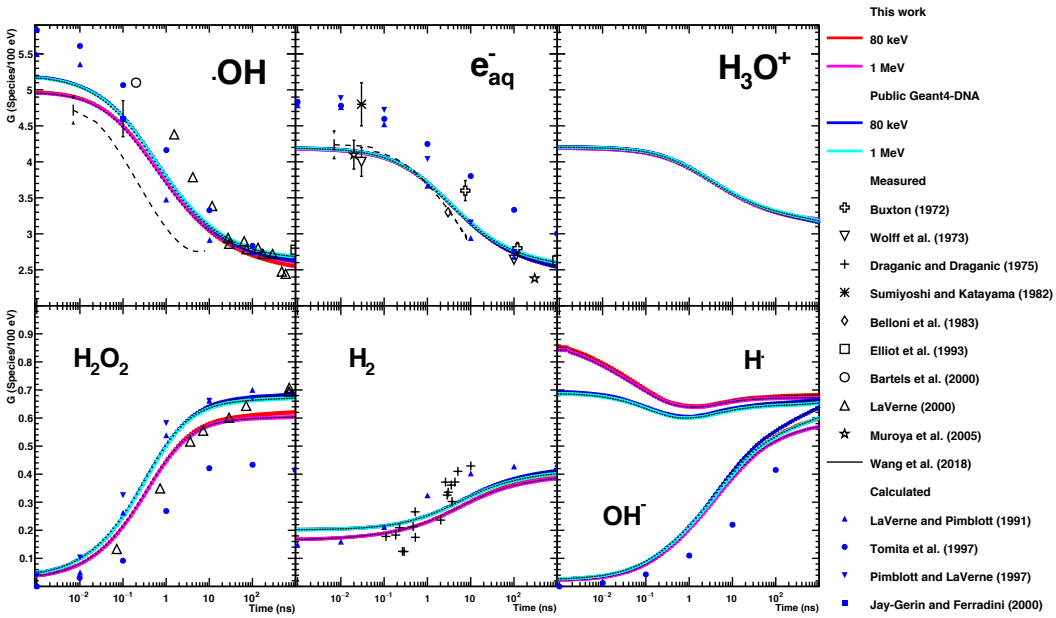


Figure 3.20: The time evolution of G-values for incident electrons.

We first present in Figure 3.20 radiochemical yields simulated as a function of time, using the public version of Geant4-DNA (blue and cyan curves) or the improved version of the pre-chemical stage of Geant4-DNA (including the corrections of the Auger effect, the electron attachment, the dissociation channels, and the displacement) proposed in this work (red and magenta curves). The LET values of all the references calculated and measured are in the range of 0.186-0.314 keV/µm. Thus, only 80 keV and 1 MeV electron results are shown, since they have LET values of 0.61 and 0.16 keV/µm respectively.

We observe that the red and magenta curves show lower G-values for •OH radicals, H_2O_2, H_2, and OH^- than the public version of Geant4-DNA. The G-values of those species calculated in this study are closer to recent experimental data such as the ones of Wang et al. (2018) (dashed line, on •OH and e_{aq}^-). In the case of H_2 molecules, the initial G-values at 1 ps are very similar to the values of Cobut et al. (1998), 0.15 molecules/100eV for ^{60}Co at 1 ps, supposedly based on the experiment of Schwarz et al. (1959), 0.42 molecules/100 eV at 1 µs. Moreover, the G-values at 1 µs match well the literature data.

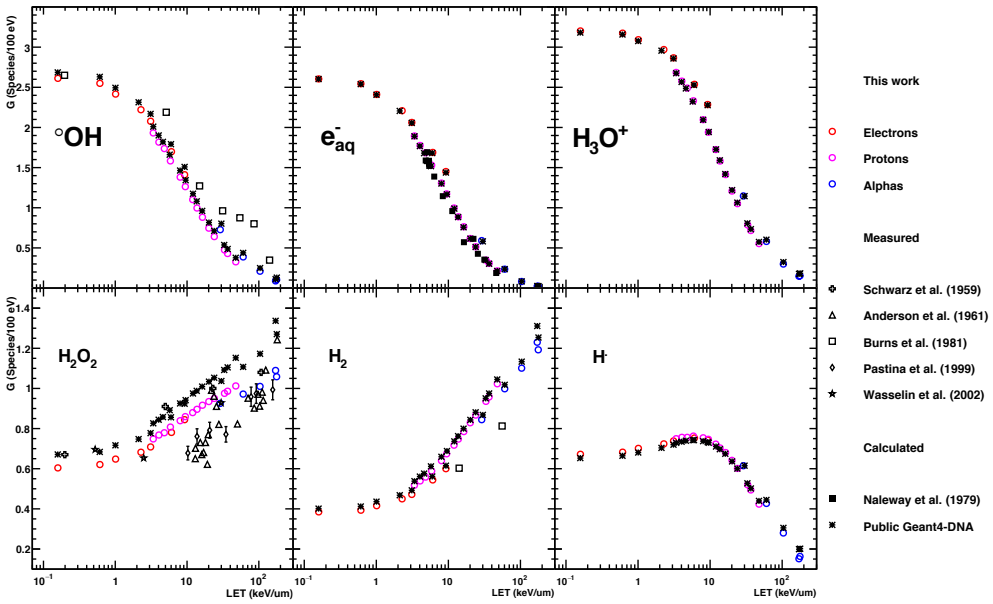


Figure 3.21: G-values at the end of the chemical stage as a function of LET.

Then, the G-values as a function of LET in the range of 0.16-177.85 keV/um are calculated at 1 μs as shown in Figure 3.21. The results show significant improvement on the G-values of H_2O_2. We can observe a slight change of $^{\bullet}OH$, H_2, and $H^{\bullet}$ due to the correction of the dissociation probabilities in the recombination process and B^1A_1 excitation. The main improvement here is that the G-values of H_2O_2, which previously showed higher values than the literature data using the public version of Geant4-DNA, are smaller due to the decrease of $^{\bullet}OH+^{\bullet}OH$ reactions.

3.3. Conclusions

In this chapter, we presented an overview of the pre-chemical stage in Geant4-DNA. In the first section, we verified the SBS method available in current Geant4-DNA for radiochemical yield simulation, and evaluated the influence of molecular species distribution on radiochemical yields. For that, we developed the new “chem6” example which is an extension of the “chem4” example, including time step limitation, LET calculation, and user macros.

The influence of electron elastic scattering is not significant on water radiolysis simulation. However, the results show that the default physics option (G4EmDNAPhysics_option2) calculates too diffusive electron transportation compared to the recent experimental data.

The electron thermalization models available in Geant4-DNA are improved thanks to this study. Specifically, in the case of the Ritchie and Terrissol electron thermalization models, we tried to reproduce the models as proposed in their original paper, and the water radiolysis simulations with Geant4-DNA leads to reasonable results. However, we underline that the Ritchie model currently implemented in Geant4-DNA is not reasonable above 7 eV because it is based on projected range (and not 3D range) and it assumes a linear correlation between energy and thermalization range. The Terrissol model uses cross-sections of amorphous-solid water without any correction for the phase influence between liquid and amorphous-solid, therefore, it always calculates longer thermalization range compared to the more realistic Meesungnoen model. For completeness, the “Meesungnoen amorphous" model using more recent cross section data and Monte Carlo method is implemented in this study and it is proposed in replacement of the Terrissol model. The G-values of the Kreipl model shows also similar results with the Meesungnnoen model, however, we decide to keep the Meesungnoen model as default model in order to be coherent with the original paper of Meesungnoen et al. (2002). In addition, we decide to use a kinetic energy of 1.7 eV for electrons produced from the auto-ionization process, and to thermalize these electrons with the same electron thermalization model used in the simulation for sub-excitation electrons.

The chemistry lists affect G-values of •OH radical and H_2O_2 significantly. We observed that the current version of Geant4-DNA with chemistry option 1 can lead to a good agreement with experimental data. Moreover, not only reaction rates but also diffusion coefficients significantly impact the amount of chemical reactions.

In the second section, we identified several errors on the displacement of fragments, on the dissociation probabilities, and on the sampling method of the electron attachment process, present in the public version of Geant4-DNA. We fixed these issues and evaluated their influence on the initial and final radiochemical yields as a function of LET.

The new sampling method proposed for the electron attachment process leads to frequent interactions when using the cross-section model (Melton) which is applicable to the vapour phase, currently available in the public version of Geant4-DNA. However, the newly proposed cross-

section (Rowntree) for amorphous ice also leads to an overestimation and requires to use a scaling factor if one wishes to obtain similar results as the predictions of the public version of Geant4-DNA. This parameter will be optimized when new electron attachment cross-section data for liquid water becomes available.

The new dissociation scheme for the electron-hole recombination process directly influences the G-values of $^{\bullet}OH$, H_2, and $H^{\bullet}$ species. Fortunately, our results become closer to recent experimental data, especially for H_2O_2 generated by $^{\bullet}OH+^{\bullet}OH$ reactions.

It remains however difficult to fully state which model is the most accurate, due to the large variety of experimental data which exact conditions cannot be easily simulated. Nevertheless, this work at least helped to maintain a better consistency with the original papers describing the models usable in Geant4-DNA.

As a further development step, the dissociation of H_2O^+ cation (Montenegro et al., 2007) and multi-ionized water molecules (Meesungnoen and Jay-Gerin, 2005) could be considered based on the most recent literature for the accurate modeling of pre-chemical stage.

References

Anderson, A. R. & Hart, E. J. 1961, Molecular product and free radical yields in the decomposition of water by protons, deuterons, and helium ions. *Radiat. Res.,* 14**,** 689-704.

Ballarini, F., Biaggi, M., Merzagora, M., et al. 2000, Stochastic aspects and uncertainties in the prechemical and chemical stages of electron tracks in liquid water: a quantitative analysis based on Monte Carlo simulations. *Radiat. Environ. Bioph.,* 39**,** 179-188.

Bartels, D. M., Cook, A. R., Mudaliar, M., et al. 2000, Spur decay of the solvated electron in picosecond radiolysis measured with time-correlated absorption spectroscopy. *J. Phys. Chem. A,* 104**,** 1686-1691.

Beaudre, A. 1988. *Simulation spatio-temporelle sur ordinateur des processus radiolytiques induits dans l'eau par des électrons.* Ph.D. thesis, Toulouse University 3.

Belli, M., Cherubini, R., Finotto, S., et al. 2009, RBE-LET Relationship for the Survival of V79 Cells Irradiated with Low Energy Protons. *Int. J. Radiat. Biol.,* 55**,** 93-104.

Belloni, J., Billiau, F., Delaire, J. A., et al. 1983, Ionizing radiation-liquid interactions: a comparative study of polar liquids. *Radiat. Phys. Chem.,* 21**,** 177-183.

Berg, H. C. 1993. *Random walks in biology,* New Jersey, Princeton University Press.

Berger, M. J., Coursey, J., Zucker, M., et al. 1998. *Stopping-power and range tables for electrons, protons, and helium ions*, NIST Physics Laboratory Gaithersburg, MD.

Bernal, M. A., Bordage, M. C., Brown, J. M. C., et al. 2015, Track structure modeling in liquid water: A review of the Geant4-DNA very low energy extension of the Geant4 Monte Carlo simulation toolkit. *Phys. Med.,* 31**,** 861-874.

Bisby, R. H., Cundall, R. B. & Sims, H. E. 1977, Linear energy transfer (LET) effects in the radiation-induced inactivation of papain. *Faraday Discuss. Chem. Soc.,* 63**,** 237-247.

Blohm, R. & Harder, D. 1985, Restricted LET: still a good parameter of radiation quality for electrons and photons. *Radiat. Prot. Dosim.,* 13**,** 337-381.

Bordage, M. C., Bordes, J., Edel, S., et al. 2016, Implementation of new physics models for low energy electrons in liquid water in Geant4-DNA. *Phys. Med.,* 32**,** 1833-1840.

Boscolo, D., Krämer, M., Durante, M., et al. 2018, TRAX-CHEM: A pre-chemical and chemical stage extension of the particle track structure code TRAX in water targets. *Chem. Phys. Lett.,* 698**,** 11-18.

Burns, W. G. & Marsh, W. R. 1981, Radiation chemistry of high-temperature (300-410 deg) water. *J. Chem. Soc. Faraday Trans. 1,* 77**,** 197-215.

Burns, W. G., May, R., Buxton, G. V., et al. 1981, Nanosecond proton pulse radiolysis of aqueous solutions. *J. Chem. Soc. Faraday Trans.,* 77**,** 1543-1551.

Burns, W. G. & Sims, H. E. 1981, Effect of radiation type in water radiolysis. *J. Chem. Soc. Faraday Trans.,* 77**,** 2803-2813.

Buxton, G. V. 1972, Nanosecond pulse radiolysis of aqueous solution containing oproton and hydroxyl radical scavengers. *Proc. R. Soc. A,* 328**,** 9-21.

Buxton, G. V. 1982. *Basic radiation chemistry of liquid water,* Dordrecht, Springer.

Buxton, G. V., Greenstock, C. L., Helman, W. P., et al. 1988, Critical review of rate constants for reactions of hydrated electrons, hydrogen atoms and hydroxyl radicals (·OH/·O− in Aqueous Solution. *J. Phys. Chem. Ref. Data,* 17**,** 513-886.

Champion, C., Incerti, S., Aouchiche, H., et al. 2009, A free-parameter theoretical model for describing the electron elastic scattering in water in the Geant4 toolkit. *Radiat. Phys. Chem.,* 78**,** 745-750.

Chen, C. Z. & Watt, D. E. 1985, Biophysical mechanism of radiation damage to Mammalian cells by X- and γ-rays. *Int. J. Radiat. Biol. Relat. Stud. Phys. Chem. Med.,* 49**,** 131-142.

Cobut, V., Frongillo, Y., Patau, J. P., et al. 1998, Monte Carlo simulation of fast electron and proton tracks in liquid water-I. Physical and physicochemical aspects. *Radiat. Phys. Chem.,* 51**,** 229-243.

Cobut, V., Jay-Gerin, J. P., Frongillo, Y., et al. 1996, On the dissociative electron attachment as a potential source of molecular hydrogen in irradiated liquid water. *Radiat. Phys. Chem.,* 47**,** 247-250.

Cortes-Giraldo, M. A. & Carabe, A. 2015, A critical study of different Monte Carlo scoring methods of dose average linear-energy-transfer maps calculated in voxelized geometries irradiated with clinical proton beams. *Phys. Med. Biol.,* 60**,** 2645-69.

Curtis, M. G. & Walker, I. C. 1992, Dissociative electron attachment in water and methanol (5-14 eV). *J. Chem. Soc. Faraday Trans.,* 88**,** 2805-2810.

de Berg, M., van Kreveld, M., Overmars, M., et al. 2008. *Computational geometry,* Berlin, Heidelberg, Springer.

Dingfelder, M., Hantke, D., Inokuti, M., et al. 1998, Electron inelastic-scattering cross sections in liquid water. *Radiat. Phys. Chem.,* 53**,** 1-18.

Draganić, Z. D. & Draganić, I. G. 1975, Formation of primary reducing yields (Geaq- and GH2) in the radiolysis of aqueous solutions of some positive ions. *Int. J. Radiat. Phys. Chem.,* 7**,** 381-386.

Dubochet, J., Adrian, M., Chang, J.-J., et al. 1988, Cryo-electron microscopy of vitrified specimens. *Q. Rev. Biophys,* 21**,** 129-228.

Elliot, A. J. 1994. Rate constants and G-values for the simulation of the radiolysis of light water over the range 0-300 deg. Ontario, Canada: Chalk River Laboratories.

Elliot, A. J., Chenier, M. P. & Ouellette, D. C. 1993, Temperature dependence of g values for H2O and D2O irradiated with low linear energy transfer radiation. *J. Chem. Soc. Faraday Trans.,* 89**,** 1193-1197.

Emfietzoglou, D., Kyriakou, I., Abril, I., et al. 2012, Inelastic scattering of low-energy electrons in liquid water computed from optical-data models of the Bethe surface. *Int. J. Radiat. Biol.,* 88**,** 22-8.

Francis, Z., Incerti, S., Capra, R., et al. 2011, Molecular scale track structure simulations in liquid water using the Geant4-DNA Monte-Carlo processes. *Appl. Radiat. Isotopes,* 69**,** 220-6.

Friedland, W., Dingfelder, M., Kundrat, P., et al. 2011, Track structures, DNA targets and radiation effects in the biophysical Monte Carlo simulation code PARTRAC. *Mutat. Res.,* 711**,** 28-40.

Frongillo, Y., Goulet, T., Fraser, M.-J., et al. 1998, Monte Carlo simulation of fast electron and proton tracks in liquid water- II. Nonhomogeneous chemistry. *Radiat. Phys. Chem.,* 51**,** 245-254.

Goulet, T. & Jay-Gerin, J. P. 1989, Thermalization of subexcitation electrons in solid water. *Radiat. Res.,* 118**,** 46-62.

Goulet, T. & Jay-Gerin, J. P. 1992, On the reactions of hydrated electrons with OH·and H3O+. Analysis of photoionization experiments. *J. Chem. Phys.,* 96**,** 5076-5087.

Green, N. J. B., Pilling, M. J., Pimblott, S. M., et al. 1990, Stochastic modeling for fast kinetics in a radiation track. *J. Phys. Chem.,* 94**,** 251-258.

Han, P. & Bartels, D. M. 1990, H/D isotope effects in water radiolysis. 2. dissociation of electronically excited water. *J. Phys. Chem.,* 94**,** 5824-5833.

ICRU 1970. Linear energy transfer. *ICRU Report 16.* Journal of the ICRU: The International Commission on Radiation Units and Measurements.

ICRU 2014. Key data for ionizing-radiation dosimetry: measurement standards and applications. *ICRU Report 90.* Journal of the ICRU: The International Commission on Radiation Units and Measurements.

Incerti, S., Kyriakou, I., Bernal, M., et al. 2018, Geant4-DNA example applications for track structure simulations in liquid water: a report from the Geant4-DNA Project. *Med. Phys.,* 45**,** e722-e739.

Jay-Gerin, J.-P. & Ferradini, C. 2000, A new estimate of the OH radical yield at early times in the radiolysis of liquid water. *Chem. Phys. Lett.,* 317**,** 388-391.

Jonah, C. D., Matheson, M. S., Miller, J. R., et al. 1976, Yield and decay of the hydrated electron from 100 ps to 3 ns. *J. Phys. Chem.,* 80**,** 1267-1270.

Jonah, C. D. & Miller, J. R. 1977, Yield and decay of the OH radical from 200 ps to 3 ns. *J. Phys. Chem.,* 81**,** 1974-1976.

Kaplan, I. G., Miterev, A. M. & Sukhonosov, V. Y. 1990, Simulation of the primary stage of liquid water radiolysis. *Radiat. Phys. Chem.,* 36**,** 493-498.

Karamitros, M., Luan, S., Bernal, M. A., et al. 2014, Diffusion-controlled reactions modeling in Geant4-DNA. *J. Comput. Phys.,* 274**,** 841-882.

Karamitros, M., Mantero, A., Incerti, S., et al. 2011, Modeling radiation chemistry in the Geant4 toolkit. *Prog. Nucl. Sci. Tech.,* 2**,** 503-508.

Konovalov, V. V., Pusep, A. Y., Raitsimring, A. M., et al. 1986, Pulsed photoemission from metals into electrolyte solutions: studies on slow-electron thermalization lengths by the method of eaq- recombination. *Chem. Phys. Lett.,* 126**,** 472-476.

Konovalov, V. V., Raitsimring, A. M., Tsvetkov, Y. D., et al. 1985, The thermalization length of low-energy electrons determined by nanosecond photoemission into aqueous electrolyte solutions. *Chem. Phys.,* 93**,** 163-170.

Kreipl, M. S., Friedland, W. & Paretzke, H. G. 2009, Time- and space-resolved Monte Carlo study of water radiolysis for photon, electron and ion irradiation. *Radiat. Environ. Bioph.,* 48**,** 11-20.

Kreitus, I. V., Benderskii, V. A. & Tiliks, Y. E. 1982, Photoelectron emission from metal into concentrated electrolytes: Part II. Thermalization path-length of low-energy electrons and eaq- bulk reactions. *J. Electroanal. Chem. Interf. Electrochem.,* 140**,** 311-328.

LaVerne, J. A. 2000, OH radicals and oxidizing products in the gamma radiolysis of water. *Radiat. Res.,* 153**,** 196-200.

LaVerne, J. A. & Pimblott, S. M. 1991, Scavenger and time dependences of radicals and molecular products in the electron radiolysis of water: examination of experiments and models. *J. Phys. Chem.,* 95**,** 3196-3206.

Lu, H., Long, F. H., Bowman, R. M., et al. 1989, Femtosecond studies of electron-cation geminate recombination in water. *J. Phys. Chem.,* 93**,** 27-28.

Martin, C. 2003. *Modélisation des dommages radioinduits sur l'ADN: prise en compte des radicaux libres et des réparations primaires.* Ph.D. thesis, Toulouse University.

Meesungnoen, J., Jay-Gerin, J.-P., Filali-Mouhim, A., et al. 2002, Low-energy electron penetration range in liquid water. *Radiat. Res.,* 158**,** 657-660.

Meesungnoen, J. & Jay-Gerin, J. P. 2005, High-LET radiolysis of liquid water with 1H+, 4He2+, 12C6+, and 20Ne9+ ions: effects of multiple ionization. *J. Phys. Chem. A,* 109**,** 6406-6419.

Melton, C. E. 1972, Cross sections and interpretation of dissociative attachment reactions producing OH−, O−, and H− in H2O. *J. Chem. Phys.,* 57**,** 4218-4225.

Michalik, V., Begusova, M. & Bigildeev, E. A. 1998, Computer-aided stochastic modeling of the radiolysis of liquid water. *Radiat. Res.,* 149**,** 224-236.

Michaud, M. & Sanche, L. 1987a, Absolute vibrational excitation cross sections for slow-electron (1-18 eV) scattering in solid H2O. *Phys. Rev. A,* 36**,** 4684.

Michaud, M. & Sanche, L. 1987b, Total cross sections for slow-electron (1–20 eV) scattering in solidH2O. *Phys. Rev. A,* 36**,** 4672-4683.

Michaud, M., Wen, A. & Sanche, L. 2003, Cross sections for low-energy (1–100 eV) electron elastic and inelastic scattering in amorphous ice. *Radiat. Res.,* 159**,** 3-22.

Montenegro, E. C., Scully, S. W. J., Wyer, J. A., et al. 2007, Evaporation, fission and auto-dissociation of doubly charged water. *J. Electron Spectrosc.,* 155**,** 81-85.

Muroya, Y., Lin, M., Wu, G., et al. 2005, A re-evaluation of the initial yield of the hydrated electron in the picosecond time range. *Radiat. Phys. Chem.,* 72**,** 169-172.

Naleway, C. A., Sauer Jr, M. C., Jonah, C. D., et al. 1979, Theoretical analysis of the LET dependence of transient yields observed in pulse radiolysis with ion beams. *Radiat. Res.,* 77**,** 47-61.

Neff, H., Sass, J. K., Lewerenz, H. J., et al. 1980, Photoemission studies of electron localization at very low excess energies. *J. Phys. Chem.,* 84**,** 1135-1139.

Nikjoo, H., O'Neill, P., Goodhead, D. T., et al. 1997, Computational modelling of low-energy electron-induced DNA damage by early physical and chemical events. *Int. J. Radiat. Biol.,* 71**,** 467-483.

Nikjoo, H., Uehara, S., Emfietzoglou, D., et al. 2006, Track-structure codes in radiation research. *Radiat. Meas.,* 41**,** 1052-1074.

Pastina, B. & LaVerne, J. A. 1999, Hydrogen peroxide production in the radiolysis of water with heavy ions. *J. Phys. Chem. A,* 103**,** 1592-1597.

Peudon, A. 2007. *Prise en compte de la structure moléculaire pour la modélisation des commages biologiques radio-induits.* Ph.D. thesis, Toulouse 3.

Peukert, D., Incerti, S., Kempson, I., et al. 2019, Validation and investigation of reactive species yields of Geant4-DNA chemistry models. *Med. Phys.,* 46**,** 983-998.

Pimblott, S. M. & LaVerne, J. A. 1997, Stochastic simulation of the electron radiolysis of water and aqueous solutions. *J. Phys. Chem. A,* 101**,** 5828-5838.

Pimblott, S. M. & LaVerne, J. A. 1998, Effect of electron energy on the radiation chemistry of liquid water. *Radiat. Res.,* 150**,** 159-169.

Plante, I. & Devroye, L. 2017, Considerations for the independent reaction times and step-by-step methods for radiation chemistry simulations. *Radiat. Phys. Chem.,* 139**,** 157-172.

Ramos-Mendez, J., Perl, J., Schuemann, J., et al. 2018, Monte Carlo simulation of chemistry following radiolysis with TOPAS-nBio. *Phys. Med. Biol.,* 63**,** 105014.

Rawat, P., Prabhudesai, V. S., Aravind, G., et al. 2007, Absolute cross sections for dissociative electron attachment to H2O and D2O. *J. Phys. B-At. Mol. Opt.,* 40**,** 4625-4636.

Ritchie, R. H., Hamm, R. N., Turner, J. E., et al. 1994. Interactions of low-energy electrons with condensed matter: relevance for track structure. *Computational approaches in molecular radiation biology.* Springer.

Rotenberg, Z. A. & Gurevich, Y. Y. 1975, Photodiffusion phenomena stimulated by photoelectron emission in solutions. *J. Electroanal. Chem. Interf. Electrochem.,* 66**,** 165-182.

Rowe, B. R., Vallée, F., Queffelec, J. L., et al. 1988, The yield of oxygen and hydrogen atoms through dissociative recombination of H2O+ ions with electrons. *J. Chem. Phys.,* 88**,** 845-850.

Rowntree, P., Parenteau, L. & Sanche, L. 1991, Electron stimulated desorption via dissociative attachment in amorphous H2O. *J. Chem. Phys.,* 94**,** 8570-8576.

Salvat, F., Jablonski, A. & Powell, C. J. 2005, ELSEPA - Dirac partial-wave calculation of elastic scattering of electrons and positrons by atoms, positive ions and molecules. *Comput. Phys. Commun.,* 165**,** 157-190.

Sauer Jr, M. C., Schmidt, K. H., Hart, E. J., et al. 1977, LET dependence of transient yields in the pulse radiolysis of aqueous systems with deuterons and alpha particles. *Radiat. Res.,* 70**,** 91-106.

Schwarz, H. A. 1969, Applications of the spur diffusion model to the radiation chemistry of aqueous solutions. *J. Phys. Chem.,* 73**,** 1928-1937.

Schwarz, H. A., Caffrey Jr, J. M. & Scholes, G. 1959, Radiolysis of neutral water by cyclotron produced deuterons and helium ions. *J. American Chem. Soc.,* 81**,** 1801-1809.

Shin, W. G., Bordage, M. C., Emfietzoglou, D., et al. 2018, Development of a new Geant4-DNA electron elastic scattering model for liquid-phase water using the ELSEPA code. *J. Appl. Phys.,* 124.

Shin, W. G., Ramos-Mendez, J., Faddegon, B., et al. 2019, Evaluation of the influence of physical and chemical parameters on water radiolysis simulations under MeV electron irradiation using Geant4-DNA. *J. Appl. Phys.,* 126.

Sumiyoshi, T. & Katayama, M. 1982, The yield of hydrated electrons at 30 picoseconds. *Chem. Lett.,* 11**,** 1887-1890.

Taube, H. 1956, Photochemical reactions of ozone in solution. *Trans. Faraday Soc.,* 53**,** 656-665.

Terrissol, M. & Beaudre, A. 1990, Simulation of space and time evolution of radiolytic species induced by electrons in water. *Radiat. Prot. Dosim.,* 31**,** 175-177.

Thomsen, C. L., Madsen, D., Keiding, S. R., et al. 1999, Two-photon dissociation and ionization of liquid water studied by femtosecond transient absorption spectroscopy. *J. Chem. Phys.,* 110**,** 3453-3462.

Tomita, H., Kai, M., Kusama, T., et al. 1997, Monte Carlo simulation of physicochemical processes of liquid water radiolysis. *Radiat. Environ. Bioph.,* 36**,** 105-116.

Tran, H. N., Karamitros, M., Ivanchenko, A., et al. 2016, Geant4 Monte Carlo simulation of absorbed dose and radiolysis yields enhancement from a gold nanoparticle under MeV proton irradiation. *Nucl. Instrum. Meth. B,* 373**,** 126-139.

Trumbore, C. N., Short, D. R., Fanning, J. E. J., et al. 1978, Effects of pulse dose on hydrated electron decay kinetics in the pulse radiolysis of water. A computer modeling study. *J. Phys. Chem.,* 82**,** 2762-2767.

Turner, J. E., Magee, J. L., Wright, H. A., et al. 1983, Physical and chemical development of electron tracks in liquid water. *Radiat. Res.,* 96**,** 437-449.

Uehara, S., Nikjoo, H. & Goodhead, D. T. 1993, Cross-sections for water vapour for the Monte Carlo electron track structure code from 10 eV to the MeV region. *Phys. Med. Biol.,* 37**,** 1841-1858.

Wang, F., Schmidhammer, U., Larbre, J. P., et al. 2018, Time-dependent yield of the hydrated electron and the hydroxyl radical in D2O: a picosecond pulse radiolysis study. *Phys. Chem. Chem. Phys.*

Wasselin-Trupin, V., Baldacchino, G., Bouffard, S., et al. 2002, Hydrogen peroxide yields in water radiolysis by high-energy ion beams at constant LET. *Radiat. Phys. Chem.,* 65**,** 53-61.

Wolff, R. K., Bronskill, M. J., Aldrich, J. E., et al. 1973, Picosecond pulse radiolysis. IV. Yield of the solvated electron at 30 picoseconds. *J. Phys. Chem.,* 77**,** 1350-1355.

Chapter 4

Chemical stage

Table of contents

4.1. Green Function Diffusion Equation……136
4.1.1. Totally diffusion-controlled reactions (Types I and III) ……...138
4.1.2. Partially diffusion-controlled reactions (Types II and IV) ……140
4.1.3. First-order reactions and background reactions (Type VI reactions) ……146
4.1.4. Contact reactions……..……147
4.2. The independent reaction time technique……148
4.2.1. Search range……..……150
4.2.2. Sampling of diffusion-controlled reactions……...151
4.2.3. Sampling of type VI reactions and contact reactions……153
4.2.4. Reaction site and position of secondaries……...154
4.3. Implementation of IRT in the chemistry framework of Geant4-DNA……155
4.4. Radiochemical yield simulation in water……156
4.5. Validation of implementation……...157
4.6. Results for G-values versus time……160
4.7. Results for G-values versus LET……163
4.8. Conclusions……164
References……165

During the chemical stage, Geant4-DNA allows to simulate the transportation of molecular species and the chemical reactions occurring between them, as described in detail in the paper of Karamitros et al. (2014). Briefly, the chemistry algorithm of Geant4-DNA is based on the step-by-step (SBS) method using the Debye-Smoluchowski Brownian diffusion equation from which encounters between reactants are dynamically sampled. In addition, it adopts a Brownian bridge technique to determine encounter times between discrete time steps. The SBS method is able to simulate radiochemical yields and concentrations of molecular species, however, the simulation time needed for calculating the distance between species at every time step is a huge computational burden.

Because of such limitation, the independent reaction time (IRT) method has been proposed by Green et al. (1990). This method relies on the independent pair approximation to simulate the reaction times between two reactants from their initial separation distance, reaction rate coefficient, and diffusion coefficients in order to avoid the burden of diffusing each molecular species. Several MCTS tools such as TRACIRT and RITRACKS (Frongillo et al., 1998, Plante, 2011) have implemented the IRT method which showed close predictions of radiochemical yields to the SBS method (Plante, 2011, Plante and Devroye, 2017).

This study is a collaborative work between the TOPAS-nBio and the Geant4-DNA collaborations in order to implement IRT into Geant4-DNA and re-engineer the IRT approach initially proposed by Jose Ramos-Mendez (UCSF, CA, USA) in TOPAS-nBio, similarly as TOPAS-nBio implemented previously the Geant4-DNA SBS approach (Ramos-Mendez et al., 2018, Schuemann et al., 2019). In order to validate this IRT implementation, I also used the "chem6" radiation chemistry example developed in this thesis to calculate radiochemical yields as a function of time and linear energy transfer (LET), as presented in the previous chapter. Finally, the results of the IRT method have been compared to experimental data and to literature simulation results.

4.1. Green Function Diffusion Equation (GFDE)

The Debye-Smoluchowski equation is typically used for the modelling of the Brownian diffusive motion of molecular species in a solvent. The variation of spatial density ρ of a reactant as a function of time can be derived from the equation:

$$\frac{\partial \rho}{\partial t} = \nabla D(\nabla \rho - \frac{F\rho}{k_B T}) \tag{4.1}$$

where D is the diffusion coefficient of the particle, k_B is the Boltzmann constant, T is the temperature of the solvent, and F is the external force acting on the particle such as electrostatic force.

Several approaches to solve this equation have been proposed, however an exact analytic solution cannot be obtained (Delaire et al., 1981, Clifford et al., 1984). Thus, a methodology using Green's function was proposed by Tachiya (1978), Tachiya (1979b), Tachiya (1979a), Sano and Tachiya (1979). Cobut et al. (1998), Frongillo et al. (1998) applied this solution into MCTS tools,

called TRACPRO and TRACELE, from which RITRACKS has been developed (Plante and Devroye, 2017).

In this study, 15 molecular species and 72 chemical reactions are employed, from the work of Plante and Devroye (2017) based on the papers of Elliot (1994), Frongillo et al. (1998). The molecular species with their diffusion coefficients and radii are shown in Table 4.1.

Table 4.1: Diffusion coefficients (D) and radii of molecular species (R) used in this study (Plante and Devroye, 2017).

Molecular species	**Diffusion coefficients D (nm^2s^{-1})**	**Radii R (nm)**
$H^{\cdot}$	7.0×10^9	0.19
$^{\cdot}OH$	2.2×10^9	0.22
H_2O_2	2.3×10^9	0.21
H_2	4.8×10^9	0.14
e_{aq}^-	4.9×10^9	0.50
H_3O^+	9.46×10^9	0.25
OH^-	5.3×10^9	0.33
O_2	2.4×10^9	0.17
$\dot{O_2^-}$	1.75×10^9	0.22
$H\dot{O}_2$	2.3×10^9	0.21
HO_2^-	1.4×10^9	0.25
$^*O(^3P)$	2.0×10^9	0.20
$\dot{O}^-$	2.0×10^9	0.25
$\dot{O_3^-}$	2.0×10^9	0.20
O_3	2.0×10^9	0.20

* $O(^3P)$ represents atomic oxygen which has triplet P state.

The speed of a chemistry reaction between two reactants is typically described by the reaction rate constant which is the change in concentration over the change in time with unit of $M^{-1}s^{-1}$. The unit of molar concentration M, which represents the spatial density of the molecules, is equivalent to mol/dm^3.

In diffusion-controlled reactions, one can distinguish the diffusion-encounter rate constant k_{diff} and the activation rate constant k_{act}. The empirically observed rate constant k_{obs} takes into account these two rate constants as:

$$A + B \underset{\text{or}}{\overset{k_{diff}}{\longrightarrow}} AB \overset{k_{act}}{\longrightarrow} Products \quad \overset{k_{obs}}{\longrightarrow} \tag{4.2}$$

where AB is the intermediate state of an encounter pair just before the chemical reaction, k_{diff} represents the rate constant for the encounter of two reactants, and k_{act} is the rate constant of reactive loss (or gain of products) measured if it is not influenced by diffusion (Elliot et al., 1990).

The constants k_{act} and k_{diff} are correlated with k_{obs} based on the long-time limit relation of the Noyes equation (Rice, 1985, Elliot, 1994):

$$\frac{1}{k_{obs}} = \frac{1}{k_{diff}} + \frac{1}{k_{act}} \tag{4.3}$$

Based on these reaction rate constants, the diffusion-controlled reactions are classified into several types in the paper of Frongillo et al. (1998). Briefly, totally diffusion-controlled (TDC) reactions which have high enough k_{act} ($k_{act} \to \infty$, types I) are distinguished from partially diffusion-controlled (PDC) reactions ($0 < k_{act} < \infty$, types II) (Sano and Tachiya, 1979).

It is obvious that the electrical charge of reactants affects the encounter rate k_{diff} due to Coulomb force. The repulsive reactants of same charge will decrease k_{diff}, and the attractive reactants of opposite charge will increase k_{diff}. TDC and PDC reactions with electrical interaction are classified into type III and type IV reactions, respectively.

In addition, few TDC reactions considering the molecular spins of specific reactants ($H^{\bullet}$ or e_{aq}^-) (Frongillo et al., 1998, Plante and Devroye, 2017), called type V reactions, are modeled. The non-diffusion-controlled (NDC) reactions such as first-order reactions are also included in the reaction list as well as background reactions (type VI). All reaction types involved in this study are summarized in Table 4.2.

Table 4.2: Brief descriptions of reaction types.

Reaction type	Diffusion controlled	Type of reactants	Description
Type I	TDC	Neutral	Section 4.1.1
Type II	PDC	Neutral	Section 4.1.2
Type III	TDC	Charged	Section 4.1.1
Type IV	PDC	Charged	Section 4.1.2
Type V	TDC	$H^{\bullet}$ or e_{aq}^- (molecular spins)	Section 4.1.1
Type VI	NDC	H_2O, H_3O^+, and OH^- in solvent	Section 4.1.3

4.1.1. Totally diffusion-controlled reactions (Types I and III)

TDC reaction is a reaction type for which every encounter between reactants within the reaction radius σ leads to a reaction. According to the GFDE, the probability of the separation distance r for reaction type I (p_I) for the two reactants at initial distance r_0 as a function of time t can be derived as:

$$4\pi r r_0 p_I(r,t|r_0) = \frac{1}{\sqrt{4\pi Dt}}\left\{\exp\left[-\frac{(r-r_0)^2}{4Dt}\right] - \exp\left[-\frac{(r+r_0-2\sigma)^2}{4Dt}\right]\right\} \tag{4.4}$$

where D is the diffusion coefficient calculated by summing the diffusion coefficients of reactants $D_A + D_B$, and σ is their reaction radius.

Therefore, the cumulative distribution function (CDF) P_I of the reaction as a function of time can be derived as one minus the reactant's survival probability integrating equation (4.4):

$$P_I(t|r_0) = 1 - \int_{\sigma}^{\infty} 4\pi r^2 p_I(r,t|r_0)dr = \frac{\sigma}{r_0} erfc\left[\frac{r_0 - \sigma}{\sqrt{4Dt}}\right] \tag{4.5}$$

where *erfc*(*x*) is the complementary error function which can be expressed as 1-*erf*(*x*), and the reaction radius σ can be calculated by Smoluchowski reaction theory:

$$\sigma = \frac{k_{diff}}{4\pi D} \tag{4.6}$$

In the case of TDC reactions, the activation rate constant is high enough ($k_{act} \rightarrow \infty$) as mentioned before, so that $k_{diff} = k_{obs}$ and $\sigma = k_{obs}/4\pi D$. One thing has to be noted: a factor of 2 is taken into account in reaction radius if the reactants A and B are of same type, in order to avoid double sampling of reactions (this also applies to the other types of reactions).

Moreover, the probabilities of contact reaction (which will be described in section 4.1.4), when the initial separation between two reactants is closer than reaction radius, are always 100% because every collision leads to a reaction of the TDC type. In the MC simulations, the list of type I reactions and the corresponding parameters used are shown in Table 4.3.

Table 4.3: Type I reactions, corresponding observed reaction rate (k_{obs}), reaction radius (σ), and probability of contact reaction ($P_{Contact}$) (Plante and Devroye, 2017).

Reaction	k_{obs} **($M^{-1}s^{-1}$)**	σ **(nm)**	$P_{Contact}$ **(%)**
$^{*}H^{\cdot} + H^{\cdot} \rightarrow H_2$	5.03×10^9	0.10	100
$^{*}H^{\cdot} + e_{aq}^{-} + H_2O \rightarrow H_2 + OH^{-}$	2.50×10^{10}	0.28	100
$H^{\cdot} + O(^3P) \rightarrow {}^{\cdot}OH$	2.02×10^{10}	0.30	100
$H^{\cdot} + \dot{O}^{-} \rightarrow OH^{-}$	2.00×10^{10}	0.29	100
${}^{\cdot}OH + O(^3P) \rightarrow H\dot{O}_2$	2.02×10^{10}	0.64	100
$H\dot{O}_2 + O(^3P) \rightarrow O_2 + {}^{\cdot}OH$	2.02×10^{10}	0.62	100
$O(^3P) + O(^3P) \rightarrow O_2$	2.20×10^{10}	1.45	100

[*] Type V reactions.

For the TDC reactions between charged particles, the electrostatic interaction between the two reactants should be considered. The paper of Clifford et al. (1984) suggested to use the same diffusion equation for the type III reaction CDF (P_{III}) as for type I reaction using an effective distance r_{eff} and σ_{eff}:

$$P_{III}(t|r_0) = \frac{\sigma_{eff}}{r_{eff}} erfc\left[\frac{r_{eff} - \sigma_{eff}}{\sqrt{4Dt}}\right] \tag{4.7}$$

The effective distance r_{eff} and effective reaction radius σ_{eff} are the distances scaled by the electric field between particles:

$$r_{eff} = \frac{-r_c}{1-\exp(r_c/r)} \text{ and } \sigma_{eff} = \frac{-r_c}{1-\exp(r_c/\sigma)} \tag{4.8}$$

Equation (4.7) is consistent with equation (4.5) of TDC when $r_c \rightarrow 0$, when there is no electrostatic interaction between two reactants, r_{eff} and σ_{eff} converge to r and σ, respectively. r_c is the Onsager's radius, defined as the distance between two charged reactants which respectively

have electrical charge e_A and e_B when the electrostatic and thermal energy are equal (Karamitros et al., 2014):

$$r_c = \frac{e_A e_B}{4\pi\varepsilon k_B T} \tag{4.9}$$

In our case, a value of ± 0.712 nm is used for r_c at temperature $T = 293.15\ K$ (20 °C), k_B is the Boltzmann constant, and ε is the relative permittivity of the solvent (water). When the polarities of the charges are different ($e_A \cdot e_B < 0$) and the Onsager's radius is minus, the effective reaction radius might be bigger than reaction radius (more reactive) due to the electronic attraction. It is opposite when the polarities of the charges are the same ($e_A \cdot e_B > 0$).

Unlike equation (4.6), $k_{diff}/4\pi D$ might be equal to σ_{eff} instead of σ in type III reactions due to the influence of electrostatic interaction, but still the probabilities of contact reactions when r_{eff} is smaller than σ_{eff} are 100% (this will be described in section 4.1.4), as shown in Table 4.4.

Table 4.4: Type III reactions, corresponding observed reaction rate constant (k_{obs}), Onsager's radius (r_c), effective reaction radius (σ_{eff}), and probability of contact reaction ($P_{Contact}$) (Plante and Devroye, 2017).

Reaction	k_{obs} **($M^{-1}s^{-1}$)**	r_c **(nm)**	σ_{eff} **(nm)**	$P_{Contact}$ **(%)**
$^{*}e_{aq}^{-} + e_{aq}^{-} + H_2O \rightarrow H_2 + 2OH^{-}$	6.36×10^{9}	0.71	0.17	100
$H_3O^{+} + OH^{-} \rightarrow 2H_2O$	1.13×10^{11}	-0.71	1.01	100
$H_3O^{+} + \dot{O_3^{-}} \rightarrow {}^{\cdot}OH + O_2 + H_2O$	9.00×10^{10}	-0.71	1.04	100

* Type V reaction.

The molecular spins of hydrogen radical $H^{\bullet}$ and solvated electron e_{aq}^{-} are kept even after 1 μs which is typically the end-time of chemistry simulation (Fessenden et al., 1981). The spin statistical factor could affect the radiolysis results on these reactions, called type V reactions, because only the singlet state (50% probability of existence in nature) leads to the chemical reactions. Unlike the equivalent approach of RITRACKS, which randomly samples with a factor of ¼ based on the weighted σ (or σ_{eff}), we just used the reaction radius σ (or σ_{eff}) calculated by k_{obs} and equation (4.6).

4.1.2. Partially diffusion-controlled reactions (Types II and IV)

PDC reactions, for which the activation rate constant is not high enough ($0 < k_{act} < \infty$), aren't activated for every single encounter between reactants. It is reported that the probability of the PDC reactions between two reactants encountered within reaction radius depends on the probability of geminate recombination (Tachiya, 1978). The GFDE of PDC reactions when there is no electrostatic interaction is:

$$\begin{aligned} 4\pi r r_0 p_{II}(r,t|r_0) = {} & \frac{1}{\sqrt{4\pi D t}}\left\{\exp\left[-\frac{(r-r_0)^2}{4Dt}\right] + \exp\left[-\frac{(r+r_0-2\sigma)^2}{4Dt}\right]\right\} \\ & - \alpha W\left(\frac{r+r_0-2\sigma}{\sqrt{4Dt}}, \alpha\sqrt{Dt}\right) \end{aligned} \tag{4.10}$$

where σ in PDC can be simply calculated by the summation of radii of two reactants R_A+R_B from

Table 4.1, α is (k_{act}+4$\pi\sigma D$)/4$\pi\sigma^2 D$, and k_{act} is calculated by using equation (4.3) and (4.6). The function $W(x,y)$ is defined as:

$$W(x,y) \equiv exp(2xy + y^2)erfc(x + y) \tag{4.11}$$

In the same spirit as TDC, the CDF of the reaction as a function of time can be derived by the integration of equation (4.10):

$$\begin{aligned} P_{II}(t|r_0) &= 1 - \int_{\sigma}^{\infty} 4\pi r^2 p_{II}(r,t|r_0)dr \\ &= \frac{\sigma}{r_0}\frac{k_{obs}}{k_{diff}}\left[erfc\left(\frac{\mathrm{b}}{\sqrt{Dt}}\right) - W\left(\frac{b}{\sqrt{Dt}}, a\sqrt{Dt}\right)\right] \end{aligned} \tag{4.12}$$

where the time coefficients a and b are -α and (r_0 - σ)/2, respectively.

The type II reactions and corresponding parameters employed in the current study are shown in Table 4.5.

Table 4.5: Type II reactions, corresponding observed reaction rate constant (k_{obs}), reaction radius (σ), diffusion-encounter rate constant (k_{diff}), activation rate constant (k_{act}), probability of contact reaction ($P_{Contact}$), and α value (Plante and Devroye, 2017).

Reaction	k_{obs} ($M^{-1}s^{-1}$)	σ (nm)	k_{diff} ($M^{-1}s^{-1}$)	k_{act} ($M^{-1}s^{-1}$)	$P_{Contact}$ (%)*	α (nm^{-1})
$H^{\cdot} + {}^{\cdot}OH \rightarrow H_2O$	1.55×10^{10}	0.41	2.85×10^{10}	3.39×10^{10}	33.42	5.34
$H^{\cdot} + H_2O_2 \rightarrow {}^{\cdot}OH + H_2O$	3.50×10^{7}	0.40	2.82×10^{10}	3.50×10^{7}	0.05	2.50
$H^{\cdot} + OH^- \rightarrow e^-_{aq} + H_2O$	2.51×10^{7}	0.52	4.84×10^{10}	2.51×10^{7}	0.02	1.92
$H^{\cdot} + O_2 \rightarrow H\dot{O}_2$	2.10×10^{10}	0.36	2.56×10^{10}	1.17×10^{11}	67.44	15.4
$H^{\cdot} + H\dot{O}_2 \rightarrow H_2O_2$	1.00×10^{10}	0.40	2.82×10^{10}	1.55×10^{10}	19.10	3.88
$H^{\cdot} + \dot{O}_2^- \rightarrow HO_2^-$	1.00×10^{10}	0.41	2.71×10^{10}	1.58×10^{10}	19.77	3.86
${}^{\cdot}OH + {}^{\cdot}OH \rightarrow H_2O_2$	5.50×10^{9}	0.44	7.33×10^{9}	2.21×10^{10}	54.98	9.12
${}^{\cdot}OH + H_2O_2 \rightarrow H\dot{O}_2 + H_2O$	2.88×10^{7}	0.43	1.91×10^{10}	3.29×10^{7}	0.08	2.33
${}^{\cdot}OH + H_2 \rightarrow H^{\cdot} + H_2O$	3.28×10^{7}	0.36	1.91×10^{10}	3.29×10^{7}	0.08	2.78
${}^{\cdot}OH + e^-_{aq} \rightarrow OH^-$	2.95×10^{10}	0.72	3.87×10^{10}	1.24×10^{11}	48.57	5.85
${}^{\cdot}OH + OH^- \rightarrow \dot{O}^- + H_2O$	6.30×10^{9}	0.55	3.12×10^{10}	7.89×10^{9}	8.19	2.28
${}^{\cdot}OH + H\dot{O}_2 \rightarrow O_2 + H_2O$	7.90×10^{9}	0.43	1.46×10^{10}	1.72×10^{10}	32.50	5.05
${}^{\cdot}OH + \dot{O}_2^- \rightarrow O_2 + OH^-$	1.07×10^{10}	0.44	1.32×10^{10}	5.74×10^{10}	63.88	12.2
${}^{\cdot}OH + HO_2^- \rightarrow H\dot{O}_2 + OH^-$	8.32×10^{9}	0.47	1.28×10^{10}	2.38×10^{10}	41.96	6.08
${}^{\cdot}OH + \dot{O}^- \rightarrow HO_2^-$	1.00×10^{9}	0.47	1.49×10^{10}	1.07×10^{9}	2.72	2.28
${}^{\cdot}OH + \dot{O}_3^- \rightarrow \dot{O}_2^- + H\dot{O}_2$	8.50×10^{9}	0.42	1.33×10^{10}	2.34×10^{10}	42.21	6.55
$H_2O_2 + e^-_{aq} \rightarrow OH^- + {}^{\cdot}OH$	1.10×10^{10}	0.71	3.87×10^{10}	1.54×10^{10}	10.56	1.97
$H_2O_2 + OH^- \rightarrow HO_2^- + H_2O$	4.71×10^{8}	0.54	3.11×10^{10}	4.78×10^{8}	0.55	1.88
$H_2O_2 + O(^3P) \rightarrow H\dot{O}_2 + {}^{\cdot}OH$	1.60×10^{9}	0.41	1.33×10^{10}	1.82×10^{9}	5.44	2.77
$H_2O_2 + \dot{O}^- \rightarrow H\dot{O}_2 + OH^-$	5.55×10^{8}	0.46	1.50×10^{10}	5.76×10^{8}	1.50	2.26
$H_2 + O(^3P) \rightarrow H^{\cdot} + {}^{\cdot}OH$	4.77×10^{3}	0.34	1.75×10^{10}	4.77×10^{3}	0.00	2.94

$H_2 + \dot{O}^- \rightarrow \dot{H} + OH^-$	1.21×10^8	0.39	2.01×10^{10}	1.22×10^8	2.63	2.58
$e_{aq}^- + O_2 \rightarrow \dot{O}_2^-$	1.74×10^{10}	0.67	3.70×10^{10}	3.28×10^{10}	21.53	2.82
$e_{aq}^- + H\dot{O}_2 \rightarrow HO_2^-$	1.29×10^{10}	0.71	3.87×10^{10}	1.94×10^{10}	12.94	2.11
$OH^- + H\dot{O}_2 \rightarrow \dot{O}_2^- + H_2O$	6.30×10^9	0.54	3.11×10^{10}	7.90×10^9	8.33	2.32
$OH^- + O(^3P) \rightarrow HO_2^-$	4.20×10^8	0.53	2.93×10^{10}	4.26×10^8	0.52	1.91
$O_2 + O(^3P) \rightarrow O_3$	4.00×10^9	0.37	1.23×10^{10}	5.92×10^9	17.71	4.00
$O_2 + \dot{O}^- \rightarrow \dot{O}_3^-$	3.70×10^9	0.42	1.40×10^{10}	5.03×10^9	13.04	3.24
$H\dot{O}_2 + H\dot{O}_2 \rightarrow H_2O_2 + O_2$	9.80×10^5	0.42	7.31×10^9	9.80×10^5	0.01	2.38
$H\dot{O}_2 + \dot{O}_2^- \rightarrow HO_2^- + O_2$	9.70×10^7	0.43	1.32×10^{10}	9.77×10^7	0.30	2.34
$HO_2^- + O(^3P) \rightarrow \dot{O}_2^- + \dot{}OH$	5.30×10^9	0.45	1.16×10^{10}	9.77×10^9	25.24	4.10

* The probabilities of contact reactions will be described in section 4.1.4.

As described before for reactions between charged particles, the consideration of electrostatic potential should be taken into account. We use equation (4.7) combined with equation (4.12) (Plante, 2011) :

$$P_{IV}(t|r_0) = \frac{\sigma_{eff}}{r_{eff}} \frac{k_{obs}}{k_{diff}} \left[erfc\left(\frac{b}{\sqrt{Dt}}\right) - W\left(\frac{b}{\sqrt{Dt}}, a\sqrt{Dt}\right)\right] \quad (4.13)$$

where the time coefficients a and b are:

$$a = \frac{4\sigma^2\alpha}{r_c^2 D} sinh^2\left(\frac{r_c}{2\sigma}\right) \text{ and } b = \frac{r_c}{4}\left[coth\left(\frac{r_c}{2r_0}\right) - coth\left(\frac{r_c}{2\sigma}\right)\right] \quad (4.14)$$

α in equation (4.14) is:

$$\alpha = \frac{k_{act}}{4\pi\sigma^2} \exp\left(r_c/\sigma\right) + \frac{r_c D}{\sigma^2(1 - exp(-r_c/\sigma))} \quad (4.15)$$

Table 4.6 shows the chemical reactions of type IV, and the parameters used in this study.

Table 4.6: Type IV reactions, corresponding observed reaction rate constant (k_{obs}), reaction radius (σ), effective reaction radius (σ_{eff}), diffusion-encounter rate constant (k_{diff}), activation rate constant (k_{act}), and probability of contact reaction ($P_{Contact}$) (Plante and Devroye, 2017).

Reaction	k_{obs} ($M^{-1}s^{-1}$)	σ (nm)	σ_{eff} (nm)	k_{diff} ($M^{-1}s^{-1}$)	k_{act} ($M^{-1}s^{-1}$)	$P_{Contact}$ (%)*
$e_{aq}^- + H_3O^+ \rightarrow H^{\cdot} + H_2O$	2.11×10^{10}	0.75	1.16	1.26×10^{11}	2.53×10^{10}	3.97
$e_{aq}^- + O_2^{\cdot -} + 2H_2O \rightarrow H_2O_2 + 2OH^-$	1.29×10^{10}	0.72	0.42	2.13×10^{10}	3.28×10^{10}	39.06
$e_{aq}^- + HO_2^- \rightarrow O^{\cdot -} + OH^-$	3.51×10^{9}	0.75	0.45	2.15×10^{10}	4.20×10^{9}	7.25
$e_{aq}^- + O^{\cdot -} \rightarrow 2OH^-$	2.31×10^{10}	0.75	0.45	2.35×10^{10}	1.33×10^{12}	95.77
$H_3O^+ + O_2^{\cdot -} \rightarrow HO_2^{\cdot}$	4.78×10^{10}	0.47	0.91	7.73×10^{10}	1.25×10^{11}	28.64
$H_3O^+ + HO_2^- \rightarrow H_2O_2 + H_2O$	5.00×10^{10}	0.50	0.94	7.70×10^{10}	1.43×10^{11}	31.04
$H_3O^+ + O^{\cdot -} \rightarrow {}^{\cdot}OH + H_2O$	4.78×10^{10}	0.50	0.94	8.12×10^{10}	1.16×10^{11}	25.77
$O_2^{\cdot -} + O^{\cdot -} + H_2O \rightarrow O_2 + 2OH^-$	6.00×10^{8}	0.47	0.20	5.71×10^{9}	6.70×10^{8}	6.57
$HO_2^- + O^{\cdot -} \rightarrow O_2^{\cdot -} + OH^-$	3.50×10^{8}	0.50	0,23	5.82×10^{9}	3.72×10^{8}	3.52
$O^{\cdot -} + O^{\cdot -} + 2H_2O \rightarrow H_2O_2 + 2OH^-$	1.00×10^{8}	0.50	0.23	3.43×10^{9}	1.03×10^{8}	1.69
$O^{\cdot -} + O_3^{\cdot -} \rightarrow 2O_2^{\cdot -}$	7.00×10^{8}	0.45	0.18	5.59×10^{9}	8.00×10^{8}	8.14

* The probabilities of contact reactions will be described in section 4.1.4.

4.1.3. First-order reactions and background reactions (Type VI reactions)

First-order reactions are the simplest case in which the reaction rate is determined by the concentration of one species as in Table 4.7. The changes of the concentration as a function of time $A(t)$ can be derived as:

$$A(t) = A_0 e^{-k_{sca}t} \tag{4.16}$$

where A_0 is the initial concentration and k_{sca} is the scavenging power determining the reaction rate.

Table 4.7: First-order reaction and scavenging power (k_{sca}) (Plante and Devroye, 2017).

	k_{sca} (s^{-1})
$\dot{O_3^-} \rightarrow \dot{O^-} + O_2$	2.66×10^3

Chemical species can react with background molecules in the solvent, however simulating the entire molecules of background is a huge burden. Thus, we need to simplify the background reactions with the following assumptions:

1) The background molecules are homogeneously concentrated in the medium.
2) The number of background molecules is large enough, thus the reactions with background molecules do not affect their concentration.
3) The survival fraction follows exponential form as first-order reaction given by equation (4.16).

The concentrations of the background molecules used in this study for H_2O, H_3O^+, and OH^- in pure liquid water are of 5.53×10^1, 9.9×10^{-8}, and 9.9×10^{-8} M at 25 °C, respectively (Plante, 2011). With those values and the above assumptions, the scavenging power or capacity k_{sca} is calculated as the multiplication of k_{obs} by the concentration of background molecules. We used reaction rate constants of same diffusion-controlled reactions of type I-V which use H_3O^+ or OH^-, and additional reaction rate constants with H_2O molecules, as shown in Table 4.8.

Table 4.8: Background reactions, corresponding reaction rate constant (k_{obs}) and scavenging power (k_{sca}) (Plante and Devroye, 2017).

Reaction	k_{obs} ($M^{-1}s^{-1}$)	k_{sca} (s^{-1})
$H\dot{O}_2 + H_2O \rightarrow H_3O^+ + \dot{O}_2^-$	1.29×10^4	7.15×10^5
$H^{\cdot} + H_2O \rightarrow e_{aq}^- + H_3O^+$	1.07×10^{-1}	5.94×10^0
$e_{aq}^- + H_2O \rightarrow H^{\cdot} + OH^-$	2.86×10^{-1}	1.58×10^1
$\dot{O}_2^- + H_2O \rightarrow H\dot{O}_2 + OH^-$	2.71×10^{-3}	1.50×10^{-1}
$HO_2^- + H_2O \rightarrow H_2O_2 + OH^-$	2.46×10^4	1.36×10^6
$O(^3P) + H_2O \rightarrow 2^{\cdot}OH$	1.81×10^1	1.00×10^3
$O^- + H_2O \rightarrow {}^{\cdot}OH + OH^-$	2.46×10^4	1.36×10^6
$e_{aq}^- + H_3O^+ \rightarrow H^{\cdot} + H_2O$	2.11×10^{10}	2.09×10^3
$\dot{O}_2^- + H_3O^+ \rightarrow H\dot{O}_2 + H_2O$	4.78×10^{10}	4.73×10^3
$^*OH^- + H_3O^+ \rightarrow 2H_2O$	1.13×10^{11}	1.12×10^4
$HO_2^- + H_3O^+ \rightarrow H_2O_2 + H_2O$	5.00×10^{10}	4.95×10^3
$O^- + H_3O^+ \rightarrow {}^{\cdot}OH + H_2O$	4.78×10^{10}	4.73×10^3
$\dot{O}_3^- + H_3O^+ \rightarrow {}^{\cdot}OH + O_2 + H_2O$	9.00×10^{10}	8.91×10^3
$H^{\cdot} + OH^- \rightarrow e_{aq}^- + H_2O$	2.52×10^7	2.48×10^0
${}^{\cdot}OH + OH^- \rightarrow \dot{O}^- + H_2O$	6.30×10^9	6.24×10^2
$H_2O_2 + OH^- \rightarrow HO_2^- + H_2O$	4.71×10^8	4.66×10^1
$H\dot{O}_2 + OH^- \rightarrow \dot{O}_2^- + H_2O$	6.30×10^9	6.24×10^2
$O(^3P) + OH^- \rightarrow HO_2^-$	4.20×10^8	4.16×10^1

* Both reactants OH^- and H_3O^+ can be background species.

4.1.4. Contact reactions

RITRACKS defines contact reaction when the distance r is smaller than reaction radius σ because GFDE can't describe that. The contact reaction immediately occurs at the position with the probability of contact reaction $P_{contact}$.

In the case of TDC reactions, $P(\infty|r_0)$, which is the cumulated reaction probability until infinite time, is identical to $P_{contact} = 1$ at the boundary condition (when $r_0=\sigma$), because all collisions lead to chemical reactions.

Besides, the $P_{contact}$ of PDC reactions should take into account the influence of unsuccessful encounters and re-encounters (Goulet and Jay-Gerin, 1992):

$$P_{contact} = \frac{\exp(-r_c/\sigma) - \exp(-r_c/(\sigma + R_s))}{\exp(-r_c/\sigma) - \exp(-r_c/(\sigma + R_s)) - (k_{dif}/k_{act})(1 - \exp(-r_c/\sigma))} \quad (4.17)$$

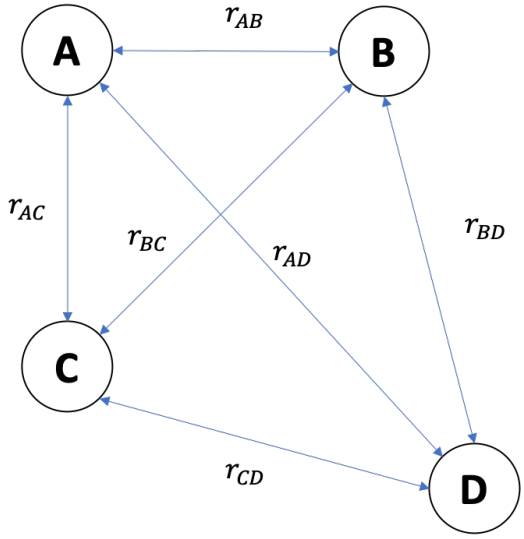
where R_S is the limitation of the separation distance after unsuccessful encounter which possibly makes next re-encounter. This approximate distance between water molecules in RITRACKS (Plante and Devroye, 2017) is taken as 0.3 nm. Note that for TDC with $k_{act} \rightarrow \infty$, $P_{contact}$ converges to 1. For reactions without electrostatic interactions ($r_c \rightarrow 0$), $P_{contact}$ can be simplified as:

$$P_{contact} = \frac{R_S}{R_S - (k_{dif}/k_{act})(\sigma + R_S)} \quad (4.18)$$

P_{contact} of PDC is always less than $P(\infty|\sigma)$ because it is possible that one of the reactants react with other reactants in solvent (Frongillo et al., 1998).

4.2. The Independent Reaction Time technique

The IRT technique simulates the time of a reaction for a reactant pair in a solvent assuming that the pair is at isolation. In other words, for a reactive pair, we don't need to diffuse the molecules and recalculate the reactions between every reactant at every time step prior to react. IRTs between all reactants are calculated based on the reaction type and the distance, and the reactions occur following the order of IRTs. For instance, in Figure 4.1, the earliest reaction A + C is sampled at first. And then, the next reaction might be the next earliest reaction which doesn't use the reactants A and C.



Reaction	Time (ps)	Order of reactions
A+B	63	X
A+C	3	1
A+D	111	X
B+C	192	X
B+D	15	2
C+D	90	X

Figure 4.1: A schematic illustration of an example for IRT sampling (Green et al., 1990).

Figure 4.2 shows the implementation of IRT in Geant4-DNA I developed in this study. Following the existing chemistry interface of Geant4-DNA, reactions are initialized in the chemistry constructor which defines molecules and chemical reactions considered in the chemistry simulation. The definitions of the reactions are set in the *G4DNAMolecularReactionTable* class. During the physical and pre-chemical stages of Geant4-DNA, the initial species at 1 ps are stored in the *G4ITTrackHolder* class.

At the first step of IRT technique, possible reactions between all the initial reactants are stored in the *G4ITReactionSet* class with the corresponding reaction time. The *G4ITReactionSet* class automatically sorts the reactions according to IRT. After that, all the reactions occur in order of time starting from the shortest time. In order to avoid duplicated sampling, all the possible reactions

for both reactants in reaction lists are removed after the reaction occurred between them. If there are products, the possible reactions of the products are also inserted in the reaction list. At last, the *G4MoleculeCounter* class counts the number of chemical species and calculates G-values as a function of time obtained from the chemistry simulations. A brief description of Geant4-DNA classes is shown in Table 4.9.

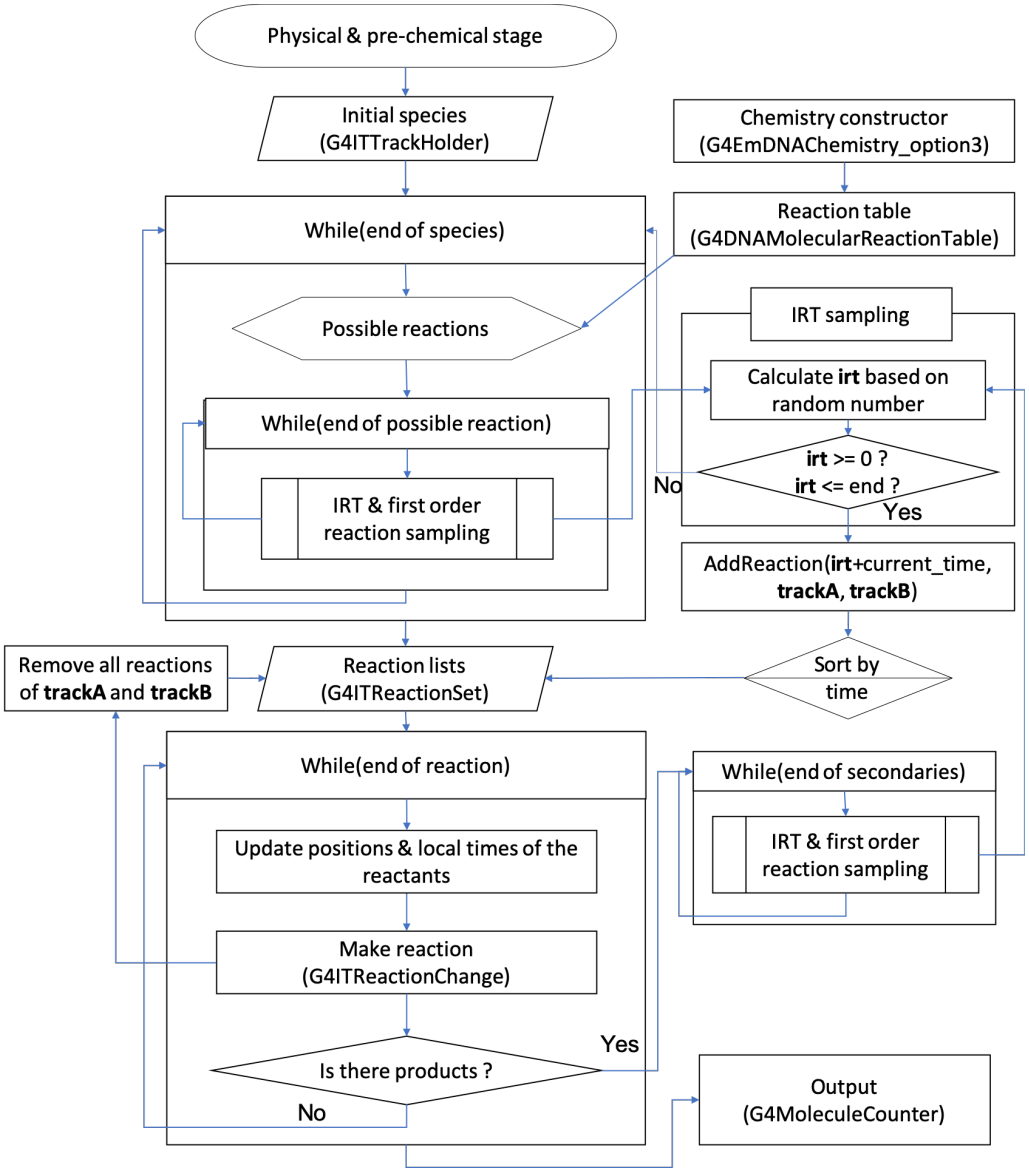


Figure 4.2: Flowchart of IRT implemented in Geant4-DNA.

Table 4.9: Geant4-DNA classes used in chemistry simulation.

G4DNA classes	Description
Chemistry constructor (G4EmDNAChemistry_option3)	Defines molecules, chemical reactions, and dissociation scheme. It is equivalent to physics constructor.
G4DNAMolecularReactionTable	Contains a table for the chemical reactions and parameters defined in the chemistry constructor.
G4ITTrackHolder	Stores chemical species generated during pre-chemical and chemical stages.
G4ITReactionSet	Stores upcoming reactions.
G4ITReactionChange	A class for updating reactants during a chemical reaction. This class is equivalent to the G4ParticleChange class used in the physical stage simulation.
G4MoleculeCounter	Counts the number of molecular species.

4.2.1. Search range

MCTS tools usually limit the search range of a reactive molecule with its neighbors based on a confidence level (Karamitros et al., 2011) in order to reduce unnecessary calculation time. The probability $P(r_0 \leq \sigma)$ that the separation distance r_0 between two reactants is smaller than the reaction radius σ can be expressed based on the 1-dimensional Smoluchowski diffusion equation at time t as:

$$P(r_0 \leq \sigma) = \mathrm{erf}\left(\frac{d_{95}}{2 \times \sqrt{Dt}}\right) \tag{4.19}$$

where d_{95} is the distance for which 95% of the distance distribution is shorter than d_{95} (P=95% confidence level) with diffusion coefficient D. $erf(\sqrt{2})$ can be approximated by 0.95, therefore, d_{95} can be obtained from:

$$d_{95} = 2\sqrt{2Dt} \tag{4.20}$$

If the initial separation between the two reactants is r_0 at initial time, then the distance d_Δ at which they can react at time t_Δ is:

$$d_\Delta = r_0 - \sigma = 2\sqrt{2D_A t_\Delta} + 2\sqrt{2D_B t_\Delta} \tag{4.21}$$

The initial distance between two reactants at which they can possibly react with each other until end-time of chemistry, 1 µs in our case (note that this maximum value can be changed by the user), depends on the reaction radius σ and the diffusion coefficient D. Therefore, the simulation should conservatively search the cutoff range σ_{off} calculated with the maximum reaction rate σ_{max} and maximum diffusion coefficient D_{max} (1.45 nm and 9.46×10^9 nm^2/s, respectively in this study) as:

$$\sigma_{off} = \sigma_{max} + 4\sqrt{2 \times D_{max} \times 1\ \mu s} \tag{4.22}$$

As a simplification of the search algorithm, a water phantom is virtually voxelized based on σ_{off} as shown in Figure 4.3. Thus, the reactions with the molecules in the 27 blue voxels (in 3-dimensions) are searched in order to find the molecule highlighted as a red star.

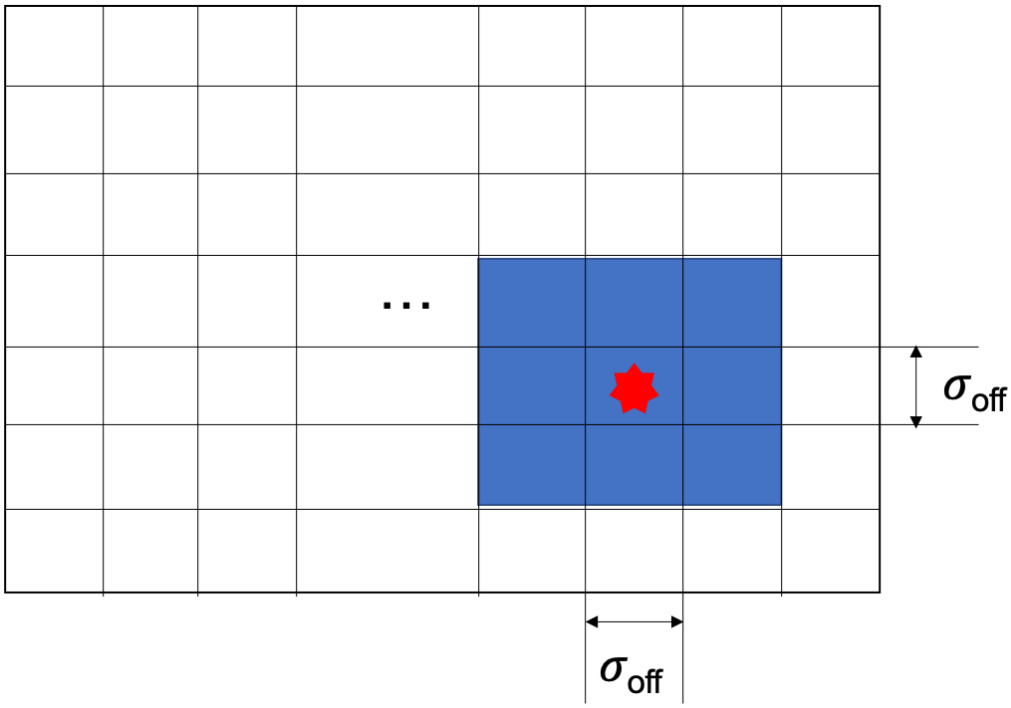


Figure 4.3: A schematic illustration of a voxelized water phantom. The red star indicates a target molecule, and the blue voxels represent regions searched for the red star molecule.

4.2.2. Sampling of diffusion-controlled reactions

In order to sample the chemical reactions by time, the CDF of reaction as derived in equation (4.5) and equation (4.7) is inverted as an equation of IRT for type I and type III reactions, respectively:

$$IRT_I = \frac{1}{4D}\left[\frac{r_0-\sigma}{erfc^{-1}\left(\frac{r_0}{\sigma}U\right)}\right]^2 \text{ and } IRT_{III} = \frac{1}{4D}\left[\frac{r_{eff}-\sigma_{eff}}{erfc^{-1}\left(\frac{r_{eff}}{\sigma_{eff}}U\right)}\right]^2 \quad (4.23)$$

A random number U is uniformly sampled, and the IRT is calculated from this random number and a look-up-table for calculating inverse error function $erfc^{-1}(x)$, which is impossible to calculate analytically.

There is an alternative way using inverse gamma distribution to avoid the calculation of inverse error function. The probability density function (PDF) of reaction probability as a function of time dP/dt can be derived from the CDF of the reaction, equation (4.4) as:

$$\frac{dP_{TDC}}{dt} = \frac{\sigma}{r_0} \times \frac{r_0-\sigma}{2\sqrt{D}\sqrt{\pi}} t^{-3/2} \exp\left[-\frac{(r_0-\sigma)^2}{4Dt}\right] \quad (4.24)$$

This PDF corresponds to the inverse gamma distribution of α=1/2, β=$(r_0-\sigma)^2/4D$. In the case of type III, the same expression replacing r_0 and σ by $r_{0\text{eff}}$ and σ_{eff} is used. Unfortunately, CLHEP doesn't provide the sampling function for inverse gamma distribution with an alpha value smaller than 1, thus in this case, MATLAB is employed in order to evaluate the plausibility and the

calculation efficiency of the gamma sampling method (MathWorks, 2005).

In the case of PDC, analytically resolving equation (4.12) and equation (4.13) as an equation of IRT is impossible. Thus, we implemented the rejection method proposed by Plante and Devroye (2017), and the Brent's method (Brent, 1971) in order to numerically solve the equations, and to calculate the IRT.

The Brent's method is a numerical approach to find a solution x_0 for $f(x_0)=0$ in very complex equations (Brent, 1971). The initial interval gets closer during the iterations using bisection method, secant method, and inverse quadratic interpolation until the interval is less than the pre-defined tolerance. Equation (4.12) and (4.13) which represent the CDFs of the type II and IV reactions as a function of time are subtracted by a random number U, and the Brent's algorithm starts to find the time t_0 of $P_{PDC}(t_0) - U = 0$. The initial interval between the minimum time bin (0.1 ps) and end time (1 us) with 1 fs tolerance and 10^5 maximum number of iterations is used in this study.

The acceptance-rejection method is usually employed to generate a random value from a distribution (Casella et al., 2004). At first, arbitrary PDF $q(x)$, which is fully bounding the target distribution $p(x)$, is defined. After that, a random number is generated and the corresponding x_0 is calculated. Again, another random number U in a range of $[0, q(x_0)]$ is generated. If U is less than $p(x_0)$, the accepted x_0 is sampled. If U is larger than $p(x_0)$, the value x_0 is rejected and the algorithm continues the iteration.

To determine proper $q(x)$ is very important for acceptance-rejection method. Too large $q(x)$ leads to high rejection ratio and calculation time, too small $q(x)$ will distort the sampling result. Plante et al. (2013), Plante and Devroye (2017) already proposed $q(x)$ in their previous paper. The PDF of the reaction time for PDC reactions can be derived by differentiation of CDF in equation (4.12) and (4.13):

$$\frac{dP_{PDC}}{dt} = P(\infty|r_0) \times \frac{Da}{\sqrt{\pi}\sqrt{Dt}} \exp\left[-\left(\frac{b}{\sqrt{Dt}}\right)^2\right]\left[1 - a\sqrt{\pi}\sqrt{Dt} \times erfcx\left(\frac{b}{\sqrt{Dt}} - a\sqrt{Dt}\right)\right] \quad (4.25)$$

In this equation, the latter part is always less than 1. Thus, we could say that dP/dt is fully bounded by $q(x) = P(\infty|r_0) \times Da/\sqrt{\pi Dt}$. In order to improve the calculation efficiency, Plante and Devroye (2017) suggested the pre-determined $q(x)$ as:

$$q(x) = \begin{cases} C/x^{1/2} & x < 2b/a \\ CM/x^{3/2} & x > 2b/a \end{cases} \quad (4.26)$$

where $x=Dt$, C is a constant, $M = max(1/a^2, 3b/a)$.

With this $q(x)$ and the acceptance-rejection algorithm, the random sampled IRT could be generated based on the PDF of equation (4.25) as shown in Figure 4.4.

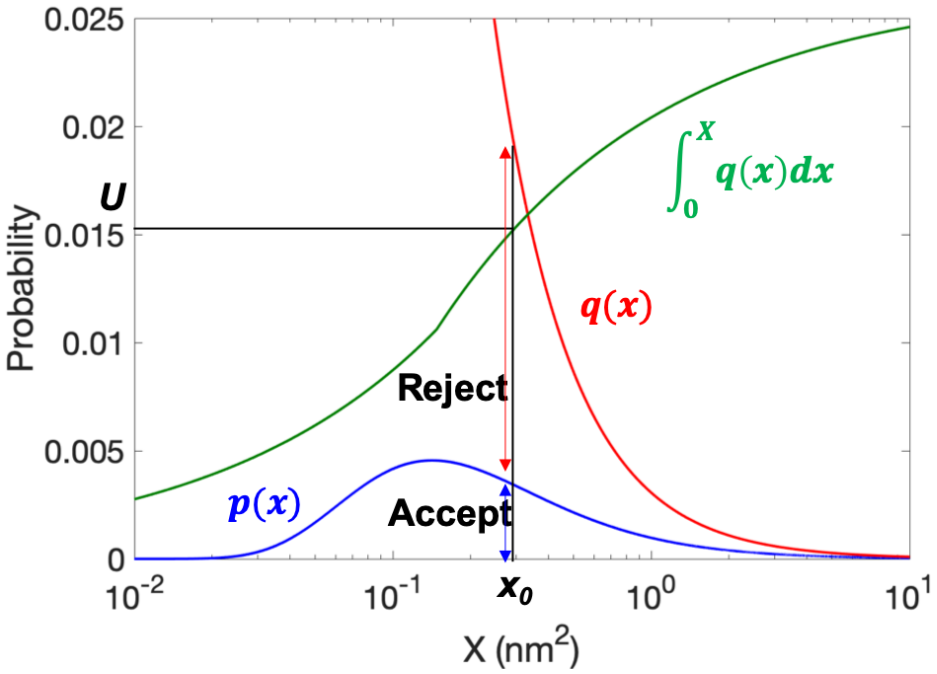


Figure 4.4: A schematic outline of acceptance-rejection method for $H^{\cdot} + {}^{\cdot}OH \rightarrow H_2O$ type II reaction when r_0 = 1 nm. The blue, red, green curves represent target distribution $p(x)$, bounding distribution $q(x)$, and CDF of $q(x)$, respectively.

A validation study with these sampling methods has been performed using the IRT code with a specific chemical reaction and an initial distance r_0 for each reaction type (e.g. •H + OH• -> H_2O reaction with the initial distance of 1 nm for the validation of sampling type II reactions). The calculation efficiency of each sampling method with 10^9 maximum iterations has been evaluated, and the sampling results are compared with theoretical results, as will be shown in section 4.5.

4.2.3. Sampling of type VI reactions and contact reactions

For both of first-order reactions and background reactions, IRT of type VI reactions can be calculated by simply inverting equation (4.16):

$$IRT = -\frac{\ln(1 - P_{VI})}{k_{sca}} \tag{4.27}$$

The contact reaction only occurs when two reactants encounter each other in the distance of reaction radius σ (or σ_{eff}). Basically, this means that the reaction occurs immediately as the IRT is equal to 0.

4.2.4. Reaction site and position of secondaries

The determination of reaction site $\vec{r_r}$ is typically based on the diffusion coefficients D and the positions r of both reactants A and B:

$$r_r = \frac{\sqrt{D_B}}{\sqrt{D_A}+\sqrt{D_B}} r_A + \frac{\sqrt{D_A}}{\sqrt{D_A}+\sqrt{D_B}} r_B \tag{4.28}$$

Equation (4.27) is reasonable when the time step Δt is very small like SBS method, however in the IRT method, the positions of r_A and r_B need to be updated for the reaction time because the IRT method doesn't consider diffusion of reactants during the simulation. There is an alternative approach developed for the IRT method (Clifford et al., 1986) with the assumption that all the molecular species freely diffuse for the time step Δt based on normal distribution as:

$$r(t) = r(0) + N_3(0, \sigma^2) \tag{4.29}$$

where $r(x)$ is the position of the particle at time x, and $N_3(0, \sigma^2)$ is a 3-dimensional random vector with mean is 0 and variance $\sigma^2=2D\Delta t$.

For the calculation of the positions of reactants at the reaction time, the separation vector S_1 and the diffusion vector S_2 are defined as:

$$S_1 \equiv r_A - r_B;\ S_2 \equiv r_A + b r_B \tag{4.30}$$

where the constant b is $2D_A\Delta t/2D_B\Delta t$ for the statistical independence between S_1 and S_2, in other words, the covariance is 0.

The magnitude of the separation vector $S_1(t)$ at the reaction time should be equal to the reaction radius σ. Thus, the random direction of $S_1(t)$ described with angle θ and ϕ in a spherical coordinate system is sampled. θ can be sampled by a uniform random number because of the cylindrical symmetry at the axis of r_0, and ϕ can be sampled by:

$$\phi = \cos^{-1}\{1 + \frac{1}{\alpha}\ln[1 - U\{1 - \exp(-2\alpha)\}]\} \tag{4.31}$$

where U is a uniform random number, and $\alpha=\sigma r_0/2D\Delta t$.

The S_2 vector independently diffuses regardless of S_1, because their covariance is 0. It can be described with 3-dimensional normal distribution as:

$$S_2 = r_A + b r_B + N_3\left[0, \left(2D_A t + \frac{D_A}{D_B}\right)\right] \tag{4.32}$$

Based on the sampled S_1, S_2, and equation (4.29), the positions of the reactants at reaction time can be described as:

$$r_A(t) = \frac{D_A S_1 + D_B S_2}{D_A + D_B};\ r_B(t) = \frac{D_B(S_2 - S_1)}{D_A + D_B} \tag{4.33}$$

Basically, the secondary species should be generated in the reaction site r_r of equation (4.28), however, this position should be rearranged when the products are more than 2 in order to avoid the immediate recombination of the products by contact reaction. In this study, the secondaries are located at $r_A(t)$ and $r_B(t)$ when the number of products is two, and located at $r_A(t)$, $r_B(t)$, $r_r(t)$ when three products are generated.

4.3. Implementation of IRT in the chemistry framework of Geant4-DNA

In order to synchronize the IRT method with the chemistry framework of Geant4-DNA, the new developments have to be incorporated in Geant4 kernel classes. Figure 4.5 shows the structure of Geant4 kernel classes for chemistry simulation, including the new classes of the IRT method.

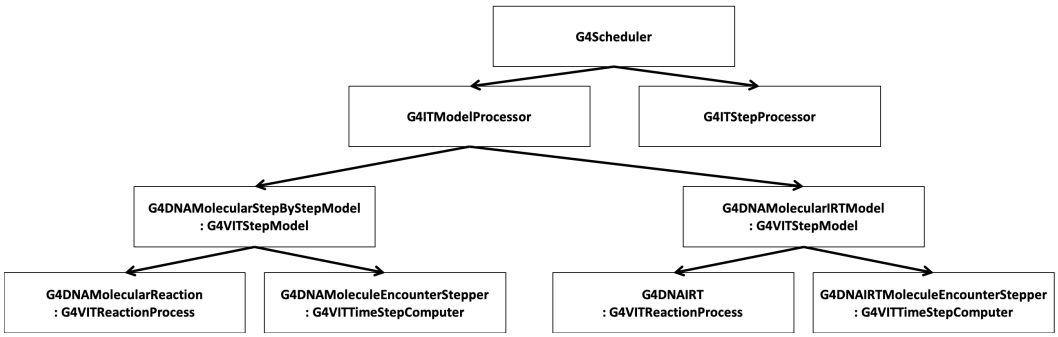


Figure 4.5 A schematic illustration of Geant4-DNA kernel classes for chemistry simulation.

The *G4Scheduler* class is in charge of managing the time evolution during the simulation of the chemical stage. This class calls two processor classes for different types of reaction. At first, the interactions with medium, such as molecular dissociation processes of water molecules during the pre-chemical stage and Brownian transportation, are described in the *G4ITStepProcessor* class. One thing should be noted: the IRT method doesn't need to take into account the diffusion process of species (Brownian transportation) because the reaction time is calculated based on the initial separation distance only. Moreover, *G4ITStepProcessor* takes much time to scan every molecular species. Thus, the class is deactivated after the dissociation processes are finished in order to avoid unnecessary waste of computational time.

Besides, the chemical reactions between reactants are controlled by the *G4ITModelProcessor* class. The *G4DNAMolecularStepByStepModel* class or the *G4DNAMolecularIRTModel* class, which are inherited from the *G4VITStepModel* class, are called in the chemistry constructor in order to be employed in the *G4ITModelProcessor* class. Thus, users can easily switch between the SBS and the IRT method in the chemistry constructor. These classes determine the methodologies used to calculate time steps (*G4VITTimeStepComputer*) and to find or make chemical reactions (*G4VITReactionProcess*). In this study, *G4DNAIRTMoleculeEncounterStepper* and *G4DNAIRT* classes are developed by mimicking *G4DNAMoleculeEncounterStepper* and *G4DNAMolecularReaction* classes, respectively.

4.4. Radiochemical yield simulation in water

In order to validate our implementation, water radiolysis simulations are performed to obtain G-values as a function of time and LET. For electron simulations, a large enough water phantom ($1\times1\times1$ km^3) is used, as well as the same electron energies and same primary killer technique for selecting a small segment of the entire physical track in order to simulate large enough numbers of tracks in reasonable calculation times, as proposed by previous studies for SBS approach (Ramos-Mendez et al., 2018). Protons and alphas are generated at the edge of a $5\times5\times5$ cm^3 water phantom as same condition with typical water radiolysis studies (without primary killer technique). In the physical and pre-chemical stages, we combine the Geant4-DNA physics constructor "option 2" (consisting of the default inelastic models of Geant4-DNA and the elastic scattering model developed during this thesis using Dirac partial wave analysis (Shin et al., 2018) for electrons) with the Meesungnoen electron thermalization model (Meesungnoen et al., 2002) and with the dissociation scheme including displacement of fragments proposed by Kreipl et al. (2009), Shin et al. (2019), as we previously explained (Shin et al., 2019).

Sets of calculation and experimental data presented by Buxton (1972), Wolff et al. (1973), Draganić and Draganić (1975), Burns et al. (1981), Sumiyoshi and Katayama (1982), LaVerne and Pimblott (1991), Belloni et al. (1983), Elliot et al. (1993), Tomita et al. (1997), Pimblott and LaVerne (1997), Bartels et al. (2000), Jay-Gerin and Ferradini (2000), LaVerne (2000), LaVerne et al. (2005), Muroya et al. (2005), Wang et al. (2018) are used for comparison with the simulated G-values versus time. For G-values versus LET, calculation and experimental data obtained by Schwarz et al. (1959), Anderson and Hart (1961), Naleway et al. (1979), Burns and Sims (1981), Pastina and LaVerne (1999), Wasselin-Trupin et al. (2002), Shin et al. (2019) are used. It should be noted that most of the experimental data measured the G-values under solvent concentration arbitrary controlled, whereas the simulations are performed for pure liquid water.

4.5. Validation of the implementation

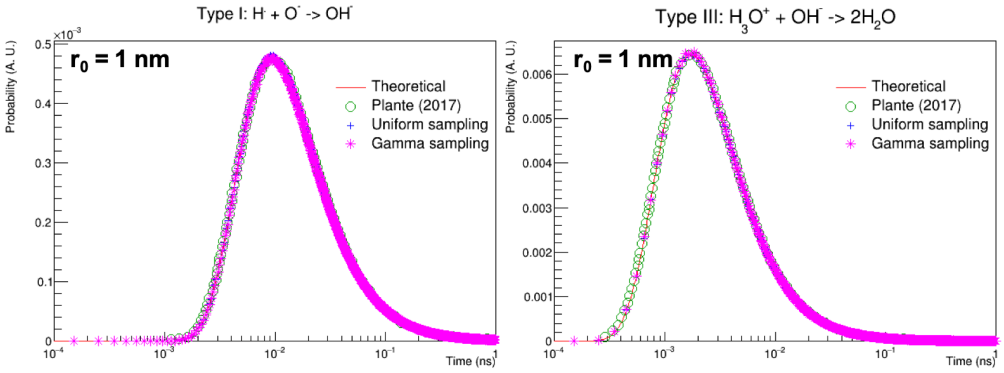


Figure 4.6: The PDFs of TDC reactions as a function of IRT according to the different sampling methods.

Both of the results obtained using the uniform and the gamma sampling method for TDC reactions are identical compared to theoretical model and to the sampling results of Plante and Devroye (2017) as shown in Figure 4.6. The calculation time with uniform sampling is 10 times faster than that with gamma sampling using MATLAB. The performance of the calculation probably depends on the function assessing $erfc^{-1}(x)$ and inverse gamma distribution, however, we decided to keep the uniform sampling method as default.

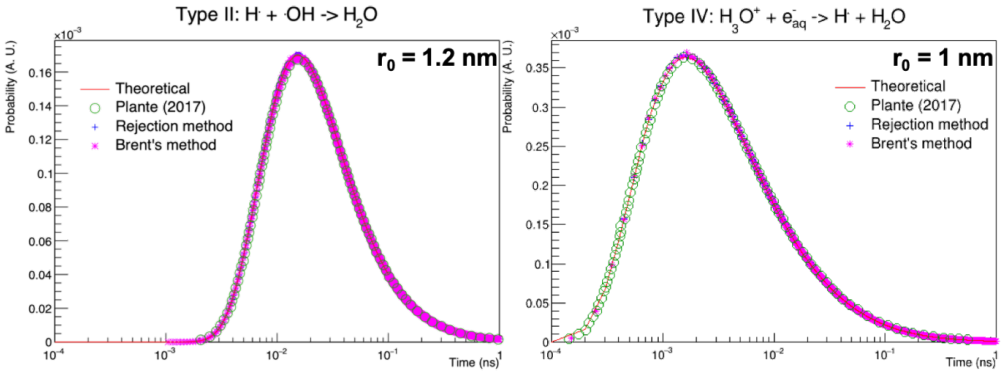


Figure 4.7: The PDFs of PDC reactions as a function of IRT according to the different sampling methods.

In the case of PDC reactions, the Brent's method and the rejection method also give identical PDFs as shown in Figure 4.7. The calculation efficiency could be changed by modifying the tolerance and the maximum number of iterations. However, the Brent's method takes more calculation time than the rejection method, up to 5 times. Thus, the rejection method is taken as a default sampling method for PDC reactions.

Table 4.10: Particle energies, number of particles, LET values, standard deviation of LET values, and calculation times obtained in this study.

Electron energy (keV)	**N of electrons**	**LET (keV/um)**	**σ of LET (keV/um)**	**Calculation time (CPU-hr)**
2	30,000	9.24	5.79E-4	0.06
3.5	30,000	5.97	5.94E-4	0.09
7.5	30,000	3.10	1.86E-4	0.16
12.5	5,000	2.11	9.51E-4	0.07
30	3,000	1.01	3.49E-4	0.06
80	3,000	0.61	3.80E-3	0.08
1,000	3,000	0.16	4.58E-5	0.10
Proton energy (keV)	**N of protons**	**LET (keV/um)**	**σ of LET (keV/um)**	**Calculation time (CPU-hr)**
500	300	47.66	5.43E-3	39.71
700	300	36.77	5.08E-3	19.30
800	300	32.93	4.90E-3	13.48
1,200	300	23.89	4.98E-3	5.56
1,500	300	20.19	4.90E-3	3.20
2,000	300	16.24	5.12E-3	1.72
2,500	300	13.64	4.65E-3	1.02
3,000	300	12.09	4.59E-3	0.67
4,000	300	9.56	4.90E-3	0.38
5,000	300	8.01	5.01E-3	0.24
7,500	300	5.73	4.31E-3	0.13
10,000	300	4.61	4.46E-3	0.08
12,000	3,000	4.02	4.76E-4	0.65
15,000	3,000	3.38	4.90E-4	0.48
Alpha energy (keV)	**N of alphas**	**LET (keV/um)**	**σ of LET (keV/um)**	**Calculation time (CPU-hr)**
1,000	30	171.25	5.46E-2	194.09
2,000	30	177.85	1.05E-1	189.14
4,000	30	104.07	6.40E-2	36.72
8,000	30	60.48	1.09E-1	5.94
20,000	30	30.06	9.33E-2	0.69

The number of particles and the particle energies are based on the previous study performed by Ramos-Mendez et al. (2018) with SBS method for a fair comparison as shown in Table 4.10. The calculation time is exponentially increased according to the LET of the particle, due to not only the number of chemical reactions but also to the number of physical interactions. The escape of the particle from the water phantom also affects the calculation time because the range of the protons and alphas are longer than the phantom volume of 5×5×5 μm^3 when the energies are more than 300 and 800 keV, respectively (Seltzer et al., 2014).

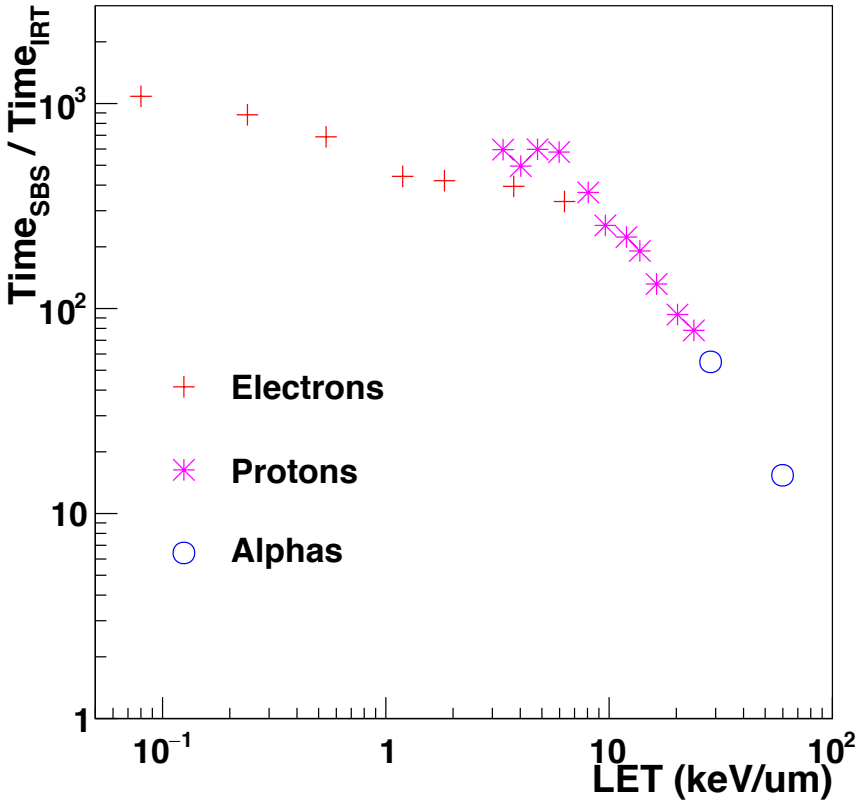


Figure 4.8: The improvement of calculation efficiency as a function of LET comparing to SBS method. The simulation time of chemical stage including initialization time is measured.

The IRT method can reduce the simulation time up to 1,000 times compared to SBS method for low-LET simulations such as 1 MeV electrons as shown in Figure 4.8. Then, the ratio of the calculation efficiency between SBS and IRT decreases as a function of LET even though the ratio only includes chemical stage simulation. That means, the IRT method is more sensitive to the spatial density of chemical species because this method calculates all the possible reactions in the search range. Nevertheless, the simulation time for chemical stage is remarkably reduced using the IRT method.

4.6. Results for G-values versus time

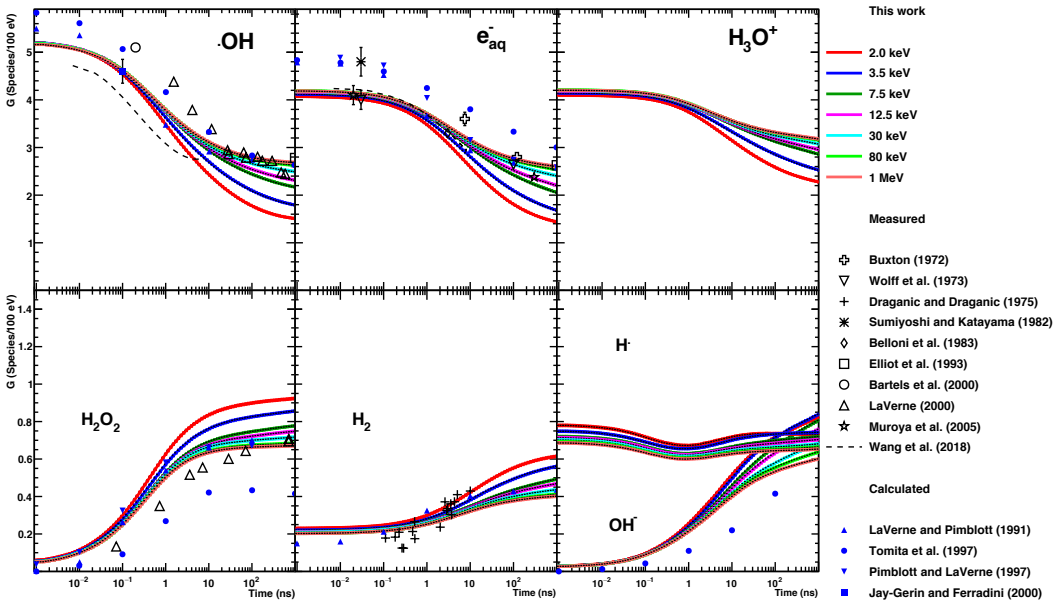


Figure 4.9: G-values as a function of time for various energies of incident electrons.

Figure 4.9 shows the G-values as a function of time according to different electron energies. The case of 2.0 keV electrons which has the highest LET shows the strongest time evolution, in other words, this is the configuration for which the number of chemical reactions is the largest. The major species ($^{\bullet}OH$, e_{aq}^-, H_3O^+) undergoing chemical reactions decrease according to time unlike the minor species (H_2O_2, H_2) produced by chemical reactions. In any case, the G-values are saturated when the reactants have sufficiently diffused. The simulation results of low LET radiation especially at 80 keV and 1 MeV are acceptable comparing with experimental data performed for low LET radiation (0.18-0.3 keV/µm). Especially, the G-values of solvated electrons e_{aq}^- calculated in this study match well recent experimental data such as Muroya et al. (2005), Wang et al. (2018). However, the calculated results in the other studies are higher than our results. We could assume that these disagreements come from the modeling of the pre-chemical stage and the initial radiochemical yields which are difficult to measure.

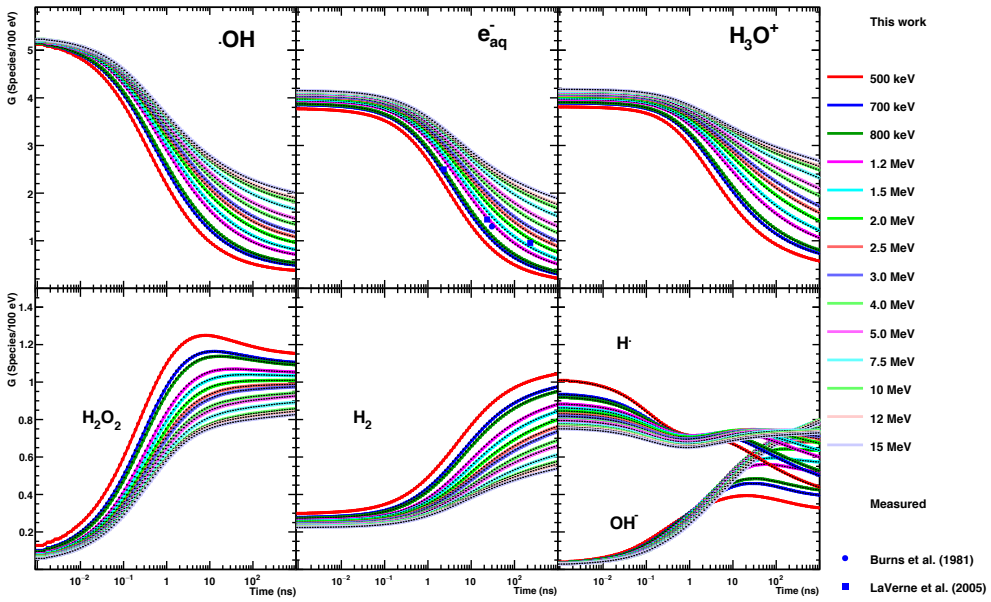


Figure 4.10: G-values as a function of time with various energies of protons.

The G-value results for protons also show good agreement with the experimental data of Burns et al. (1981), LaVerne et al. (2005) for 2 and 3 MeV protons respectively, as shown in Figure 4.10. Basically, the results show similar tendency with electron cases rapidly changing for high LET particles. Interestingly, we observe that the initial G-values of e_{aq}^- and H_3O^+ are affected by LET values due to the electron-hole recombination process controlled by the *G4DNAElectronHoleRecombination* class. This process occurs at very early pre-chemical stage between ionized water H_2O^+ and solvated electrons e_{aq}^-. The survival fraction of this process is provided by Lu et al. (1989) at femtosecond scale, and the probability of the recombination process could be assessed by one minus the survival fraction:

$$P_{e_{aq}^- + H_2O^+} = 1 - \exp\left(-\frac{r_c}{r_0}\right) \tag{4.34}$$

This recombination produces H_2O^* molecules which dissociate as •OH radicals and hydrogen products, and the number of solvated electrons and H_2O^+ will decrease (H_2O^+ is dissociated as H_3O^+ and •OH radicals). Based on equation (4.34), the number of electron-hole recombinations depends on the distance between H_2O^+ and e_{aq}^-, thus the influence of the recombination process is remarkable for high LET.

In addition, unlike in the electron cases for which the G-values of H_2O_2 are generally saturated after few nanoseconds, we observe that the G-values are distorted at high LET. We could guess that, because the number of residual •OH radicals which generate H_2O_2 via the $^{\cdot}OH + {}^{\cdot}OH \rightarrow H_2O_2$ reaction decreases rapidly before they diffuse enough to escape, due to the fast chemical kinetics compared to the electron case.

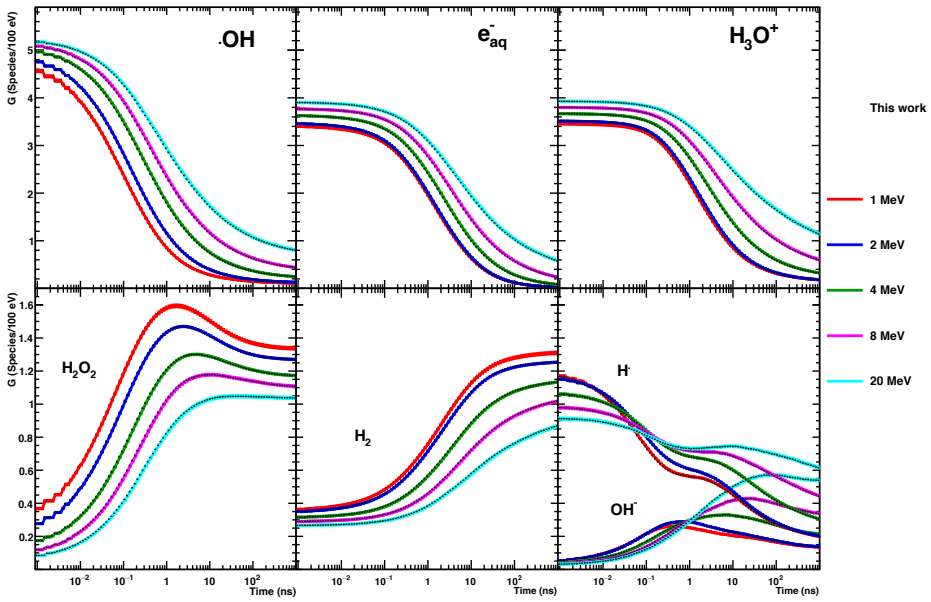


Figure 4.11: G-values as a function of time with various energies of alpha particles.

Figure 4.11 shows the G-value results for alpha particles. It shows similar tendency with that of high LET proton case, such as strong time evolution and initial yields of e_{aq}^- and H_3O^+. Most of the major species are rapidly spent with faster chemical kinetics due to the dense distribution of molecular species. Thus, the distortions of minor species such as H_2O_2, $H^\bullet$ radical, and OH^- are getting significant according to the LET due to the same reason as for the high LET proton case. Unfortunately, for full-validation, experimental data for alpha particles has not been found in the literature.

4.7. Results for G-values versus LET

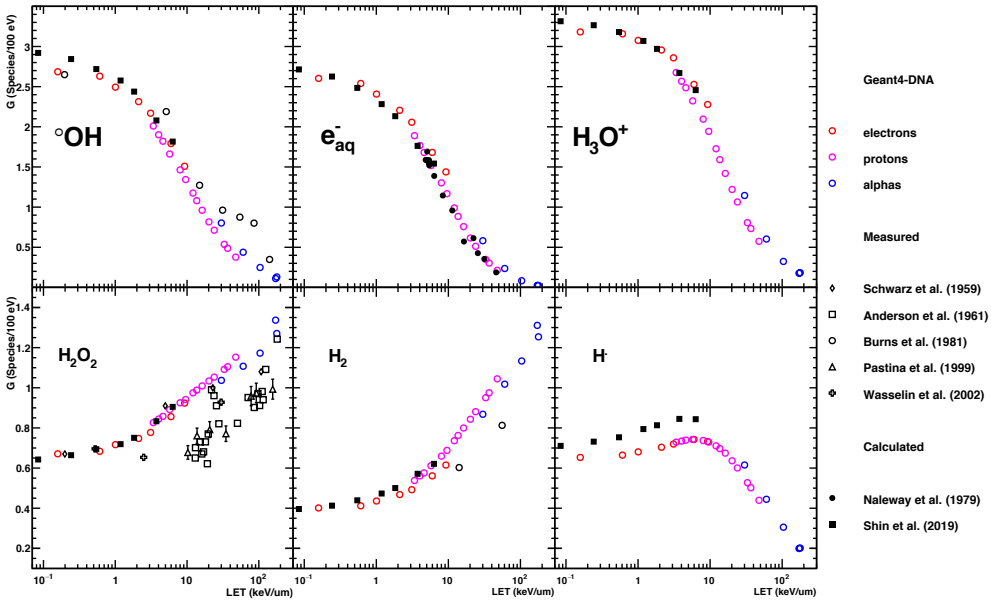


Figure 4.12: G-values at the end of the chemical stage as a function of LET. The red, magenta, and blue markers indicate electron, proton, and alpha case, respectively.

The G-values as a function of LET in the range of 0.16-177.85 keV/μm are calculated as shown in Figure 4.12, because it is reported that the LET values have a close relationship with the radiation effectiveness (Blohm and Harder, 1985, Belli et al., 2009). The radiolysis simulations reproduce the G-values of •OH radicals and solvated electrons of the experimental and calculated data of Naleway et al. (1979), Burns et al. (1981). The G-values of the H_2O_2 and hydrogen molecules at high LET are quite different from the experimental data of Anderson and Hart (1961), Pastina and LaVerne (1999), Wasselin-Trupin et al. (2002) because Geant4-DNA doesn't take the multi-ionization process into account (Meesungnoen and Jay-Gerin, 2005).

The G-values versus LET curves do not completely match between different particle types. It could be guessed that the LET is not sufficient to evaluate radiation effectiveness. The other parameters such as velocity, charge, and energies of secondary particles could impact chemistry kinetics.

4.8. Conclusions

In the present chapter, the IRT technique which simulates the reaction time of a certain reaction based on the distance between two reactants without following the spatio-temporal steps of all species as the already existing Geant4-DNA chemistry model (SBS model) has been implemented in Geant4-DNA by following the GFDE employed in the RITRACKS MCTS toolkit, and the IRT algorithm of TOPAS-nBio, in full collaboration with J. Ramos-Mendez. After the implementation, the validation study on water radiolysis simulations has been performed for electron, proton, and alpha simulations.

The calculation efficiency is strongly improved reaching 1,000 times faster for low LET particles, and this now allows to simulate radiolysis with reasonable calculation time. However, it is still challenging to simulate chemistry for very high LET particles like for low energy protons and alpha particles. For those cases, further research on the pair-searching algorithm which replaces virtual voxels such as k-d tree or octree algorithm is needed (Finkel et al., 1977, de Berg et al., 2008). In addition, LET-specific search range could enable the chemistry simulation to avoid waste of computational resource.

This IRT implementation has been publicly released in Geant4 in June 2020.

References

Anderson, A. R. & Hart, E. J. 1961, Molecular product and free radical yields in the decomposition of water by protons, deuterons, and helium ions. *Radiat. Res.,* 14**,** 689-704.

Bartels, D. M., Cook, A. R., Mudaliar, M., et al. 2000, Spur decay of the solvated electron in picosecond radiolysis measured with time-correlated absorption spectroscopy. *J. Phys. Chem. A,* 104**,** 1686-1691.

Belli, M., Cherubini, R., Finotto, S., et al. 2009, RBE-LET Relationship for the Survival of V79 Cells Irradiated with Low Energy Protons. *Int. J. Radiat. Biol.,* 55**,** 93-104.

Belloni, J., Billiau, F., Delaire, J. A., et al. 1983, Ionizing radiation-liquid interactions: a comparative study of polar liquids. *Radiat. Phys. Chem.,* 21**,** 177-183.

Blohm, R. & Harder, D. 1985, Restricted LET: still a good parameter of radiation quality for electrons and photons. *Radiat. Prot. Dosim.,* 13**,** 337-381.

Brent, R. P. 1971, An algorithm with guaranteed convergence for finding a zero of a function. *The Computer Journal,* 14**,** 422-425.

Burns, W. G., May, R., Buxton, G. V., et al. 1981, Nanosecond proton pulse radiolysis of aqueous solutions. *J. Chem. Soc. Faraday Trans.,* 77**,** 1543-1551.

Burns, W. G. & Sims, H. E. 1981, Effect of radiation type in water radiolysis. *J. Chem. Soc. Faraday Trans.,* 77**,** 2803-2813.

Buxton, G. V. 1972, Nanosecond pulse radiolysis of aqueous solution containing oproton and hydroxyl radical scavengers. *Proc. R. Soc. A,* 328**,** 9-21.

Casella, G., Robert, C. P. & Wells, M. T. 2004. A Festschrift for Herman Rubin. *Generalized accept-reject sampling schemes.* Institute of Mathematical Statistics.

Clifford, P., Green, N. J. B., Oldfield, M. J., et al. 1986, Stochastic models of multi-species kinetics in radiation-induced spurs. *J. Chem. Soc. Faraday Trans. 1,* 82**,** 2673-2689.

Clifford, P., Green, N. J. B. & Pilling, M. J. 1984, Analysis of the Debye-Smoluchowski equation. Approximations for high-permittivity solvents. *J. Phys. Chem.,* 88**,** 4171-4176.

Cobut, V., Frongillo, Y., Patau, J. P., et al. 1998, Monte Carlo simulation of fast electron and proton tracks in liquid water-I. Physical and physicochemical aspects. *Radiat. Phys. Chem.,* 51**,** 229-243.

de Berg, M., van Kreveld, M., Overmars, M., et al. 2008. *Computational geometry,* Berlin, Heidelberg, Springer.

Delaire, J. A., Croc, E. & Cordier, P. 1981, Numerical solution of the Smoluchowski equation

applied to the radiolysis of aliphatic amines and hydrazine. *J. Phys. Chem.,* 85**,** 1549-1555.

Draganić, Z. D. & Draganić, I. G. 1975, Formation of primary reducing yields (Geaq- and GH2) in the radiolysis of aqueous solutions of some positive ions. *Int. J. Radiat. Phys. Chem.,* 7**,** 381-386.

Elliot, A. J. 1994. Rate constants and G-values for the simulation of the radiolysis of light water over the range 0-300 deg. Ontario, Canada: Chalk River Laboratories.

Elliot, A. J., Chenier, M. P. & Ouellette, D. C. 1993, Temperature dependence of g values for H2O and D2O irradiated with low linear energy transfer radiation. *J. Chem. Soc. Faraday Trans.,* 89**,** 1193-1197.

Elliot, A. J., McCracken, D. R., Buxton, G. V., et al. 1990, Estimation of rate constants for near-diffusion-controlled reactions in water at high temperatures. *J. Chem. Soc. Faraday Trans.,* 86**,** 1539-1547.

Fessenden, R. W., Hornak, J. P. & Venkataraman, B. 1981, Electron spin-lattice relaxation times of transient free radicals. *J. Chem. Phys.,* 74**,** 3694-3704.

Finkel, R., Friedman, J. & Bentley, J. 1977, An algorithm for finding best matches in logarithmic expected time. *ACM T. Math. Software***,** 200-226.

Frongillo, Y., Goulet, T., Fraser, M.-J., et al. 1998, Monte Carlo simulation of fast electron and proton tracks in liquid water- II. Nonhomogeneous chemistry. *Radiat. Phys. Chem.,* 51**,** 245-254.

Goulet, T. & Jay-Gerin, J. P. 1992, On the reactions of hydrated electrons with OH·and H3O+. Analysis of photoionization experiments. *J. Chem. Phys.,* 96**,** 5076-5087.

Green, N. J. B., Pilling, M. J., Pimblott, S. M., et al. 1990, Stochastic modeling for fast kinetics in a radiation track. *J. Phys. Chem.,* 94**,** 251-258.

Jay-Gerin, J.-P. & Ferradini, C. 2000, A new estimate of the OH radical yield at early times in the radiolysis of liquid water. *Chem. Phys. Lett.,* 317**,** 388-391.

Karamitros, M., Luan, S., Bernal, M. A., et al. 2014, Diffusion-controlled reactions modeling in Geant4-DNA. *J. Comput. Phys.,* 274**,** 841-882.

Karamitros, M., Mantero, A., Incerti, S., et al. 2011, Modeling radiation chemistry in the Geant4 toolkit. *Prog. Nucl. Sci. Tech.,* 2**,** 503-508.

Kreipl, M. S., Friedland, W. & Paretzke, H. G. 2009, Time- and space-resolved Monte Carlo study of water radiolysis for photon, electron and ion irradiation. *Radiat. Environ. Bioph.,* 48**,** 11-20.

LaVerne, J. A. 2000, OH radicals and oxidizing products in the gamma radiolysis of water. *Radiat. Res.,* 153**,** 196-200.

LaVerne, J. A. & Pimblott, S. M. 1991, Scavenger and time dependences of radicals and molecular products in the electron radiolysis of water: examination of experiments and models. *J. Phys. Chem.,* 95**,** 3196-3206.

LaVerne, J. A., Štefanić, I. & Pimblott, S. M. 2005, Hydrated electron yields in the heavy ion radiolysis of water. *J. Phys. Chem. A,* 109**,** 9393-9401.

Lu, H., Long, F. H., Bowman, R. M., et al. 1989, Femtosecond studies of electron-cation geminate recombination in water. *J. Phys. Chem.,* 93**,** 27-28.

MathWorks 2005. *MATLAB: The language of technical computing. Getting started with MATLAB, version 7*, MathWorks, Incorporated.

Meesungnoen, J. & Jay-Gerin, J. P. 2005, High-LET radiolysis of liquid water with 1H+, 4He2+, 12C6+, and 20Ne9+ ions: effects of multiple ionization. *J. Phys. Chem. A,* 109**,** 6406-6419.

Meesungnoen, J., Jay-Gerin, J. P., Filali-Mouhim, A., et al. 2002, Low-energy electron penetration range in liquid water. *Radiat. Res.,* 158**,** 657-60.

Muroya, Y., Lin, M., Wu, G., et al. 2005, A re-evaluation of the initial yield of the hydrated electron in the picosecond time range. *Radiat. Phys. Chem.,* 72**,** 169-172.

Naleway, C. A., Sauer Jr, M. C., Jonah, C. D., et al. 1979, Theoretical analysis of the LET dependence of transient yields observed in pulse radiolysis with ion beams. *Radiat. Res.,* 77**,** 47-61.

Pastina, B. & LaVerne, J. A. 1999, Hydrogen peroxide production in the radiolysis of water with heavy ions. *J. Phys. Chem. A,* 103**,** 1592-1597.

Pimblott, S. M. & LaVerne, J. A. 1997, Stochastic simulation of the electron radiolysis of water and aqueous solutions. *J. Phys. Chem. A,* 101**,** 5828-5838.

Plante, I. 2011, A Monte-Carlo step-by-step simulation code of the non-homogeneous chemistry of the radiolysis of water and aqueous solutions. Part I: theoretical framework and implementation. *Radiat. Environ. Bioph.,* 50**,** 389-403.

Plante, I. & Devroye, L. 2017, Considerations for the independent reaction times and step-by-step methods for radiation chemistry simulations. *Radiat. Phys. Chem.,* 139**,** 157-172.

Plante, I., Devroye, L. & Cucinotta, F. A. 2013, Random sampling of the Green's Functions for reversible reactions with an intermediate state. *J. Comput. Phys.,* 242**,** 531-543.

Ramos-Mendez, J., Perl, J., Schuemann, J., et al. 2018, Monte Carlo simulation of chemistry following radiolysis with TOPAS-nBio. *Phys. Med. Biol.,* 63**,** 105014.

Rice, S. A. 1985. *Comprehensive chemical kinetics,* Amsterdam, Elsevier.

Sano, H. & Tachiya, M. 1979, Partially diffusion-controlled recombination. *J. Chem. Phys.,* 71**,** 1276.

Schuemann, J., McNamara, A. L., Ramos-Mendez, J., et al. 2019, TOPAS-nBio: An Extension to the TOPAS Simulation Toolkit for Cellular and Sub-cellular Radiobiology. *Radiat. Res.,* 191**,** 125-138.

Schwarz, H. A., Caffrey Jr, J. M. & Scholes, G. 1959, Radiolysis of neutral water by cyclotron produced deuterons and helium ions. *J. American Chem. Soc.,* 81**,** 1801-1809.

Seltzer, S. M., Fernández-Varea, J. M., Andreo, P., et al. 2014. Key data for ionizing-radiation dosimetry: measurement standards and applications. *ICRU report.* Journal of the ICRU: The International Commission on Radiation Units and Measurements.

Shin, W. G., Bordage, M. C., Emfietzoglou, D., et al. 2018, Development of a new Geant4-DNA electron elastic scattering model for liquid-phase water using the ELSEPA code. *J. Appl. Phys.,* 124.

Shin, W. G., Ramos-Mendez, J., Faddegon, B., et al. 2019, Evaluation of the influence of physical and chemical parameters on water radiolysis simulations under MeV electron irradiation using Geant4-DNA. *J. Appl. Phys.,* 126.

Sumiyoshi, T. & Katayama, M. 1982, The yield of hydrated electrons at 30 picoseconds. *Chem. Lett.,* 11**,** 1887-1890.

Tachiya, M. 1978, General method for calculating the escape probability in diffusion-controlled reactions. *J. Chem. Phys.,* 69**,** 2375.

Tachiya, M. 1979a, Effect of an external electric field on the scavenging of geminate ions. *J. Chem. Phys.,* 70**,** 2021.

Tachiya, M. 1979b, Recombination of geminate species in the prescence of a scavenger. *J. Chem. Phys.,* 70.

Tomita, H., Kai, M., Kusama, T., et al. 1997, Monte Carlo simulation of physicochemical processes of liquid water radiolysis. *Radiat. Environ. Bioph.,* 36**,** 105-116.

Wang, F., Schmidhammer, U., Larbre, J. P., et al. 2018, Time-dependent yield of the hydrated electron and the hydroxyl radical in D2O: a picosecond pulse radiolysis study. *Phys. Chem. Chem. Phys.*

Wasselin-Trupin, V., Baldacchino, G., Bouffard, S., et al. 2002, Hydrogen peroxide yields in water radiolysis by high-energy ion beams at constant LET. *Radiat. Phys. Chem.,* 65**,** 53-61.

Wolff, R. K., Bronskill, M. J., Aldrich, J. E., et al. 1973, Picosecond pulse radiolysis. IV. Yield of the solvated electron at 30 picoseconds. *J. Phys. Chem.,* 77**,** 1350-1355.

Chapter 5

DNA damage

Table of contents

5.1. DNA geometry……171
5.2. DNA damage and scoring…..……172
5.2.1. Source classification……..173
5.2.2. Complexity of breaks……176
5.3. Implementation of IRT method into "molecularDNA" example……177
5.4. Verification of the applicability of the IRT approach ("cylinders" approach) ……178
5.5. Evaluation of DNA damage in a simplified human cell ("human_cell" approach)…....179
5.6. Cell repair model……181
5.7. Results for verification of the applicability of the IRT approach……182
5.8. Results for human cell nucleus……183
5.9. Conclusions……191
References……192

The evaluation of biological effects induced by ionizing radiation using MCTS codes combined with the simulation of water radiolysis and geometrical models of DNA, has been studied using PARTRAC (Friedland et al., 2011), KURBUC (Nikjoo et al., 2016), TOPAS-nBio (McNamara et al., 2017), and, recently, RITCARD (Plante et al., 2019). These research groups have developed independent DNA geometrical models and damage repair models based on theoretical approaches or experimental data from the literature. In the case of the Geant4-DNA collaboration, two geometrical models have been proposed, one based on an external program called DNAFabric (Meylan et al., 2016, Meylan et al., 2017), and the other one based on a Python script (Lampe, 2017, Lampe et al., 2018a, Lampe et al., 2018b). Both models have been recently validated by comparing with predictions of other MCTS tools and experimental data (Villagrasa et al., 2017, Sakata et al., 2019, Tang, 2019).

In these works, two main limitations appear: the computing time penalty and the tuning of simulation parameters, which differ between codes. At first, the step-by-step (SBS) method, typically used in combination with MCTS codes for radiolysis simulation, is extremely slow compared to the simulation of the physical stage, for example, reaching several days for the proton case in Geant4-DNA (Meylan et al., 2017). Moreover, most of the MCTS codes have tuned the simulation parameters instead of using the values or settings available in the original papers on which they are based.

In this study, in an attempt to solve such issues, we thus propose to apply the developments of this thesis, such as the electron elastic model (Shin et al., 2018), the calibrated pre-chemical and chemical parameters (Shin et al., 2019), and the independent reaction time (IRT) approach (Ramos-Mendez et al., 2020), into the “molecularDNA” Geant4-DNA example which has been initially developed by Lampe (2017) during his PhD thesis but not released yet. As a first step, a simplified DNA geometry (the so-called "cylinders" approach) is employed in order to verify the results predicted by the "molecularDNA" example when we replace the SBS approach by the IRT approach for the simulation of water radiolysis. In a second step, DNA damage induced in a simplified human cell nucleus is evaluated including the proposed changes, and compared with several experimental data.

5.1. DNA geometry

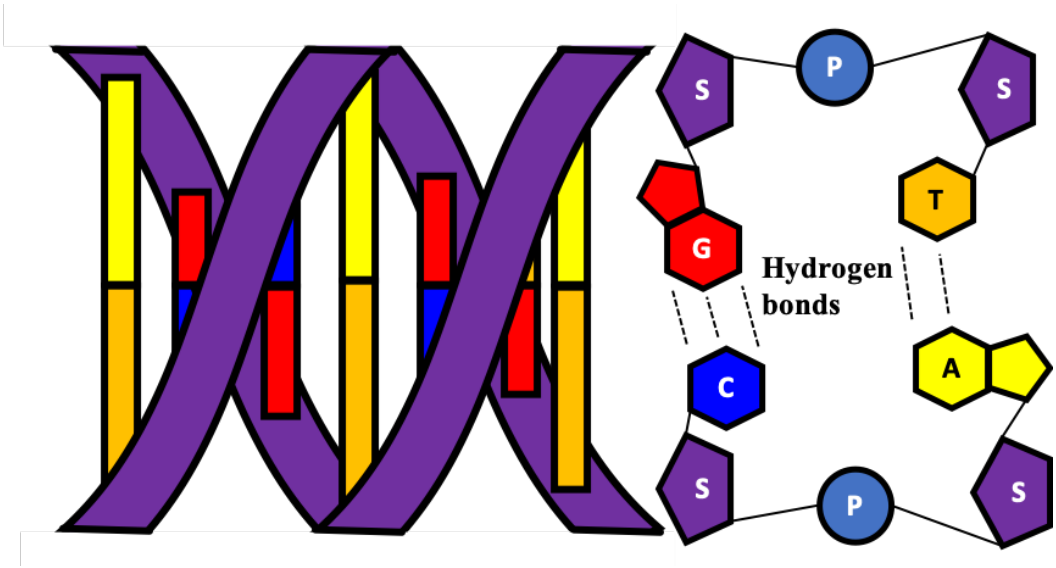


Figure 5.1: A schematic illustration of DNA double helix. A, T, G, C, S, and P represent Adenine ($C_5H_5N_5$, Yellow), Thymine ($C_5H_6N_2O_2$, Orange), Guanine ($C_5H_5N_5O$, Red), Cytosine ($C_4H_5N_3O$, Blue), Sugar ($C_5H_{10}O_4$, deoxyribose, Purple), and Phosphate (H_3PO_4, Light blue), respectively.

The molecular structure of deoxyribonucleic acid (DNA) double-helix was firstly proposed by Watson and Crick (1953). DNA strand consists of a repetition of nucleotides, which deoxyribose ("sugar" below) combined with a phosphate and a nucleobase ("base" below) as shown in Figure 5.1. The DNA bases consist of purine bases (Adenine and Guanine) and pyrimidine bases (Thymine and Cytosine). The basic unit of DNA, called "pair", consists of two sugar-phosphate backbone groups and a couple of DNA bases. Guanine and Cytosine complementarily combine each other with three hydrogen bonds (C≡G), Adenine and Thymine combine with two hydrogen bonds (A=T). In addition, the biological role of DNA was revealed by the experiment of Crick et al. (1961). The genetic information, the "codon" which is the combination of nucleotides triplet, carries the specific information for amino-acids or termination (Shu, 2017). The fractal bunches of DNA double helix construct chromatin fiber with histone proteins in the case of mammalian chromosome (Luger et al., 1997).

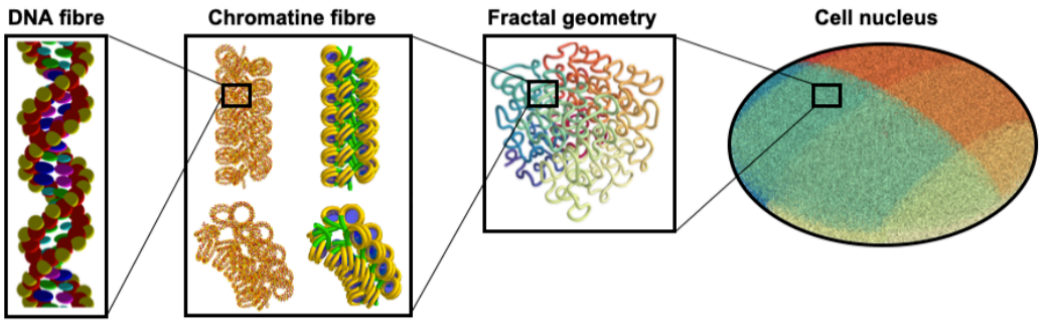


Figure 5.2: Structure of simplified human cell nucleus geometrical model taken from Sakata et al. (2020).

In order to model this complex DNA geometry in Geant4-DNA, Lampe (2017), Lampe et al. (2018a), Lampe et al. (2018b) have proposed a simplified geometrical model of DNA from a fractal

structure based on a Python script. This Python script allows lots of extendibility by modifying the geometrical parameters, and it is very practical to model different types of DNA geometry. Moreover, the "molecularDNA" application is user-friendly and allows to implement the corresponding geometries and to easily test damage parameters using macro commands. Thanks to its easy usage and extensibility, we thus chose to employ the Python script for the modeling of the DNA geometry.

The geometrical information on chromosomes is provided in two parts: the geometrical configuration of nucleotides in a chromatin segment, and the configuration of chromatin fibers, as shown in Figure 5.2. At first, a chromatin segment is defined in an input file based on a simple input structure (macro command: */dnageom/placementVolume segment_model input_file*). The input file contains the shape, position, and size of DNA molecules in a voxel. In order to scheme the realistic inter-connection between chromatin segments, three segment models, “straight”, “turned”, and “turned-twisted”, are available. Then, the fractal structure of the chromosome is generated from a Hilbert curve typically used for a continuous fractal space-filling (Hilbert, 1935). This function is also defined in another input file with the user macro command (*/dnageom/definitionFile input_file*) based on the region-of-interest (usually the cell nucleus), pre-defined by another user macro command (*/chromosome/add name shape size position unit rotation*) and allowing to define chromosome shape.

5.2. DNA damage and scoring

We follow the classification of DNA damage proposed by Nikjoo et al. (1997). Depending on the source of damage such as physical interaction or chemical reaction, direct and indirect damage is defined, respectively. Direct damage is induced by the physical interaction between a particle and DNA, and indirect damage occurs from the chemical reaction between a radical and DNA. Regarding the complexity of the damage, DNA strand breaks are classified as not only single strand breaks (SSB) and double strand breaks (DSB), but also as complex damage. In this section, we present a brief description of all parameters used and the methodology for scoring direct and indirect damage.

5.2.1. Source classification

Direct damage induction

The criterion of DNA damage induction by energy deposition has been studied by many groups and it is now well-known that direct damage is induced by energy deposition of physical particles (Nias, 1998). Especially, for MCTS simulations, the phenomenological parameters of the criterion should be carefully modeled for the plausibility of the damage prediction.

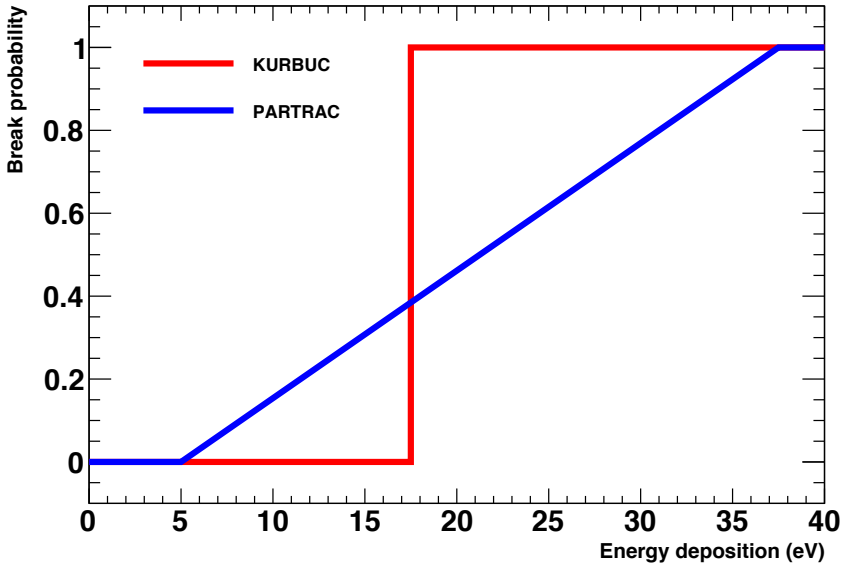


Figure 5.3: The probability of direct DNA damage induction as a function of energy deposition with the energy threshold 17.5 eV as used in KURBUC (Nikjoo et al., 1996), and with the parameters of 5 eV (E_{lower}) and 37.5 eV (E_{higher}) as used in PARTRAC (Friedland et al., 2003).

At first, we should define how much energy deposition induces direct DNA damage. Two energy threshold models for directly breaking the DNA strand are proposed by KURBUC and PARTRAC (Nikjoo et al., 1996, Friedland et al., 2003). KURBUC assumed a single energy threshold of 17.5 eV, and all the energy depositions above the threshold induce direct damage based on the phenomenological estimation of Charlton et al. (1989). On the other side, PARTRAC proposed a linearly proportional model between a lower energy threshold E_{lower} of 5 eV and a higher threshold E_{higher} of 37.5 eV from experiments and parameter adjustment (Prise et al., 2000, Boudaïffa et al., 2000), as shown in Figure 5.3. Unfortunately, there is still no clear evidence which model is more accurate, but we can imagine that a DNA strand break is not determined by a single threshold energy as KURBUC approximated.

Then, in order to determine whether the energy deposit impacts on DNA or not, the effective target volume and radius have to be defined. In the case of PARTRAC, the effective target volume is taken as the backbone sugar-phosphate group with a van-der-Waals radius multiplied by 2 in order to take the charge transfer effect into account in the water shell around DNA. On the other side, KURBUC designed a simple arch structure of B-DNA with 3.4 × 6.5 Å cross-section area

representing an effective volume of sugar-phosphate backbone and one nucleotide pair. Geant4-DNA also assumed an effective radius R_{direct} larger than the van-der-Waals radius of sugar and phosphate (2.28 and 2.63 Å, respectively) in order to consider the water shell effect, and the R_{direct} value was optimized in a previous study using Geant4-DNA (Sakata et al., 2020) as shown in Table 5.1.

Table 5.1: The parameters used for predicting direct DNA damage induction.

	This work (Sakata et al., 2020)	Geant4-DNA (Sakata et al., 2019)	PARTRAC (Friedland et al., 2003)	KURBUC (Nikjoo et al., 1996)
R_{direct}	3.5 Å	4.5 Å	4.56-5.26 Å	1.7-3.25 Å
E_{lower}	5 eV	5 eV	5 eV	17.5 eV
E_{higher}	37.5 eV	37.5 eV	37.5 eV	17.5 eV

Indirect damage induction

The calculation of indirect damage requires water radiolysis simulation. In brief, the molecular species generated during the pre-chemical stage undergo not only chemical reactions with sugar-phosphate backbone inducing so-called indirect strand breaks, but are also scavenged by each other or by histone proteins.

The main chemical reaction inducing indirect strand breaks is between a •OH radical and a sugar-phosphate group. In this work, the probability P_{OH} of this reaction to induce a SSB is set to 40%, causing around 13% of all reactions between DNA and •OH to induce a double strand break, which is in agreement with previous experimental studies and simulations (Lampe et al., 2018b).

In order to reduce computational time for radiolysis simulation, two key chemical parameters have been identified in the previous study of Sakata et al. (2020), derived from the work of Lampe et al. (2018b). First, all the molecular species generated farther away than a "radical kill distance" d_{kill} from DNA molecule are scavenged, with an assumption that only radicals in the hydration shells can cause indirect damage (Ljungman et al., 1991, Daly, 2012). Then, the proper end time T_{chem} of the chemical stage is determined from the equivalent time of •OH radical based on the maximum distance (d_{kill}) this radical can travel. All the chemical parameters taken in this work and used in the other works are given in Table 5.2.

Table 5.2: The parameters used for predicting indirect DNA damage.

	This work (Sakata et al., 2020)	Geant4-DNA (Sakata et al., 2019)	PARTRAC (Friedland et al., 2003)	KURBUC (Nikjoo et al., 2001)
P_{OH} (%)	40.5	40	70	13
d_{kill} (nm)	9	4.5	12.5	4
T_{chem} (ns)	5	2.5	10	1
T_{step} (ns)	0.5	0.5	-	-

One thing should be noted: the maximum time step T_{step} listed in Table 5.2 is only applicable to the IRT method in order to compensate the disagreement between SBS method and IRT. This will be further discussed in section 5.3.

The chemical reactions between the DNA molecule and the radicals are proposed by Buxton et al. (1988). The reactions shown in Table 5.3 have been implemented into the Geant4-DNA chemistry constructor "G4EmDNAChemistry_option3" specifically for this thesis.

Table 5.3: Chemical reactions, reaction rates and reaction radii between molecular species and DNA components, proposed by Buxton et al. (1988).

Reaction	k_{obs} **($M^{-1}s^{-1}$)**	**Reaction radius (nm)**
$^{\cdot}OH + A$	6.1×10^{9}	0.372
$^{\cdot}OH + T$	6.4×10^{9}	0.384
$^{\cdot}OH + G$	9.2×10^{9}	0.553
$^{\cdot}OH + C$	6.1×10^{9}	0.366
$^{\cdot}OH + Sugar$	1.8×10^{9}	0.108
$e^{-}_{aq} + A$	0.9×10^{10}	0.541
$e^{-}_{aq} + T$	1.8×10^{10}	1.081
$e^{-}_{aq} + G$	1.4×10^{10}	0.841
$e^{-}_{aq} + C$	1.3×10^{10}	0.781
$e^{-}_{aq} + Sugar$	1.0×10^{7}	6.01×10^{-4}
$H^{\cdot} + A$	1.0×10^{8}	6.01×10^{-3}
$H^{\cdot} + T$	5.7×10^{8}	3.42×10^{-2}
$H^{\cdot} + C$	9.2×10^{7}	5.53×10^{-3}
$H^{\cdot} + Sugar$	2.9×10^{7}	1.74×10^{-3}

5.2.2. Complexity of breaks

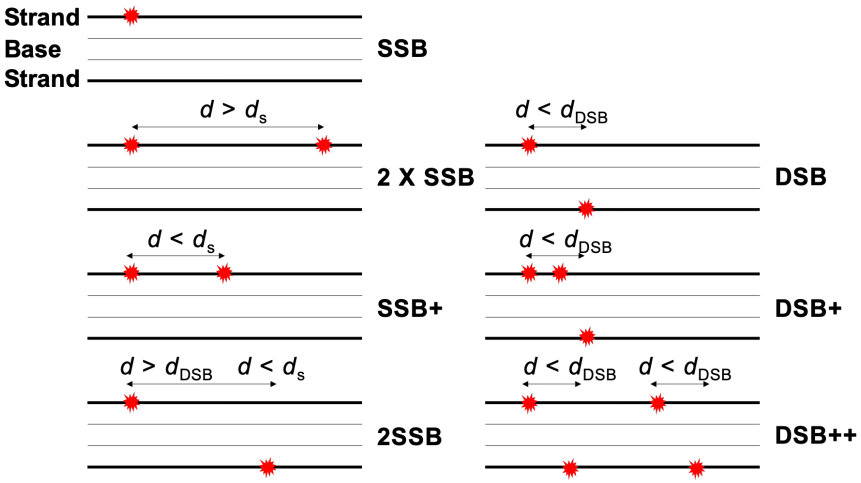


Figure 5.4: A schematic illustration of damage classification for complexity of SSB (left) and DSB (right) based on the paper of Nikjoo et al. (1996).

The breaks, regardless of the source (direct or indirect), are classified as a function of their complexity as shown in Figure 5.4. The DSB represents two opposite strands that are simultaneously damaged within the distance d_{DSB} between two damages, taken as 10 bp. If the separation distance between two breaks is larger than d_{DSB}, the breaks are taken as two independent SSBs. Nikjoo et al. (1996) also proposed fragment gap d_s which determines complex SSBs. In Geant4-DNA, a fragment gap d_s value of 100 bp is used (Lampe et al., 2018b). The complex SSBs are denoted as SSB+ or 2SSB, if the strand damages are located in same or opposite strands, respectively. The complex DSBs, DSB+ and DSB++, are more important than the complex SSBs because those determine the irreparable DNA damages (further described in section 5.8). DSB+ classification requires one DSB and one more break within d_{DSB}. Besides, two DSBs within the fragment gap d_s are classified as DSB++.

In addition, the single and double strand breaks can also be classified as direct damage, denoted as SSB_d and DSB_d, or indirect damage, denoted as SSB_i and DSB_i, or mixed damage, denoted as SSB_m and DSB_m. DSB_{hyb} is the mixed damage classification, however, only when the indirect damage is the core damage for DSB.

5.3. Implementation of IRT method into the "molecularDNA" example

We recall that the IRT method approximates the reaction time based on the reaction rate and the separation distance between a pair of molecular species, without considering the diffusion, as previously described in Chapter 4. Thus, the approach presents the advantage of much faster simulation times compared to the SBS method. However, the simulation of chemistry kinetics without considering diffusion may cause imprecision in simulated DNA damage.

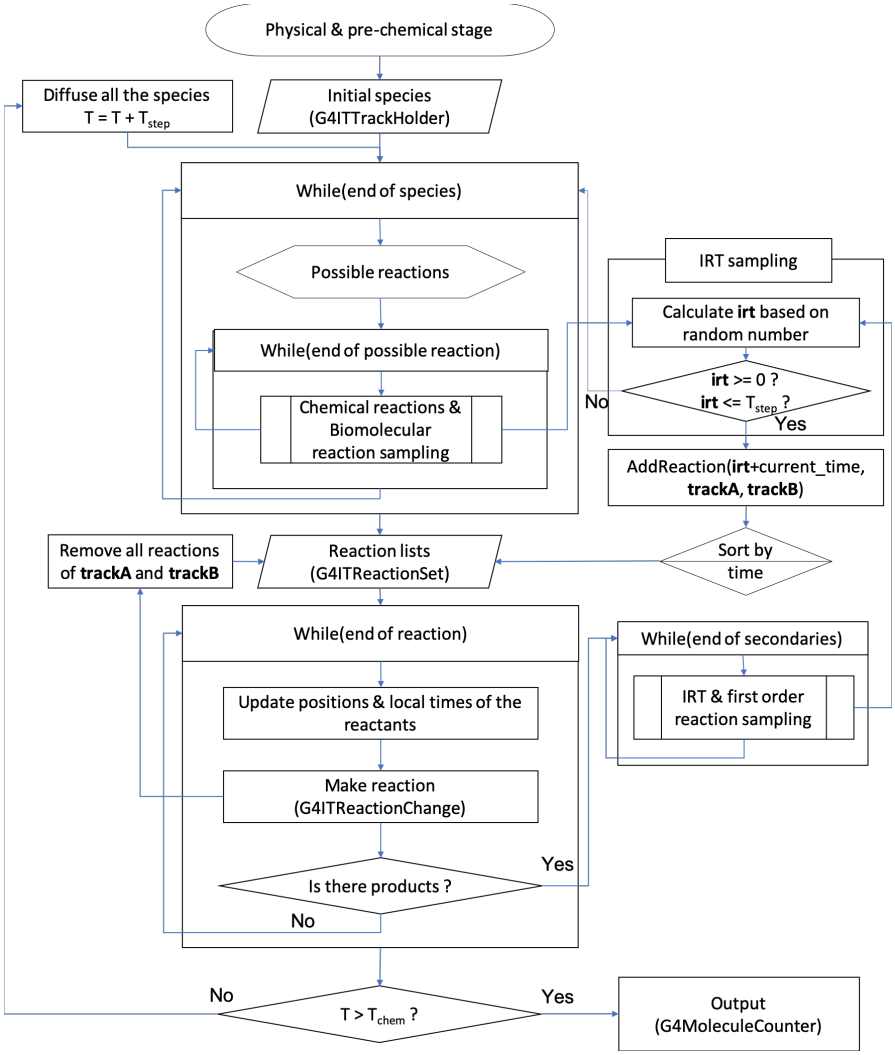


Figure 5.5: Flowchart of IRT with time step T_{step} as implemented in this thesis.

In order to minimize the distortion induced by the IRT approximation, time steps T_{step} are employed in this study, as shown in Figure 5.5. This idea was initially proposed in previous Geant4-DNA works for the simulation of E. coli bacterium DNA damage induction from ionizing radiation (Lampe et al., 2018b, Lampe, 2017, Karamitros et al., 2020). The whole chemistry simulation until T_{chem} is thus split into several time steps of duration T_{step}, and the reaction times are sampled at each time step.

The search range between a molecular species and a DNA molecule is determined with 95% confidence level similarly as for chemical reaction between reactants:

$$d_{95} = 2\sqrt{2D\Delta t} \tag{5.1}$$

However, *D* should here be taken as the diffusion coefficient of molecular species, instead of the summation of diffusion coefficients. The octree node approach is employed as search algorithm (Finkel et al., 1977). Octree data structure divides a node into eight smaller nodes. The octrees of this application contain the positions of all the radicals, allowing Geant4-DNA to rapidly identify neighbor reactants in the search range of the DNA molecule.

5.4. Verification of the applicability of the IRT approach ("cylinders" approach)

In order to verify the correct combination of our IRT with DNA geometry, we used the simple geometrical approach (so-called "cylinders" approach, available in the "molecularDNA" example) developed by Lampe et al. (2018a). This is a useful approach to verify the method, that has already been used by other research groups, such as KURBUC and TOPAS-nBio (Charlton et al., 1989, Nikjoo et al., 1996, McNamara et al., 2017).

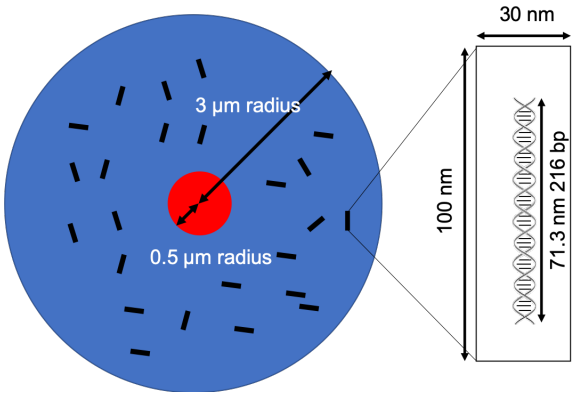


Figure 5.6: A spherical water phantom of 3 µm radius is filled with 200,000 individual chromosomes (blue volume, left), it includes a source spherical volume of 0.5 µm radius placed at its center (red, left). A DNA segment of 216 bp placed in 100×30×30 nm^3 voxels is shown in the right panel.

A spherical water phantom of 3 µm radius is uniformly filled with 200,000 individual straight DNA segments of 216 bp (total 43.2 Mbp) placed in 100×30×30 nm^3 voxels, as shown in Figure 5.6. For this test, 4.5 keV electrons are isotropically generated in a smaller sphere of 500 nm radius placed at the center of the water phantom. In order to reproduce the results of Lampe (2017), we took exactly the same parameters as he did. The number of initial electrons was set to 10^6. The G4EmDNAPhysics_option4 physics constructor which uses Uehara screened Rutherford elastic model (Uehara et al., 1993) and the dielectric inelastic model (Kyriakou et al., 2016) was used, in combination with the default G4EmDNAChemistry chemistry constructor based on the chemistry model of PARTRAC (Kreipl et al., 2009) used by the IRT method. The physical and chemical parameters for damage scoring are shown in Table 5.4.

Table 5.4: The physical and chemical parameters used for damage scoring, in order to reproduce the results of Lampe (2017).

Physical parameters		Chemical parameters	
R_{direct}	7 Å	P_{OH}	65%
E_{lower}	17.5 eV	d_{kill}	0-9 nm
E_{higher}	17.5 eV	T_{chem}	1 and 1,000 ns
		T_{step}	0.5 ns

The number of breaks classified by complexity and source is evaluated as a function of radical kill distance d_{kill}. In addition, the amount of chemical reactions between radicals and DNA molecules (base and strand) as a function of radical kill distance d_{kill} is also calculated. The results are compared with the previous results of Lampe (2017).

5.5. Evaluation of DNA damage in a simplified human cell ("human_cell" approach)

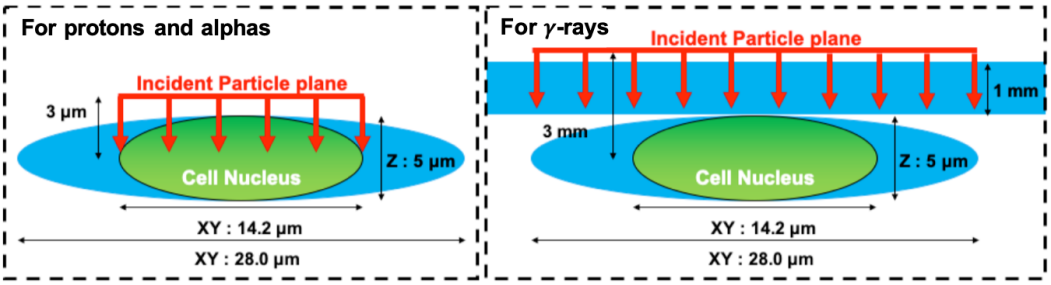


Figure 5.7: A schematic illustration of geometrical configuration of human cell nucleus and source term taken from Sakata et al. (2020).

Following the work of Sakata et al. (2019), Sakata et al. (2020), we considered a simplified geometry of a human cell nucleus, consisting of an ellipsoidal cell nucleus of 14.2×14.2×5 µm^3 surrounded by an ellipsoidal water phantom of 28×28×5 µm^3 representing the cytoplasm, as shown in Figure 5.7. The total number of base pairs included in this nucleus is 6.4 Gbp (bp density of 0.012 bp/nm^3), consistent with the reported bp density of a mammalian cell (about 0.015 bp/nm^3) (Suzuki et al., 1998, Zhong et al., 2018).

The ^{137}Cs and ^{60}Co photon beams, proton beams of 0.3, 0.4, 0.7, 1.0, 1.67, 2.34, 4.0, 7.0, 50 MeV, alpha beams of 5, 10, and 15 MeV of initial energies are generated from a plane parallel to the cell nucleus (Sakata et al., 2020). In these simulations, it is almost impossible to reproduce the exact experimental setup, in particular we did not consider substrate on which cells are platted. However, in the case of gamma rays, the cells are on a flask (Asaithamby et al., 2008) and we included a water absorber of 1 mm thickness, as in the configuration of Sakata et al. (2020).

For this study, we used the G4EmDNAPhysics_option2 physics constructor which is the default physics constructor of Geant4-DNA and the G4EmDNAChemistry_option3 chemistry constructor developed in this thesis, based on RITRACKS (Plante and Devroye, 2017) with the parameters for damage scoring given in Tables 5.1 and 5.2.

In addition, the **improvements** proposed in this thesis are implemented in the simulations:

1) the new ELSEPA electron elastic scattering model developed in Chapter 2 (Shin et al., 2018),
2) the Meesungnoen electron thermalization model for thermalized and auto-ionized electrons (Shin et al., 2019),
3) the pre-chemical parameters suggested in Chapter 3,
4) our IRT method and chemical reaction table developed in Chapter 4 (Ramos-Mendez et al., 2020),
5) the other bug fixes such as initial displacement of $^{\bullet}$OH radicals and electron attachment model.

The numbers of total strand breaks, SSBs, DSBs, as a function of LET (which values are based on (ICRU, 2014)) and the corresponding SSB/DSB ratios are calculated.

One thing should be noted: the measurements of DNA damage shown in this study, (Frankenberg et al., 1999, Hoglund et al., 2000, Belli et al., 2000, Belli et al., 2001, Leloup et al., 2005, Campa et al., 2005), use gel electrophoresis approach such as agarose gel electrophoresis (AGE) (Kryndushkin et al., 2003), constant-field gel electrophoresis (CFGE), and pulsed-field gel electrophoresis (PFGE) (Iliakis et al., 2009). However, it is reported that this approach has a limitation to count very small fragments of DNA, for example less than 23 kbp (Campa et al., 2005), due to the limitation of the detection method. Thus, the yields of distant DSBs, separated by at least 10 kbp between two DSBs, are also calculated in this study

On the other hand, an approach for DNA damage measurement has been proposed through the counting of the number of immuno-fluorescent foci such as γ-H2AX (Olive, 2004, Rothkamm et al., 2015). This approach also has limitation such as the non-linear correlation between foci number and DSB yields, however, the sensitivity is much better than PFGE (Löbrich et al., 2010). Petkovic et al. (2019), Ristić Fira and Petrovic (2020) (private communication) evaluated DNA damage by measuring γ-H2AX for gamma ray and proton beams. The details of the experimental data presented in this study are given in Table 5.5.

Table 5.5: The experimental data and corresponding reference, cell line, cell type, and the approach used for the damage measurement.

Reference	Name of the cell line	Description	Damage measurement
Frankenberg et al. (1999)	-	Human fibroblasts	PFGE
Hoglund et al. (2000)	GM5758	Human fibroblasts	PFGE
Belli et al. (2000)	V79	Chinese hamster cells	CFGE
Belli et al. (2001)			PFGE
Leloup et al. (2005)	XL2-Blue MRF	Plasmid DNA of bacteria	PFGE
Campa et al. (2005)	AG1522	Human fibroblasts	PFGE
Petkovic et al. (2019), Ristić Fira and Petrovic (2020)	HTB177	Normal lung tissue	γ-H2AX

The calculated SSB and DSB results are compared with the experimental data listed above, as well as the data calculated by Nikjoo et al. (2001), Friedland et al. (2003), Meylan et al. (2017), Sakata et al. (2020).

The histogram of fragment length distribution, which represents the distance between two DSBs, is calculated with 100 Gy of 1 MeV protons, and compared with the experimental data of Belli et al. (2001), Campa et al. (2005).

Another experimentally accessible quantity is the protectable damage fraction (or scavengeable fraction). This is the fraction of the damage at the infinite dimethyl-sulfoxide (DMSO radical scavenger) concentration. Such chemicals can scavenge free molecular species, especially •OH radicals. In the simulation, the protectable damage fraction is the ratio between the DSBs induced by indirect damage and all the DSBs. We calculated these fractions as a function of LET, and compared with the experimental data of Ito et al. (2006), Hirayama et al. (2009).

5.6. Cell repair model

Our simulations can be used to predict "early" DNA damage, that is up to 5 ns in this study. Modelling radiobiological phenomena during the (longer) biological stage requires complex mathematical models (Lea, 1955). It is reported that DSBs induce deletions, translocations, and fusions of the DNA, when those are not correctly repaired (Negritto, 2010). A review of existing repair models and the implementation of some of these models are out of the scope of this thesis work (Frankenberg-Schwager, 1989), mainly due lack of time.

However, an initiative was proposed by JINR, Dubna in Russia and CENBG in 2017, in the context of the Geant4-DNA collaboration, regarding the implementation of a repair model based on the advanced mathematical model of Belov et al. (2015). This model assesses the principal “repair pathways”, named non-homologous end-joining (NHEJ), homologous recombination (HR), single-strand annealing (SSA), and alternative end-joining mechanism (Alt-NHEJ) (The details of the pathways are described in the paper of Heyer et al. (2010), Decottignies (2013), but will not be presented here).

In this model, the total yield of DSBs, N_0, as a function of time can be calculated based on the repair pathways as:

$$\frac{dN_0}{dt} = \alpha(L)\frac{dD}{dt}N_{ir} - V_{NHEJ} - V_{HR} - V_{SSA} - V_{micro-SSA} - V_{Alt-NHEJ} \quad (5.2)$$

where D is the absorbed dose (Gy), and $\alpha(L)$ is the DSB induction per dose (Gy^{-1} per cell). N_{ir} is the yield of irreparable DSBs, representing (with our damage scheme) $N_{DSB}^{+} + 2\times N_{DSB}^{++}$. V_{NHEJ}, V_{HR}, V_{SSA}, $V_{micro\text{-}SSA}$, and $V_{Alt\text{-}NHEJ}$ are the repair potential for NHEJ, HR, SSA, micro-SSA, and Alt-NHEJ repair pathways, respectively. The repair potentials are given with 29 differential equations and 54 parameters including rate constants for human fibroblasts in the paper of Belov et al. (2015). This model enables to calculate five foci yields, which refer to the response of specific proteins to DSBs (Rothkamm et al., 2015), such as Ku protein, DNA-PKcs, RPA, Rad51, and γ-H2AX.

In this study, we used the repair model of Belov et al. (2015), which has been recently implemented in the "molecularDNA" example by Sakata et al. (2020). The foci yield of γ-H2AX

as a function of time is calculated and compared with the experimental data performed by Asaithamby et al. (2008).

5.7. Results for verification of the applicability of the IRT approach

The IRT method can significantly reduce the simulation time for DNA damage prediction compared to the SBS (several days in the previous study (Meylan et al., 2017)) as shown in Table 5.6. Besides, the calculation time is also reduced at short radical kill distance d_{kill}.

Table 5.6: Simulation time according to the radical kill distance d_{kill}.

Radical kill distance d_{kill} (nm)	Simulation time (CPU-hrs)
0	42.95
1	44.18
2	44.20
3	44.79
4	44.92
5	45.68
6	46.14
7	47.17
8	48.29
9	50.27

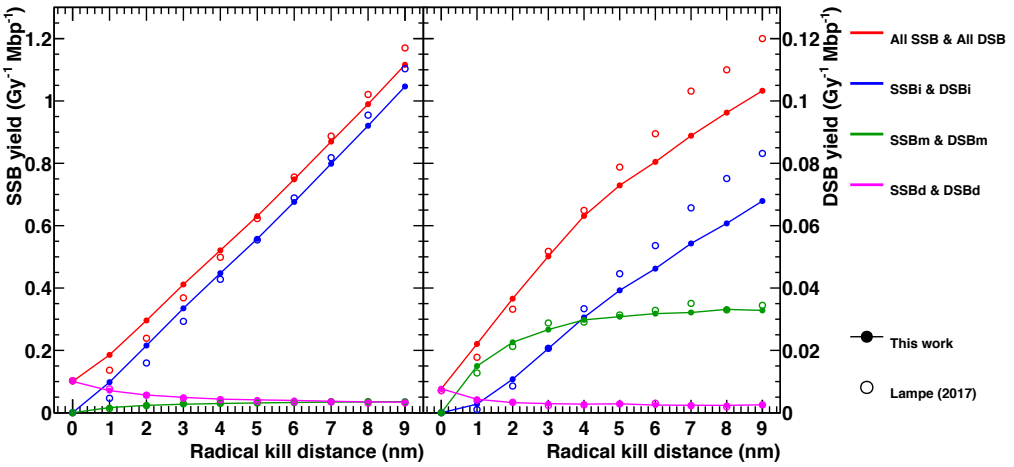


Figure 5.8: The SSB (left) and DSB (right) yields as a function of radical kill distance according to the damage source.

As shown in Figure 5.8, the direct and multi damage yields can be nicely reproduced, however, there is small disagreement between the curves of this work and those of Lampe (2017), even if both approaches use exactly the same parameters. In particular, the indirect damage yields are larger at 1-4 nm d_{kill}, and smaller at 5-9 nm d_{kill} than Lampe (2017) up to 4.6% and 13.9% for SSB and DSB yields, respectively. There is no clear evidence for the reason of this disagreement, but we could guess that it originates from the two different IRT approaches, especially on the stepping method (this work versus the study of Karamitros et al. (2020)). However, the disagreement is not

so large, especially at 5 nm d_{kill} which represents the diffusive distance of •OH radical until 5 ns, equal to 5.29 nm. Moreover, the number of particles used in Lampe (2017) is 10 times less than this work. Thus the statistical error of Lampe (2017) might be bigger than in our study.

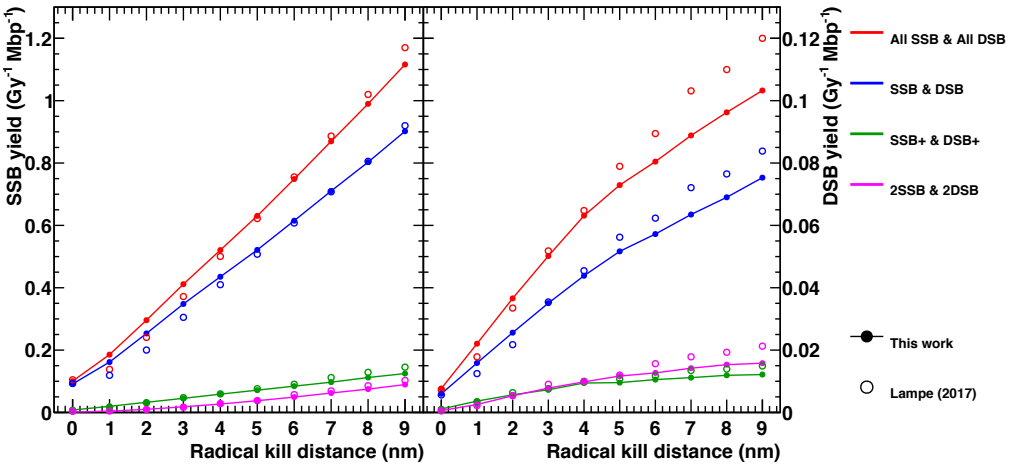


Figure 5.9: The SSB (left) and DSB (right) yields as a function of radical kill distance according to the damage complexity.

With similar tendency as in Figure 5.8, the SSB yields according to complexity are almost identical within 5% relative errors with the previous study of Lampe (2017) as shown in Figure 5.9. In the case of DSBs, the disagreements are 10.1, 18.4, and 25.7% for DSB, DSB+, and 2DSB, respectively. However, the differences at 5 nm d_{kill} are still acceptable. It should be noted that the chemistry simulation in this validation study goes up to 1 μs, and the differences are much smaller when the value of T_{chem} is decreased.

5.8. Results for human cell nucleus

Table 5.7: The cases of ^{137}Cs, ^{60}Co, protons, alphas, and the corresponding LET, number of particles, memory consumption, and simulation time.

Energy (MeV)	ICRU-90 LET (keV/um)	Number of particles	Run splitting	Simulation time per run (CPU-hrs)
^{137}Cs	0.8[a]	25,000,000[b]	25	47.43
^{60}Co	0.4[a]	5,000,000	10	56.69
0.30 (protons)	54.41	3,000	30	146.90
0.40	46.48	3,000	30	101.76
0.70	33.14	3,000	30	44.60
1.00	25.77	3,600[c]	36	27.78
1.67	18.12	3,000	30	15.46
2.34	14.31	8,000	20	100.63
4.00	9.33	10,000	1	435.54
7.00	6.11	10,000	1	654.87
50.00	1.24	10,000	1	25.67
5 (alphas)	87.54	100	10	29.67
10	52.94	100	10	21.51
15	38.96	100	10	12.86

[a] From ICRP-92 report (Valentin, 2016).
[b] corresponding to about 1 Gy dose for reproducing experimental data of repair model.

[c] corresponding to about 100 Gy dose for reproducing experimental data of fragment distribution.

As described before, the computational power for simulating all the DNA structure in a human nucleus and storing damage is still challenging. The "molecularDNA" example with IRT approach proposed in this study can reduce very significantly calculation time as shown in Table 5.7. However, the memory consumption goes up to 160 Go, in particular for the storing of DNA damage. Thus, we decided to split the number of particles and runs. The memory is cleaned up for every split, and after all the simulations are finished, the output file is merged and analyzed.

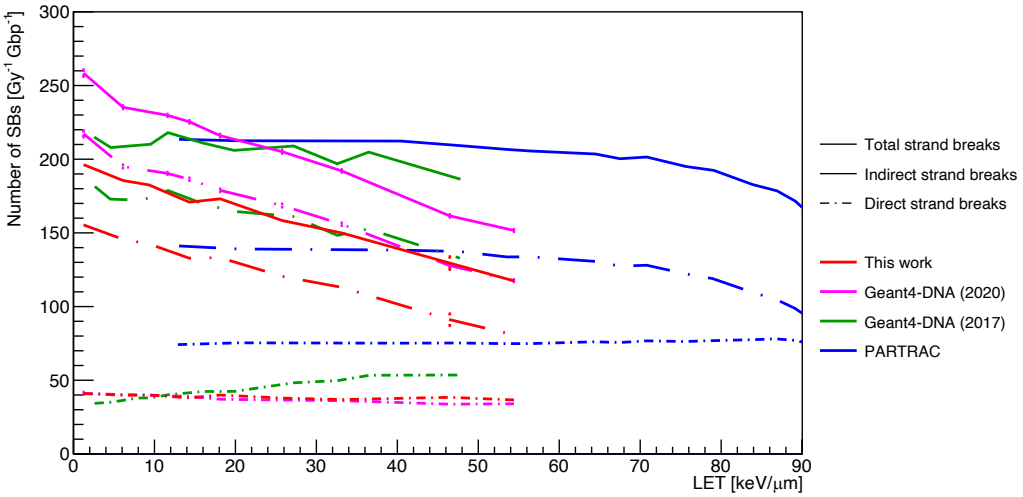


Figure 5.10: The number of total strand breaks as a function of LET calculated by Geant4-DNA (this work, Sakata et al. (2020), Meylan et al. (2017)) and PARTRAC (Friedland et al., 2003).

Figure 5.10 shows the number of total, indirect and direct strand breaks for incident protons as a function of LET calculated in this thesis and in other studies. The pink curves (Geant4-DNA 2020 (Sakata et al., 2020)) in Figure 5.10 and the red curves (this work) differ only by the improvements developed in this thesis and listed in section 2.5. The number of indirect damages (long dash - dot red curve) decreases as a function of LET unlike the number of direct damages, because the concentrated molecular species at high LET rapidly react with each other as reported before (Hirayama et al., 2009). Compared to the previous studies (Geant4-DNA 2020 (Sakata et al., 2020), Geant4-DNA 2017 (Meylan et al., 2017), PARTRAC (Friedland et al., 2003)), the curve for indirect SBs is significantly lower. This could be explained by a combination of several factors, such as the initial distribution of radicals determined by our new elastic scattering model, and the revised chemical parameters of diffusion coefficients and reaction rates. In particular, we can explain the difference with Geant4-DNA 2020 (Sakata et al., 2020) results by the fact that G-values of $^{\bullet}$OH radicals (mainly responsible for indirect damage induction) with the new pre-chemical and chemical models are smaller than that of Geant4-DNA 2020 (Sakata et al., 2020) and also 2017 (Meylan et al., 2017) as described in the Chapter 3 and 4. The direct SB yields induced by physical interaction should intuitively be independent of LET unless the radiation overkills DNA (Goodhead,

1988, Chang et al., 2014) (> 100 keV/μm in Chinese hamster cells (Mehnati et al., 2005)). All the studies show direct SB yields independent of LET except Geant4-DNA 2017 (Meylan et al., 2017). All Geant4-DNA simulations are smaller than PARTRAC results due to the small effective target volume (R_{direct} in this study). Of course, the new elastic scattering model doesn't impact the direct SB yields.

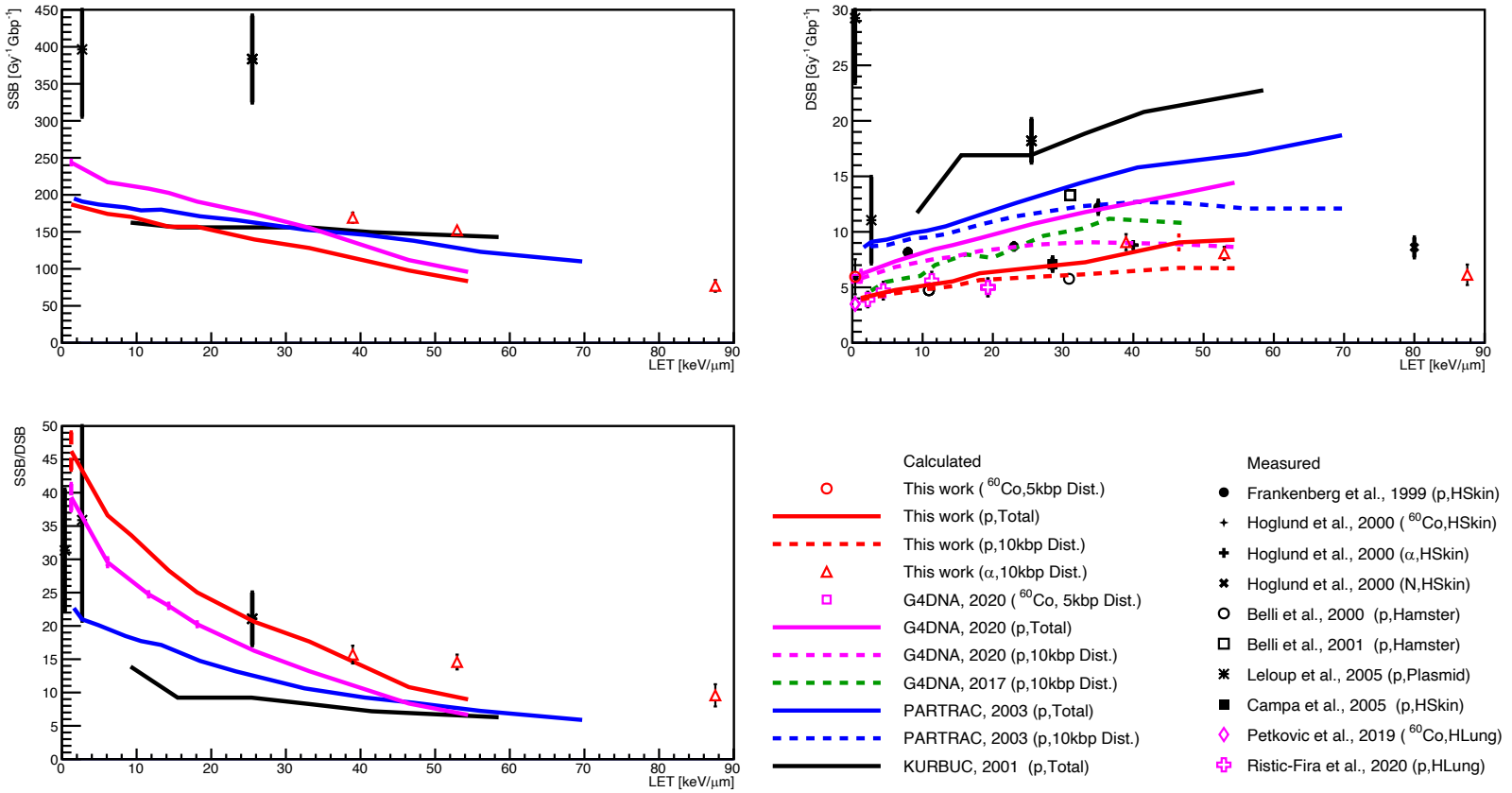


Figure 5.11: The SSB (left upper) and DSB yields (right upper), and SSB/DSB ratio (left below) as a function of LET.

Figure 5.11 shows the calculated SSB and DSB yields, and the SSB/DSB ratio for incident gamma, protons, and alphas as a function of LET. The tendency of SSB and DSB as a function of LET is indeed opposite because DSB consists two SSBs, they are competitive. With the increase of LET, in other words for higher concentration of species, DSB yield also increases.

The main change in this study is that less DSB yields are predicted compared to the other Monte Carlo tools such as Geant4-DNA 2017, 2020, PARTRAC, KURBUC, which cannot reproduce the small DSB yields measured experimentally for human fibroblast cells (HSkin in Figure 5.11) as can be observed for DSBs. That because the total number of indirect SBs are smaller as shown in Figure 5.10 due to the chemistry models used in this study.

The quantitative comparison with experimental data is a challenge because the uncertainties of the measurements are still large due to the cell cycle, measurement condition, beam properties, etc. In addition, the experimental data for the SSB and DSB yields has been measured with significant variations such as the experimental methodology and the type of cell as listed in Table 5.5.

The data of Leloup et al. (2005) shows relatively higher SSB and DSB yields compared to the other data, apparently due to the influence of base pair density and histone scavenging effect (Sakata et al., 2020). Note that: the base pair densities of the plasmid, human fibroblast cell, and the hamster cell (V79) are 9.4×10^{-6}, 0.012, 0.015 bp/nm^3, respectively (Suzuki et al., 1998, Zhong et al., 2018, Belli et al., 2001).

The DSB yields in this work show lower than the data of Frankenberg et al. (1999) down to 62%, however, the other experimental data of Hoglund et al. (2000), Campa et al. (2005) match well with the results of this work within 10% difference. In particular, our simulations are getting closer to the recent experimental data performed by Petkovic et al. (2019), Ristić Fira and Petrovic (2020) (magenta diamond and crosses) even though the data is for human lung cells and assumes that the number of foci/cell is equal to the number of DSB/nucleus. The DSB yields of alpha cases (red triangles) also show good agreement with the data of Hoglund et al. (2000) (black cross). Unlike proton cases, the alpha cases slightly decreased as a function of LET. This tendency is apparently because of the overkill effects due to the lack of the remaining normal DNA strands (Hall and Giaccia, 2018).

In addition, higher SSB/DSB ratio is calculated in this study compared to the other simulations. This result is close to the plasmid data of Leloup et al. (2005), especially at the LET of 25.5 keV/μm, even though plasmids are not cells.

As underlined in several studies (Sakata et al., 2020, Lampe et al., 2018a, Zhu et al., 2020, Nikjoo et al., 1997), one should keep in mind that the numbers of SSBs and DSBs are very sensitive to the damage scoring parameters, such as the parameters given in Table 5.1 and 5.2. New and more systematic experimental data on cellular irradiation will be clearly needed to better validate the simulations.

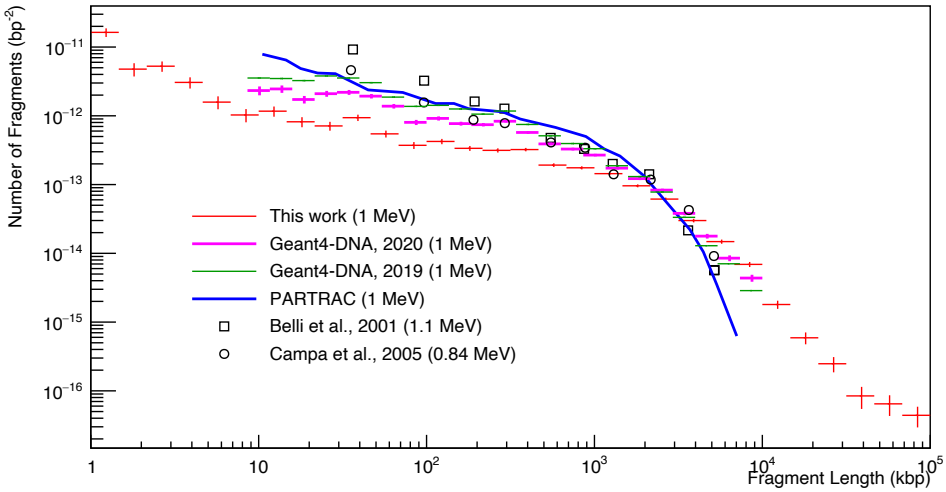


Figure 5.12: Histogram of the fragment length distribution after 100 Gy irradiation with 1 MeV protons.

The fragments distribution is shown in Figure 5.12 for 1 MeV proton irradiation. Small fragments are the most frequent, and the distribution decreases as a function of fragment length. The results of this study obtained for 100 Gy are similar to PARTRAC simulations and experimental data of Belli et al. (2001) (3 MeV incident proton beam and a corresponding energy at beam exit of 1.1 MeV) in Chinese hamster cells (V79) and Campa et al. (2005) in human fibroblasts.

We have observed that the results in this study overestimates the long fragments above 50 kbp compared to PARTRAC and experimental results even though the experimental range of fragment counting method is 23 kbp to 5.7 Mbp (Campa et al., 2005). It can be intuitively assumed that the higher number of DSB damages increases the number of fragments at short fragment length, and decreases at short fragment length if the geometry is exactly same. For example, the DSB yields of Campa et al. (2005) are much less than Belli et al. (2001) as shown in top right panel of Figure 5.11. On the other hand, the PFGE approach subtracts the fragment background, that is, the DNA fragment distribution of unirradiated cells (Pinto et al., 2000), however, this approach could lead another systematic uncertainty due to the oversimplification of background fragment distribution (Newman et al., 2000). In addition, the interparticle interference could be a reason of the disagreement. In this work, each particle is independent because we assumed radiobiological effects in low-dose region. However, the experimental data of Belli et al. (2001), Campa et al. (2005) are for the relative high dose rate of 20 and 3.5 Gy/min, respectively.

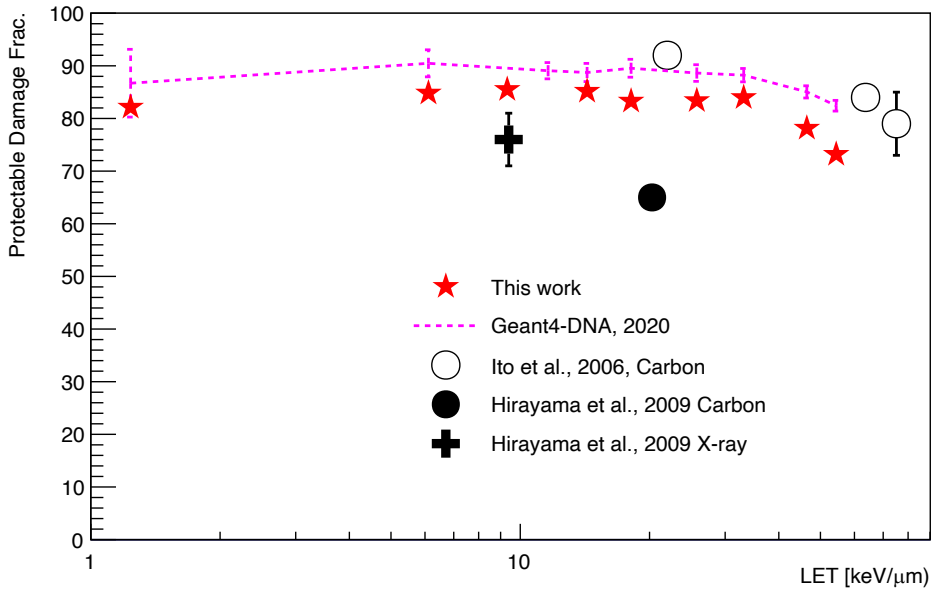


Figure 5.13: Protectable damage fraction, which is the ratio of protectable DSBs over the total number of DSBs, as a function of LET.

The protectable (or scavengeable) damage fraction as a function of LET is also calculated, as shown in Figure 5.13. As reported so far, the indirect damage fraction is dominant for low LET irradiation, and decreases as a function of increasing LET (Hirayama et al., 2009). In the LET range of 1.24-54.41 keV/μm, the results of this work are lower than the previous study of Geant4-DNA up to 10% (Sakata et al., 2020) because this work assessed less indirect damage due to the changes in pre-chemical and chemical stages as we described in Chapters 3 and 4. Besides, the results are between the data by Ito et al. (2006) and the data by Hirayama et al. (2009). There is no clear reason why, but one hypothesis to explain the observed difference could arise from the fact that experimental data of Hirayama et al. (2009) in V79 hamster cells are also significantly different with the data of Ito et al. (2006) measured in HL-60 human leukemia cell. We could guess that the cell size or base pair density could impact the fraction of indirect damage as recently reported (Tang et al., 2019).

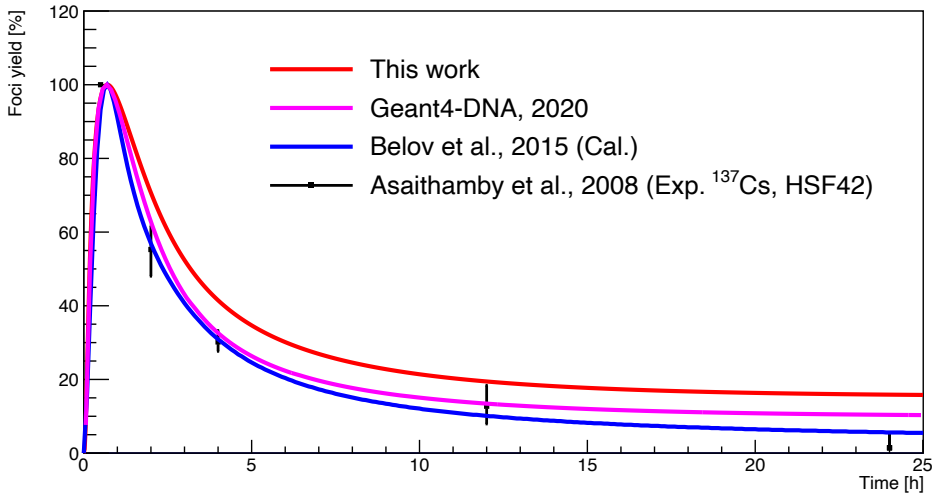


Figure 5.14: γ-H2AX yield as a function of repair time from the irradiation by ^{137}Cs at the dose of 1 Gy. The calculated repair model in this study is compared with the calculations of Belov et al. (2015), Sakata et al. (2020) and the experimental data of Asaithamby et al. (2008).

Figure 5.14 shows the results of scaled γ-H2AX yield of ^{137}Cs at the dose of 1 Gy as a function of time up to 25 hours after irradiation. The results in this work are obtained with the calculated number of DSBs (4.04 Gbp^{-1}) and the irreparable fraction (~ 0.15) by the approach of Belov et al. (2015). The foci yields calculated in this study are larger than the calculation of Belov et al. (2015), Sakata et al. (2020) and experimental data. There are two factors explaining this behavior: the initial DSB yields in this study are relatively lower than the previous ones obtained with Geant4-DNA, thus the normalized foci yields after the peak are apparently higher. In addition, the work of Belov et al. (2015) estimated the irreparable fraction as 0.01 as Asaithamby et al. (2008). We could thus simulate the biological stage, based on this repair model, with a good agreement with the experimental results within 10% difference. However, there are still some limitations of such analytical approach, for example the model were optimized for only human fibroblasts. Also, such model is not an absolute prediction model, because the model requires input with total number of DSBs and the irreparable fraction.

5.9. Conclusions

In this study, we implemented our IRT approach into the "molecularDNA" example, allowing to overcome the computational burden of radiolysis simulation. The control of the implementation has been verified with a simple geometry test. Then, we applied the developments proposed in this thesis - such as electron elastic scattering model, pre-chemical and chemical parameters, and bug fixes on displacement of •OH radicals and electron attachment model - for the simulation of early DNA damage in a simplified human fibroblast cell nucleus. We showed that the results obtained with those improvements are in reasonable agreement with recent experimental data on DSB yields as a function of LET, with acceptable simulation times. In order to further evaluate the plausibility of our simulations, we also presented various quantities such as SSB/DSB ratio as a function of LET, the histogram of fragment lengths, the scavengeable fraction as a function of LET, and repair of foci as a function of time. All underline the need for more accurate experimental data on DNA damage in irradiated cells.

There are some technical limitations remaining, such as large memory consumption (up to ~ 160 Go for only 100 high LET particles). To solve such issues, we have split simulation runs in order to prevent crashes due to the large memory consumption. In the future, technical optimization of software will be followed.

References

Asaithamby, A., Uematsu, N., Chatterjee, A., et al. 2008, Repair of HZE-particle-induced DNA double-strand breaks in normal human fibroblasts. *Radiat. Res.,* 169**,** 437-46.

Belli, M., Cherubini, R., Dalla Vecchia, M., et al. 2000, DNA DSB induction and rejoining in V79 cells irradiated with light ions: a constant field gel electrophoresis study. *Int. J. Radiat. Biol.,* 76**,** 1095-104.

Belli, M., Cherubini, R., Vecchia, D., et al. 2001, DNA fragmentation in mammalian cells exposed to various light ions. *Adv. Space Res.,* 27**,** 393-399.

Belov, O. V., Krasavin, E. A., Lyashko, M. S., et al. 2015, A quantitative model of the major pathways for radiation-induced DNA double-strand break repair. *J. Theor. Biol.,* 366**,** 115-30.

Boudaïffa, B., Cloutier, P., Hunting, D., et al. 2000, Resonant formation of DNA strand breaks by low-energy (3 to 20 eV) electrons. *Science,* 287**,** 1658-1660.

Buxton, G. V., Greenstock, C. L., Helman, W. P., et al. 1988, Critical review of rate constants for reactions of hydrated electrons, hydrogen atoms and hydroxyl radicals (·OH/·O− in Aqueous Solution. *J. Phys. Chem. Ref. Data,* 17**,** 513-886.

Campa, A., Ballarini, F., Belli, M., et al. 2005, DNA DSB induced in human cells by charged particles and gamma rays: experimental results and theoretical approaches. *Int. J. Radiat. Biol.,* 81**,** 841-54.

Chang, D. S., Lasley, F. D., Das, I. J., et al. 2014. Oxygen Effect, Relative Biological Effectiveness and Linear Energy Transfer. *Basic Radiotherapy Physics and Biology.* Springer, Cham.

Charlton, D. E., Nikjoo, H. & Humm, J. L. 1989, Calculation of initial yields of single- and double-strand breaks in cell nuclei from electrons, protons and alpha particles. *Int. J. Radiat. Biol.,* 56**,** 1-19.

Crick, F. H. C., Barnett, L., Brenner, S., et al. 1961, General nature of the genetic code for proteins. *Nature,* 192**,** 1227-1232.

Daly, M. J. 2012, Death by protein damage in irradiated cells. *DNA Repair (Amst),* 11**,** 12-21.

Decottignies, A. 2013, Alternative end-joining mechanisms: a historical perspective. *Front. Genet.,* 4**,** 48.

Finkel, R., Friedman, J. & Bentley, J. 1977, An algorithm for finding best matches in logarithmic expected time. *ACM T. Math. Software***,** 200-226.

Frankenberg, D., Brede, H. J., Schrewe, U. J., et al. 1999, Induction of DNA double-strand breaks by 1H and 4He ions in primary human skin fibroblasts in the LET range of 8 to 124 keV/μm. *Radiat. Res.,* 151**,** 540-549.

Frankenberg-Schwager, M. 1989, Review of repair kinetics for DNA damage induced in eukaryotic cells in vitro by ionizing radiation. *Radiother. Oncol.,* 14**,** 307-320.

Friedland, W., Dingfelder, M., Kundrat, P., et al. 2011, Track structures, DNA targets and radiation effects in the biophysical Monte Carlo simulation code PARTRAC. *Mutat. Res.,* 711**,** 28-40.

Friedland, W., Jacob, P., Bernhardt, P., et al. 2003, Simulation of DNA Damage after Proton Irradiation. *Radiat. Res.,* 159**,** 401-410.

Goodhead, D. T. 1988, Spatial and temporal distribution of energy. *Health Phys.,* 55**,** 231-240.

Hall, E. J. & Giaccia, A. J. 2018. *Radiobiology for the Radiologist*, Lippincott Williams & Wilkins.

Heyer, W. D., Ehmsen, K. T. & Liu, J. 2010, Regulation of homologous recombination in eukaryotes. *Annu. Rev. Genet.,* 44**,** 113-39.

Hilbert, D. 1935. Über die stetige Abbildung einer Linie auf ein Flächenstück. *Dritter Band: Analysis· Grundlagen der Mathematik· Physik Verschiedenes.* Springer.

Hirayama, R., Ito, A., Tomita, M., et al. 2009, Contributions of direct and indirect actions in cell killing by high-LET radiations. *Radiat. Res.,* 171**,** 212-8.

Hoglund, E., Blomquist, E., Carlsson, J., et al. 2000, DNA damage induced by radiation of different linear energy transfer: initial fragmentation. *Int. J. Radiat. Biol.,* 76**,** 539-47.

ICRU 2014. Key data for ionizing-radiation dosimetry: measurement standards and applications. *ICRU Report 90.* Journal of the ICRU: The International Commission on Radiation Units and Measurements.

Iliakis, G. E., Cicilioni, O. & Metzger, L. 2009, Measurement of DNA Double-strand Breaks in CHO Cells at Various Stages of the Cell Cycle Using Pulsed Field Gel Electrophoresis: Calibration by Means of125I Decay. *Int. J. Radiat. Biol.,* 59**,** 343-357.

Ito, A., Nakano, H., Kusano, Y., et al. 2006, Contribution of indirect action to radiation-induced mammalian cell inactivation: dependence on photon energy and heavy-ion LET. *Radiat. Res.,* 165**,** 703-12.

Karamitros, M., Brown, J. M. C., Lampe, N., et al. 2020, Implementing the independent reaction time method in Geant4 for radiation chemistry simulations. *arXiv preprint arXiv:2006.14225*.

Kreipl, M. S., Friedland, W. & Paretzke, H. G. 2009, Time- and space-resolved Monte Carlo study of water radiolysis for photon, electron and ion irradiation. *Radiat. Environ. Bioph.,* 48**,** 11-20.

Kryndushkin, D. S., Alexandrov, I. M., Ter-Avanesyan, M. D., et al. 2003, Yeast [PSI+] Prion Aggregates Are Formed by Small Sup35 Polymers Fragmented by Hsp104. *J. Biol. Chem.,* 278**,** 49636-49643.

Kyriakou, I., Šefl, M., Nourry, V., et al. 2016, The impact of new Geant4-DNA cross section models on electron track structure simulations in liquid water. *J. Appl. Phys.,* 119.

Lampe, N. 2017. *The long term impact of ionising radiation on living systems.* Ph.D. thesis, Université Clermont Auvergne, France.

Lampe, N., Karamitros, M., Breton, V., et al. 2018a, Mechanistic DNA damage simulations in Geant4-DNA part 1: A parameter study in a simplified geometry. *Phys. Med.,* 48**,** 135-145.

Lampe, N., Karamitros, M., Breton, V., et al. 2018b, Mechanistic DNA damage simulations in Geant4-DNA Part 2: Electron and proton damage in a bacterial cell. *Phys. Med.,* 48**,** 146-155.

Lea, D. E. 1955. *Actions of radiations on living cells,* UK, Cambridge University Press.

Leloup, C., Garty, G., Assaf, G., et al. 2005, Evaluation of lesion clustering in irradiated plasmid DNA. *Int. J. Radiat. Biol.,* 81**,** 41-54.

Ljungman, M., Nyberg, S., Nygren, J., et al. 1991, DNA-bound proteins contribute much more than soluble intracellular compounds to the intrinsic protection against radiation-induced DNA strand breaks in human cells. *Radiat. Res.,* 127**,** 171-176.

Löbrich, M., Shibata, A., Beucher, A., et al. 2010, γH2AX foci analysis for monitoring DNA double-strand break repair: strengths, limitations and optimization. *Cell Cycle,* 9**,** 662-669.

Luger, K., Mäder, A. W., Richmond, R. K., et al. 1997, Crystal structure of the nucleosome core particle at 2.8 Å resolution. *Nature,* 389**,** 251-260.

McNamara, A., Geng, C., Turner, R., et al. 2017, Validation of the radiobiology toolkit TOPAS-nBio in simple DNA geometries. *Phys. Med.,* 33**,** 207-215.

Mehnati, P., Morimoto, S., Yatagai, F., et al. 2005, Exploration of 'Over kill effect' of high-LET Ar- and Fe-ions by evaluating the fraction of non-hit cell and interphase death. *J. Radiat. Res.,* 46**,** 343-350.

Meylan, S., Incerti, S., Karamitros, M., et al. 2017, Simulation of early DNA damage after the irradiation of a fibroblast cell nucleus using Geant4-DNA. *Sci. Rep.,* 7**,** 11923.

Meylan, S., Vimont, U., Incerti, S., et al. 2016, Geant4-DNA simulations using complex DNA geometries generated by the DnaFabric tool. *Comput. Phys. Commun.,* 204**,** 159-169.

Negritto, M. C. 2010, Reparing double-strand DNA breaks. *Nat. Educ.,* 3**,** 26.

Newman, H. C., Prise, K. M. & Michael, B. D. 2000, The role of higher-order chromatin structure in the yield and distribution of DNA double-strand breaks in cells irradiated with X-rays or alpha-particles. *Int. J. Radiat. Biol.,* 76**,** 1085-93.

Nias, A. H. W. 1998. *An introduction to radiobiology,* Chichester ;New York :, Wiley.

Nikjoo, H., Emfietzoglou, D., Liamsuwan, T., et al. 2016, Radiation track, DNA damage and response-a review. *Rep. Prog. Phys.,* 79**,** 116601.

Nikjoo, H., Martin, R. F., Charlton, D. E., et al. 1996, Modelling of Auger-induced DNA damage by incorporated 125I. *Acta Oncol.,* 35**,** 849-56.

Nikjoo, H., O'Neill, P., Goodhead, D. T., et al. 1997, Computational modelling of low-energy electron-induced DNA damage by early physical and chemical events. *Int. J. Radiat. Biol.,* 71**,** 467-483.

Nikjoo, H., O'Neill, P., Wilson, W. E., et al. 2001, Computational approach for determining the spectrum of DNA damage induced by ionizing radiation. *Radiat. Res.,* 156**,** 577-583.

Olive, P. L. 2004, Detection of DNA damage in individual cells by analysis of histone H2AX phosphorylation. *Methods in Cell Biology,* 75**,** 355-373.

Petkovic, V. D., Keta, O. D., Vidosavljevic, M. Z., et al. 2019, Biological outcomes of gamma-radiation induced DNA damages in breast and lung cancer cells pretreated with free radical scavengers. *Int. J. Radiat. Biol.,* 95**,** 274-285.

Pinto, M., Newman, H. C., Prise, K. M., et al. 2000, Quantification of DNA damage by PFGE: development of an analytical approach to correct for the background distribution. *Int. J. Radiat. Biol.,* 76**,** 741-8.

Plante, I. & Devroye, L. 2017, Considerations for the independent reaction times and step-by-step methods for radiation chemistry simulations. *Radiat. Phys. Chem.,* 139**,** 157-172.

Plante, I., Ponomarev, A., Patel, Z., et al. 2019, RITCARD: Radiation-Induced Tracks, Chromosome Aberrations, Repair and Damage. *Radiat. Res.,* 192**,** 282-298.

Prise, K. M., Folkard, M., Michael, B. D., et al. 2000, 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. *Int. J. Radiat. Biol.,* 76**,** 881-90.

Ramos-Mendez, J., Shin, W. G., Dominguez-Kondo, J. N., et al. 2020, Independent reaction times method in Geant4-DNA: implementation and performance. *Med. Phys.,* 47**,** 5919-5930.

Ristić Fira, A. & Petrovic, I. 2020, Private communication. *Private communication, Data not yet published.*

Rothkamm, K., Barnard, S., Moquet, J., et al. 2015, DNA damage foci: Meaning and significance. *Environ. Mol. Mutagen.,* 56**,** 491-504.

Sakata, D., Belov, O., Bordage, M. C., et al. 2020, Fully integrated Monte Carlo simulation for evaluating radiation induced DNA damage and subsequent repair using Geant4-DNA. *Sci. Rep.,* 10**,** 20788.

Sakata, D., Lampe, N., Karamitros, M., et al. 2019, Evaluation of early radiation DNA damage in a fractal cell nucleus model using Geant4-DNA. *Phys. Med.,* 62**,** 152-157.

Shin, W. G., Bordage, M. C., Emfietzoglou, D., et al. 2018, Development of a new Geant4-DNA electron elastic scattering model for liquid-phase water using the ELSEPA code. *J. Appl. Phys.,* 124.

Shin, W. G., Ramos-Mendez, J., Faddegon, B., et al. 2019, Evaluation of the influence of physical and chemical parameters on water radiolysis simulations under MeV electron irradiation using Geant4-DNA. *J. Appl. Phys.,* 126.

Shu, J. J. 2017, A new integrated symmetrical table for genetic codes. *Biosystems,* 151**,** 21-26.

Suzuki, S., Yamanashi, T., Tazawa, S., et al. 1998, Quantitative analysis of DNA orientation in stationary AC electric fields using fluorescence anisotropy. *IEEE T. Ind. Appl.,* 34**,** 75-83.

Tang, N. 2019. *Évaluation, à partir de modélisations nanodosimétriques, de l'influence de la compaction de la chromatine sur les effets radio-induits précoces et extension aux effets tardifs (réparation des dommages à l'ADN et mort cellulaire).* Ph.D. thesis, Université de Bordeaux, France.

Tang, N., Bueno, M., Meylan, S., et al. 2019, Influence of chromatin compaction on simulated early radiation-induced DNA damage using Geant4-DNA. *Med. Phys.,* 46**,** 1501-1511.

Uehara, S., Nikjoo, H. & Goodhead, D. T. 1993, Cross-sections for water vapour for the Monte Carlo electron track structure code from 10 eV to the MeV region. *Phys. Med. Biol.,* 37**,** 1841-1858.

Valentin, J. 2016, Relative biological effectiveness (RBE), quality factor (Q), and radiation weighting factor (wR). *Annals of the ICRP,* 92**,** 1-121.

Villagrasa, C., Meylan, S., Gonon, G., et al. 2017, Geant4-DNA simulation of DNA damage caused by direct and indirect radiation effects and comparison with biological data. *EPJ Web Conf.,* 153.

Watson, J. D. & Crick, F. H. C. 1953, A structure for deoxyribose nucleic acid. *Nature,* 171**,** 737-738.

Zhong, S., Ghosh, S. K. & Jost, D. 2018, How epigenome drives chromatin folding and dynamics, insights from efficient coarse-grained models of chromosomes. *PLOS Comput. Biol.,* 14.

Zhu, H., McNamara, A. L., Ramos-Mendez, J., et al. 2020, A parameter sensitivity study for simulating DNA damage after proton irradiation using TOPAS-nBio. *Phys. Med. Biol.,* 65**,** 085015.

Conclusions & perspectives

This thesis mainly focuses on the improvement of the Geant4-DNA toolkit, especially in order:

1. to improve the physics and chemistry models allowing to more accurately predict radiobiological effects of ionizing radiation at the sub-cellular scale,
2. to develop an integrated chain platform which simultaneously simulates all the stages benefiting from those improvements.

For that, in Chapter 2, a new electron elastic scattering model for liquid-phase water was developed using the ELSEPA Dirac partial-wave code (Salvat et al., 2005). This elastic model has been validated by comparing with experimental data and with the other models available in Geant4-DNA. The impact of the new elastic model on physics simulations was evaluated using 5 Geant4-DNA examples, such as *range*, *mfp* (mean-free-path), *TestEm12* (for dose-point-kernels), *microyz* (for microdosimetric distributions), and *clustering* (for a rapid estimation of direct DNA damage). I have shown that the influence of this new elastic scattering model on the physical stage is almost negligible compared to the other physics models, because the energy depositions of elastic scattering are neglected. The model has been delivered in Geant4 10.6 in Dec. 2019.

However, the spatial distribution of secondary electrons might affect the chemical stage. Thus, we reviewed in Chapter 3 the details of the simulation of the pre-chemical stage in Geant4-DNA and in other MCTS tools. At first, the influence of the spatial distribution of molecular species was evaluated when changing the electron elastic scattering model (developed in Chapter 2), the electron thermalization models, and the chemical parameters used for molecular species (such as diffusion coefficients and reaction rates). In addition, Auger effect, electron molecular attachment, electron-hole recombination, and dissociation channels were improved based on original papers of the literature. I concluded that a combination of the ELSEPA elastic scattering model (Shin et al., 2018), the electron thermalization model (Meesungnoen et al., 2002), and the chemistry constructor based on the literature (Cobut et al., 1998) successfully reproduces the experimental radiochemical yields, especially low G-values of hydroxyl radical recently published (Wang et al., 2018). The electron thermalization model has been released in Geant4 10.6 in Dec. 2019 and the new pre-chemical model including physico-chemical interactions and the dissociation channels will be delivered in Geant4 in 2021.

In Chapter 4, the IRT approach was implemented into Geant4-DNA in order to reduce the computational time of chemical stage. A validation study on calculation efficiency and radiochemical yields was performed. This implementation showed good agreement with experimental data, and the simulation of chemical stage is faster than the SBS method by about 1,000 times for low LET incident particle. This IRT implementation has been delivered in Geant4 10.7 beta in Jun. 2020.

At last, in Chapter 5, the influence of all the developments presented in this thesis on DNA damage prediction was evaluated using the *molecularDNA* example, including a simplified human fibroblast cell geometric model. The results for gamma, proton, alpha particles were compared with the previous Geant4-DNA simulations and data from the literature. The results, in particular DSB yields as a function of LET, for human fibroblast cell nucleus are closer to recent experimental data (Campa et al., 2005, Ristić Fira and Petrovic, 2020) than the other MCTS tools (Nikjoo et al., 1997, Friedland et al., 2003), even than the previous Geant4-DNA studies (Meylan et al., 2017, Sakata et al., 2019, Sakata et al., 2020). The full release of this example in Geant4 is expected in 2021.

With the improvements described in this thesis, the agreements with experimental data have been improved at all stages (physics, chemistry and DNA damage induction).

Geant4-DNA is the first fully open access MCTS toolkit available freely to the community, being fully included in Geant4. Thus, users can easily benefit from the developments described in this thesis, and they can verify their simulation results independently and ensure their robustness. Also, these improvements could directly help Geant4-based codes such as TOPAS-nBio and GATE.

This thesis enabled to mechanistically simulate DNA damage induced by ionizing radiation using Geant4-DNA in a simplified human cell nucleus, with a close agreement with recent experimental data. Such a simulation platform, based on the *molecularDNA* example, including the proposed developments, could help several research fields where the understanding of ionizing radiation effects at the DNA scale is crucial. For example, Geant4-DNA could be employed to calculate relative biological effect in novel approaches for radiation therapy (e.g. mini/microbeam radiotherapy, FLASH radiotherapy, targeted radiotherapy), and in specific radiation environments (e.g. chronic exposure to ionizing radiation in space).

As further steps, the radiobiological response according to the structure of the geometrical cell model (e.g. fractal and rosette structures…) could be evaluated in order to shape more “realistic” human cell geometrical models.

In-depth study of repair model is also necessary. For instance, the repair model used in this thesis roughly assumed that all the complex DSBs are never repaired. A prediction model for complex DSB repair could be implemented into Geant4-DNA (Stewart, 2001).

The evaluation of radiobiological effects for heavy ions (e.g. oxygen and carbon) could be investigated for recent radiotherapy techniques such as Carbon therapy. For that, however, the ion cross-section models for MCTS simulation should be further developed (e.g. charge exchange process is neglected, as well as excitation processes).

One could investigate the possibility to go beyond the single cell scale, and try to simulate multi-cell organisms (e.g. C. elegans), microtumors or small fragments of tissue. For that, the modeling of various geometries would be needed (e.g. assembly of cells). However, the simulation time of chemical stage using IRT method may still be a huge burden. This limitation might be resolved by using GPU processors (Okada et al., 2019) and analytical calculation of chemical stage.

In addition, the construction of damage database according to the LET and cell type may allow to avoid additional lengthy MCTS simulation and might be of interest for the radiobiology community.

References

Campa, A., Ballarini, F., Belli, M., et al. 2005, DNA DSB induced in human cells by charged particles and gamma rays: experimental results and theoretical approaches. *Int. J. Radiat. Biol.,* 81**,** 841-54.

Cobut, V., Frongillo, Y., Patau, J. P., et al. 1998, Monte Carlo simulation of fast electron and proton tracks in liquid water-I. Physical and physicochemical aspects. *Radiat. Phys. Chem.,* 51**,** 229-243.

Friedland, W., Jacob, P., Bernhardt, P., et al. 2003, Simulation of DNA Damage after Proton Irradiation. *Radiat. Res.,* 159**,** 401-410.

Meesungnoen, J., Jay-Gerin, J.-P., Filali-Mouhim, A., et al. 2002, Low-energy electron penetration range in liquid water. *Radiat. Res.,* 158**,** 657-660.

Meylan, S., Incerti, S., Karamitros, M., et al. 2017, Simulation of early DNA damage after the irradiation of a fibroblast cell nucleus using Geant4-DNA. *Sci. Rep.,* 7**,** 11923.

Nikjoo, H., O'Neill, P., Goodhead, D. T., et al. 1997, Computational modelling of low-energy electron-induced DNA damage by early physical and chemical events. *Int. J. Radiat. Biol.,* 71**,** 467-483.

Okada, S., Murakami, K., Incerti, S., et al. 2019, MPEXS-DNA, a new GPU-based Monte Carlo simulator for track structures and radiation chemistry at subcellular scale. *Med. Phys.,* 46**,** 1483-1500.

Ristić Fira, A. & Petrovic, I. 2020, Private communication. *Private communication, Data not yet published.*

Sakata, D., Belov, O., Bordage, M. C., et al. 2020, Fully integrated Monte Carlo simulation for evaluating radiation induced DNA damage and subsequent repair using Geant4-DNA. *Sci. Rep.***,** in press.

Sakata, D., Lampe, N., Karamitros, M., et al. 2019, Evaluation of early radiation DNA damage in a fractal cell nucleus model using Geant4-DNA. *Phys. Med.,* 62**,** 152-157.

Salvat, F., Jablonski, A. & Powell, C. J. 2005, ELSEPA - Dirac partial-wave calculation of elastic scattering of electrons and positrons by atoms, positive ions and molecules. *Comput. Phys. Commun.,* 165**,** 157-190.

Shin, W. G., Bordage, M. C., Emfietzoglou, D., et al. 2018, Development of a new Geant4-DNA electron elastic scattering model for liquid-phase water using the ELSEPA code. *J. Appl. Phys.,* 124.

Stewart, R. D. 2001, Two-lesion kinetic model of double-strand break rejoining and cell killing. *Radiat. Res.,* 156**,** 365-378.

Wang, F., Schmidhammer, U., Larbre, J. P., et al. 2018, Time-dependent yield of the hydrated electron and the hydroxyl radical in D2O: a picosecond pulse radiolysis study. *Phys. Chem. Chem. Phys.*