UNIVERSIDADE ESTADUAL DE CAMPINAS

IFGW - Instituto de Física Gleb Wataghin

John James Tello Cajiao

The influence of the DNA conformation on the radiation-induced DNA damage probabilities.

A influência da conformação do DNA nas probabilidades de dano induzido por radiações.

Campinas

2016

John James Tello Cajiao

The influence of the DNA conformation on the radiation-induced DNA damage probabilities.

A influência da conformação do DNA nas probabilidades de dano induzido por radiações.

Dissertation presented to the Institute of Physics "Gleb Wataghin" of the University of Campinas in partial fulfillment of the requirements for the degree of Master in Physics.

Dissertação apresentada ao Instituto de Física "Gleb Wataghin" da Universidade Estadual de Campinas como parte dos requisitos exigidos para a obtenção do título de Mestre em Física.

Supervisor: Dr. Mario Antonio Bernal Rodriguez

ESTE EXEMPLAR CORRESPONDE À VERSÃO FINAL DISSERTAÇÃO DEFENDIDA PELO ALUNO JOHN JAMES TELLO CAJIAO, E ORIENTADA PELO PROF DR. MARIO ANTONIO BERNAL RODRGUEZ.

Campinas

2016

Agência(s) de fomento e nº(s) de processo(s): CAPES, 1370449/2014

Ficha catalográfica
Universidade Estadual de Campinas
Biblioteca do Instituto de Física Gleb Wataghin
Lucimeire de Oliveira Silva da Rocha - CRB 8/9174

T239i

Tello Cajiao, John James, 1990-
The influence of the DNA conformation on the radiation-induced DNA damage probabilities / John James Tello Cajiao. – Campinas, SP : [s.n.], 2016.

Orientador: Mario Antonio Bernal Rodriguez.
Dissertação (mestrado) – Universidade Estadual de Campinas, Instituto de Física Gleb Wataghin.

1. Radiobiologia. 2. DNA. 3. Monte Carlo, Método de. 4. Modelo linear-quadrático. I. Bernal Rodriguez, Mario Antonio,1972-. II. Universidade Estadual de Campinas. Instituto de Física Gleb Wataghin. III. Título.

Informações para Biblioteca Digital

Título em outro idioma: A influência da conformação do DNA nas probabilidades de dano induzido por radiações
Palavras-chave em inglês:
Radiobiology
DNA
Monte Carlo method
Linear-quadratic model
Área de concentração: Física
Titulação: Mestre em Física
Banca examinadora:
Mario Antonio Bernal Rodriguez [Orientador]
Marcelo Baptista de Freitas
Sandro Guedes de Oliveira
Data de defesa: 30-06-2016
Programa de Pós-Graduação: Física

MEMBROS DA COMISSÃO JULGADORA DA DISSERTAÇÃO DE MESTRADO DE **JOHN JAMES TELLO CAJIAO – RA: 153910** APRESENTADA E APROVADA AO INSTITUTO DE FÍSICA “GLEB WATAGHIN”, DA UNIVERSIDADE ESTADUAL DE CAMPINAS, EM 30/06/2016.

COMISSÃO JULGADORA:

- Prof. Dr. Mario Antonio Bernal Rodriguez – (Orientador) – DFA/IFGW/UNICAMP
- Prof. Dr. Marcelo Baptista de Freitas – EPM/UNIFESP
- Prof. Dr. Sandro Guedes de Oliveira – DRCC/IFGW/UNICAMP

A Ata de Defesa, assinada pelos membros da Comissão Examinadora, consta no processo de vida acadêmica do aluno.

CAMPINAS
2016

To my beautiful daughter Isabella.

Acknowledgements

Firstly, I would like to thank to my supervisor Dr. Mario B. Rodriguez. He consistently allowed this paper to be my own work, but steered me in the right direction whenever he thought I needed it. Not only his academic guidance but also his generosity and kindness throughout the project, count for me as lessons as well.

Also, I want to thank to many friends with whom I have discussed about this project. Those discussions have helped me to clarify myself and keep working. So thanks to Maria E. Tello, Daniel P., Andres Q., Pedro S., Alejandro T., Angie A. and specially to my best friend Christian T. for bring me to this field of the applied physics.

I must express my very profound gratitude to my parents James and Carmenza for providing me with unfailing support and continuous encouragement throughout my years of study and through the process of researching and writing this thesis. This accomplishment would not have been possible without them. Thank you.

Finally, I appreciate the financial support from CAPES that funded this project along its execution.

"Aquellos que eran vistos bailando,
eran considerados locos por quienes
no podían escuchar la música".
Friedrich Nietzsche.

Abstract

The aim of this work is to study the influence of the DNA conformation on the probability of direct damage induction by ionizing particles. Also, the mechanistic grounds of the Linear-Quadratic radiobiological model are investigated through the eyes of a home-made biophysical model based on the DRAT (Dual Radiation Action Theory). To this end, three geometrical models of the genetic material were constructed. The models have atomistic resolution and account for $\sim 10^9$ base pairs (bps) in the A-, B- and Z-DNA configurations. Starting from a single bp, the different organizational levels inside the cell nucleus were created by means of linear transformations. Next, the Monte Carlo (MC) code *GEANT4-DNA* was used to simulate the transport of protons of $0.5, 1, 5, 7$ and $10\ MeV$, and alpha particles of $2, 5, 7$ and $10\ MeV$. The number of particles in each case is such that the absorbed doses range between $0.5\ Gy$ and $16\ Gy$.

The three models proved to be consistent with the dimensions of the real structures. In particular, the models were compatible with the $30\ nm$ chromatin fiber diameter requirement as well as with the bp volumes reported in other works. The Total and Double Strand Break Yields (TSBY and DSBY) were obtained for every radiation quality. Also, the Site-Hit Probability (SHP) defined as the total target to the nucleus volume ratio, was computed theoretically and from the simulations. The biophysical model in conjunction with the MC simulations furnished the number of lethal lesions (N_{LL}) as a function of dose, for protons of 0.5 and $10MeV$, and for alpha particles of 2 and $10MeV$. The N_{LL} could be split into those created by a single track and those originated by interaction of two tracks.

It is concluded that the TSBY is practically determined by the SHP and depends weakly on the incident radiation quality. Nevertheless, the DSBY showed strong dependence on both the DNA conformation and the radiation quality. This is due to the interplay between the energy deposition clustering capacity of a given radiation and the DNA spatial packing. On the other hand, the analysis of the mechanisms of damage production based on the DRAT and tested with the biophysical model developed, showed that single-track (first order) effects depend linearly on the dose. Moreover, inter-track effects follows a quadratic behavior with the dose, having a linear term that influences the first order mechanism. This means that the Linear-Quadratic behavior of the N_{LL} with the dose, has mechanistic groundings at least at the first stage of the damage.

Key-words: Computational Radiobiology. DNA. Monte Carlo. DRAT. LQ-model

Resumo

O objetivo deste trabalho é estudar a influência da conformação do DNA na probabilidade de dano direto produzido por partículas ionizantes. Além disso, os fundamentos mecanicísticos do modelo Linear-Quadrático são investigadas através de um modelo biofísico desenvolvido neste trabalho, baseado na TADR (Teoria da Ação Dual da Radiação). Para este fim, três modelos geométricos do material genético foram construídos. Os modelos têm resolução atomística e levam em conta $\sim 10^9$ pares de base (bps) nas configurações A,B e Z do DNA. A partir de um único bp, os diferentes níveis organizacionais no interior do núcleo da célula foram criados por meio de transformações lineares. Em seguida, o código Monte Carlo (MC) GEANT4-DNA foi usado para simular o transporte de prótons de $0.5, 1, 5, 7$ e $10MeV$ assim como de partículas alfa de $2, 5, 7$ e $10MeV$. O número de partículas em cada caso é de tal modo que as doses absorvidas estão entre $0.5 - 16Gy$.

Os três modelos foram consistentes com as dimensões das estruturas reais. Em particular, os modelos foram compatíveis com a exigência de que o diâmetro da cromatina seja de 30 nm, bem como com os volumes bp reportados em outros trabalhos. Os rendimentos tanto das quebras totais quanto das quebras duplas (TSBY e DSBY) foram obtidos para cada qualidade de radiação. Além disso, a probabilidade de impacto (SHP) definida como a razão entre o volume do DNA e o volume núcleo, foi calculada teoricamente e a partir das simulações. O modelo biofísico em conjunto com as simulações MC forneceu o número de lesões letais (N_{LL}) em função da dose, para prótons de $0,5$ e $10MeV$, e para partículas alfa de 2 e $10MeV$. Os N_{LL} puderam ser divididos em aqueles criados por uma única trajetória e aqueles originados pela interacção de duas trajetórias.

Concluiu-se que o TSBY é praticamente determinada pela SHP e depende fracamente da qualidade de radiação incidente. No entanto, o DSBY mostrou forte dependência tanto da conformação do DNA quanto da qualidade de radiação. Isto é devido à relação entre a capacidade de agrupamento das deposições de energia para uma radiação dada e o empacotamento do DNA. Por outro lado, a análise dos mecanismos de produção de dano, baseada na TADR e testada com o modelo biofísico desenvolvido, mostraram que os efeitos de uma única trajetória (de primeira ordem) dependem linearmente com a dose. Além disso, os efeitos inter-trajetórias seguem um comportamento quadrático com a dose, com um termo linear que influencia o mecanismo de primeira ordem. Isto significa que o comportamento linear-quadrático do N_{LL} com a dose, tem fundamentos mecanicistas, pelo menos, na primeira fase do dano.

Palavras-chaves: Radiobiologia computacional. DNA. Monte Carlo. DRAT. modelo LQ

List of Figures

List of Tables

Contents

Introduction

Ionizing radiation induces damages in the genetic material of living beings. This damage may give rise to cancer (carcinogenesis) and, paradoxically, is used to treat cancer through radiotherapy as well. At first stage, the damage on DNA due to the ionizing radiation is physico-chemical and can be classified as direct or indirect. The former is caused directly by the impact of ionizing particles on DNA. The latter is generated by the attack of the reactive chemical species produced by the radiolisys of water, the most abundant compound in biological tissues. After the initial damage, a series of biochemical mechanisms start, which aim to repair the DNA damage. If such mechanisms do not succeed, the cell may die (apoptosis), be inactivated or transmit its damage to next generations.

It is well known that the interaction of radiation with matter is a stochastic process, and that cross sections can be used to quantify the probability of interaction when a particle passes through matter. The rapid development of powerful computers has promoted the use of simulations both in basic and applied research. Thus, the Monte Carlo (MC) method, initially introduced in the early 40's within the Manhattan Project, has became a fundamental tool in research related to medical applications of ionizing radiations. MC is based on the combination of probability functions, constructed from of the cross sections that describe the involved interaction processes, a pseudo-random number generator and a sampling method. The stochastic variable of interest is sampled using the corresponding probability function. The first step of this sampling procedure is the generation of a pseudo-random number, which represents the probability for observing the stochastic quantity in question. Sampling the certain quantities, the history of a particle can be constructed, including secondary particles. Thus, it is possible to simulate the transport of virtually any particle through any media using the MC method. The expected value of energy deposit, absorbed dose, mean distance traveled by a particle and so on, can be computed via MC.

At the beginning, the *GEANT4* MC code was developed at CERN to transport high energy particles. In the last decade, it has been extended down to lower energies, as those used to medical applications. Further extensions were developed in order to use this code in micro and nanodosimetry. This task was accomplihsed through the *GEANT4-DNA* European project, which aims to include the transport of heavy charged particles and electrons down to very low energies ($\sim$eV for electrons and $\sim$keV for ions). The group within which this dissertation project was carried out is part of the *GEANT4-DNA* project.

There are three main configurations of the DNA, namely, A-,B- and Z-DNA [3]. B-DNA is the most common configuration and it is found in the heterochromatin, a tightly packed form of DNA. A-DNA seems to play a role in the interaction of DNA with drugs. Less abundant but still present is Z-DNA. It is estimated that our genetic material contains about 10^5 copies of Z-DNA.

The MC method can be used in conjunction with DNA geometrical models in order to determine radiation-induced strand break probabilities. These probabilities can be used to estimate the relative biological effectiveness (RBE) of a given radiation quality (particle type and energy). Additionally, MC simulations combined with biophysical models of DNA damage may provide insight into the mechanisms of damage formation.

The main objectives of this work go through these lines of applications of the Monte Carlo method. On the one hand, to study the probability of radiation-induced damage in the genetic material relative to the incident particle energies not only in the standard configuration of the DNA (B-DNA) but also including the A- and Z-configurations as they are of importance from the radiobiological point of view as well. On the other hand, the step-by-step simulation capability of the *GEANT4-DNA* code allows to study the origin and behavior of the DNA lesions when combined with a suitable biophysical model. Thus, the aims of this work are: (1) To developed three geometrical models of the genetic material including the configurations of DNA aforementioned, (2) To study the influence of the DNA conformation in the probability of damage and (3) To analyze the mechanism of production of early damage in the genetic material due to ionizing radiation.

1 Radiobiology

What is it?

Radiobiology is the combination of two areas, namely, radiation physics and biology. Its goal is to study the effect of ionizing radiation in living systems, for instance, cells, biological tissues and organisms. The cell, as the smallest unit of protoplasm (organic and inorganic compounds dissolved in water) capable of independent existence, is the basic unit of all living beings. There are both organic and inorganic materials within a cell and its structure is basically divided into two main components: the cytoplasm, which supports all the metabolic functions of the cell, and the nucleus, which contains the genetic information (DNA). It is believed, more than 50 years ago, that the DNA is the sensitive target when ionizing radiation impacts on living organisms [4]. This is why it is so important to analyze the action of the radiation down to the molecular (DNA) level. However, it should be mentioned that this paradigm of Radiobiology, i.e. this target theory, fails to explain some observations, such as genomic instability [5] and bystander effects [6]. The study of these phenomena is very important for the understanding of radiation-induced damages and for developing corresponding study models [7, 8]. Yet, DNA is still considered as a critical target for radiation-induced damage in most of current approaches to study the radiobiological problem.

1.1 Ionizing Radiation

The term ionizing radiation (IR) refers to all kind of radiation capable to remove one or more electrons from a target (atom or molecule), leaving behind the target charged. There are several ways to produce IR. For instance, X-ray machines, cyclotrons, nuclear reactors and other devices. This radiation can be emitted by the decay of an unstable nucleus or by de-excitation of atoms and/or nuclei. Its origin can be extraterrestrial (e.g Cosmic Radiation), NORM (Naturally Occurring Radiation Material such as Uranium, Thorium and Radium) or artificial (e.g Medical facilities and Nuclear Reactors, among others).

Living beings are continuously exposed to natural IR from the outer space but also from the earth's crust, air, food and even from the human body itself. Other exposures may occur due to medical practices and occupational sources as nuclear power plants. Irrespective of its nature, ionizing radiation is able to provoke injuries to the cell due to its ability to ionize the atoms that compose it. The United Nations Scientific Committee on the Effects of Atomic Radiation published a summary of public and worker exposures to different sources of radiation (i.e natural, NORM, man-made) in 2008 [9]. The action

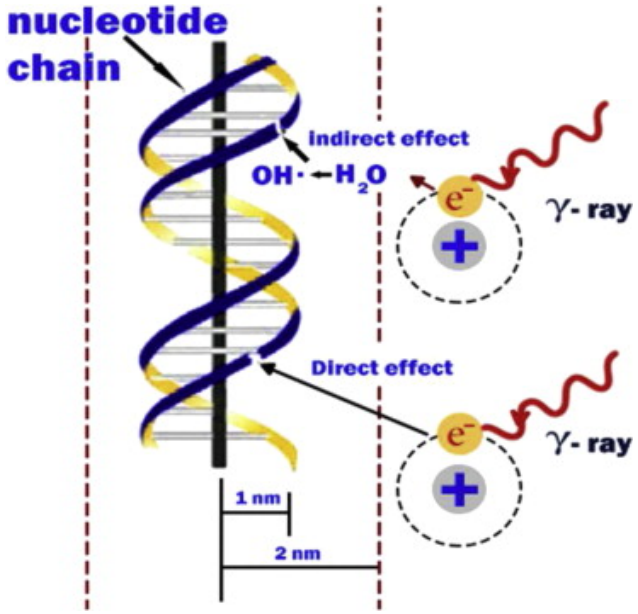


Figure 1 – Direct and Indirect action of the ionizing radiation on the genetic material. The very breakage of a bond that splits the molecule and provoke chemical reactions, constitutes a direct action. The creation of free radicals after the hit, that attack the DNA inducing changes, represents an indirect action of the radiation. Republished from [1] with permission by Elsevier.

of IR on DNA may be classified into direct and indirect. Fig. 1 depicts the direct and indirect mechanisms involved in the induction of DNA damage.

When the radiation has the sufficient energy to directly disrupt the atomic structure, it is called *directly ionizing radiation*. On the other hand, when radiation can not, by itself, alter the structure of the medium but is able to produce secondary particles (e.g secondary electrons) that do the chemical and biological damage, it is said that the radiation is *indirectly ionizing radiation*. According to its very nature, there exist two kinds of IR, namely, Electromagnetic (EM) and Particulate.

Electromagnetic radiation.
Electromagnetic radiation is nothing but photons. It is essentially characterized by its frequency (or wavelength) spectrum. For instance radio-waves, microwaves, infrared, visible light, ultraviolet, X rays, and γ rays. When EM radiation passes through a piece of matter, it may or may not interact with it. Biological effects occur when some energy is deposited in the cell or tissue. The main interactions between photons and matter are photoelectric absorption, Compton scattering, and pair production. Their cross sections depend on the atomic number of the target and on the incident energy of the photon.

In photoelectric absorption, the photon transfers all its energy to an inner-shell electron and the latter is ejected from the atom. This ejection produces a vacancy in the atom, which lead to a subsequent fluorescence emission or electron Auger emission. The kinetic energy of the electron emitted equals the incident photon energy minus the

binding energy of the involved shell. This is the dominant mechanism by which X-rays and γ-rays deposit energy in biological media below about 50 keV.

When a photon interacts with the outer electrons of a target, it is more probable that only a fraction of the photon energy be transferred to an atomic electron. This electron is ejected and the photon goes on with the remaining energy. This process is known as Compton or incoherent scattering. The atom involved in this process is left ionized. Typically, the ejected electron interacts with other atoms of the material and leads to an ionization cascade through the production of secondary electrons (also known as δ-rays). Compton scattering is the principal mechanism of interaction for X-rays and γ-rays from 100keV to 10MeV in biological media. This energy range is in the therapeutic radiation range.

Pair production refers to the process in which a photon with energy above 1.022 MeV interacts with the Coulomb field of an atomic nucleus, converting itself into a e-e^+ (electron-positron) pair. Positrons are very short-lived and they annihilate with an electron at the end of its range producing two new γ-rays of about 0.511 MeV. This process can also take place in the Coulomb field of an atomic electron. In this case, it is known as triplet-production because the involved electron is ejected together with the electron-positron pair produced. However, the threshold photon energy for triplet-production (in the atomic field) is twice higher than that for the pair-production process (in the nucleus field).

Normally, secondary particles produced in any of these processes have enough energy to cause further ionization in the medium. The effect of this secondaries is far more important from the biological point of view than that induced by the corresponding primary particles. This is why EM radiation is regarded as indirectly ionizing.

Particulate radiation.
A beam of particles may also cause damage to biological systems through ionization and excitation of the atoms of the traversed material. This kind of radiation can be split into two categories: charged and uncharged particles. In addition, charged particles can be classified as light and heavy particles.

Electrons and positrons are regarded as light charged particles while ions are considered as heavy particles. The amount of energy deposited by electrons per unit path length (see definitions below) is much lower than that for ions at the same energy. That is why electrons penetrate into a medium much more than heavy charged particles. Protons and α particles can be regarded as heavy charged particles. However, they can be considered as light ions as well.

Protons induce less DNA damage than α particles at the same energy because the former carry less charge and are lighter than the latter. Thus, protons deposit energy in the

medium at a lower rate than α particles. Electrons and positrons emitted from radioactive nuclei are known as β^-- and β^+-particles, respectively. Neutrons are uncharged particles so they are regarded as indirectly ionizing radiation. They interact with the nuclei of the medium through different processes that depend on their energy and target mass.

Ions are the result of pulling off one or more electrons from an atom. They offer advantages for cancer treatment over photon and electron beams, both from the dosimetric and radiobiological point of views. Energetic ions show a well defined range, just after a very sharp peak in the energy deposition distribution as a function of depth (Bragg peak). In addition, ion beams are more efficient to induce biological damage than electrons and photons.

1.2 Radiobiological Quantities of Interest

Absorbed Dose, Equivalent Dose, LET and RBE.
The *Absorbed Dose* is the quantity defined as the energy locally imparted (or deposited) per unit medium mass. The Gray (Gy) is the S.I. unit for the absorbed dose and 1 Gy is equivalent to 1 J/kg.

There are other units also used in Radiobiology, namely, rad, Sv and R (see Table 1 and Ref. [10]).

Table 1 – Summary of radiation doses and units.

Dose	SI Unit	Old Unit	Conversion Factor
Exposure	C/Kg air	Roentgen	1R=2.58×10^{-4} C/Kg air
Absorbed Dose	gray (Gy)	rad	100 rad=1 Gy
Equivalent Dose	sievert (Sv)	rem	100 rem=1 Sv

As will be explained later, not all radiations show the same efficiency to induce biological damage. This is why the quantity *Equivalent Dose* is frequently used in Radiation Protection, to compare the biological effectiveness of different radiation qualities.

In their passage through matter, IRs deposit energy along their path due to various interaction processes. The energy loss rate in a particular material depends not only on the energy and type of radiation but also on the density of the medium. The Linear Energy Transfer (LET) is the expected value of the energy transferred by the projectile to matter in electronic collisions per unit path length. These collisions include excitations and ionisations. The unit commonly used to express LET is $keV/\mu m$. For charged particles at a given energy, LET is defined as the quotient $-dE/dl$ where dE is the average energy transferred to the medium when the particle has traveled a distance dl. In the case of photons, the secondary particles created in the different interactions are those that deposit

energy in the medium. Thus, the LET of photon beams is defined by the LET of secondary electrons. In Radiobiology, LET is considered a measure of the quality of the IR since the biological effect of radiation depends on it (but is not totally determined by it). Radiations with different qualities may show different potential to induce biological damage. In this way, the Relative Biological Effectiveness (RBE) becomes the quantity of interest when comparing ionizing radiations from the radiobiological point of view. RBE quantifies the capacity of a given radiation quality to produce a certain biological effect when compared to a reference radiation (^{60}Co is the common reference) [11].

$$RBE = \left(\frac{Dose\ from\ standard\ radiation}{Dose\ from\ test\ radiation} \right)_{given\ biological\ effect} \tag{1.1}$$

As mentioned before, LET serves to classify the ionizing radiation. The lower the LET the sparser the energy deposition distribution and conversely, the higher the LET the higher the energy deposition density. On the one hand, X- and γ-rays are regarded as low LET particles, as well as electrons. On the other hand, ions and neutrons are considered high LET particles. In radiobiology, the frontier between low and high LET radiation is about 1 $keV/\mu m$. At first approximation, RBE increases with LET. However, RBE shows a peak at around 100 $keV/\mu m$ LET [12]. Above this LET value, RBE decreases because DNA damage yields show a saturation effect. This is termed as overkilling. IRs can be identified as densely or sparsely ionizing according to the spatial distribution of the energy depositions along their track. For instance, high LET radiations have the ability to deliver energy along their track in regions with characteristic sizes similar to those of the DNA dimensions. On the other hand, for low LET radiations the deposition sites are sparsely distributed relative to bio-molecule's dimensions $\sim nm$. Thus, high LET radiations are considered densely ionizing while low LET radiations are regarded as sparsely ionizing.

1.3 Types of Damage

Ionizations and excitations of the atoms and/or molecules of a biological system are followed by a series of chemical reactions among the species created in the processes and the ones presented in the medium. The ionized molecules may turn into free radicals in just $10^{-15}s$ to $10^{-10}s$. Later, free radicals can interact with neighbour molecules starting chain-reactions that may damage the cell. In lipids, these chain-reactions are related to damage in the cell membrane.

The very breakage of a chemical bound in a bio-molecule (BioM) due to IR represents a direct damage. This kind of damage occurs when chain reactions such as cross

linking, follow the radiation interaction.

$$IR + BioM \rightarrow R^* + H^* \tag{1.2}$$

Which can undergo, for instance,

$$R^* + R^* \rightarrow R^* - R^* \qquad (Cross\ linking) \tag{1.3}$$

R^* stands for *Radical*. H^* can react with other molecules and radicals as well. For high LET radiation the dominant effects are of this kind, i.e direct damage.

The indirect damage starts with the radiolisys of water, the most abundant component in cell. When IR impacts a water molecule, H_2O^+ radicals along with electronically excited water molecules are produced. Those products rapidly decompose to OH^* and H^* radicals through various chemical processes. Fast electrons may also be produced in IR-water interactions. These electrons thermalize and are solvated afterwards by dielectric interactions with the surrounding water molecules to give e^-_{aq} radicals. The e^-_{aq} will produce more H^* radicals when interacting with water. All these radicals, namely, H^*, OH^* and e^-_{aq} have high reactivity to cells, DNA and lipids.

2 Genetic Material

The basic component of the genetic material is the DNA (Deoxy-riboNucleic Acid) molecule which is nothing but different types of paired chemical bases: Adenine (A) with Thymine (T) and Guanine (G) with Cytosine (C), attached to two sugar-phosphate groups. A base pair attached to two sugar-phosphate groups is known as*Nucleotide*. All the instructions that control cell function are coded in the DNA. Genes are long genetic material sequences that contain specific information for the organism functioning, such as that used for protein synthesis. Genes vary in extent and content and are the fundamental objects of study of *genetics* [13]. DNA is present in the mitochondria, cytoplasm and cell nucleus. Its crucial importance to living organisms started to be understood when, in the early 40's, DNA was proposed as the carrier of the genetic information. The work of James Watson and Francis Crick [14] confirmed this hypotheses showing how genetic information could be stored in and copied from this molecule.

Back to the XIX century, chromosomes had been discovered in the nucleus of eukaryotic cells. Their structure and function were better understood once the DNA structure was unveiled. Chromosomes are long strands of DNA with associated proteins that bend and pack the genetic material in order to fit it into the nucleus dimension. This $DNA + packingproteins$ complex is called *Chromatin*. As mentioned before, long sequences of DNA with specific information are called genes, so every chromosome carries a determined number of genes. Most plant and animal cells are diploids which means that chromosomes are present in pairs within their corresponding nuclei. Two chromosomes with the same array of genes are called homologous.

In human somatic cells, there are 22 sets of homologous chromosomes and one pair of the so called *sex chromosomes*. In females, sex chromosomes are homologous unlike in the male case in which differential regions among them do exist. Thus, human somatic cells have 46 chromosomes. Within the nucleus, there are several proteins that link the genetic material to form increasing organizational levels up to the chromosome structure (see Fig. 2). These proteins are divided into *histones* and *non-histones*. The *double helix* is made up by nucleotides arranged in two long a central axis. This double helix coils around core histones to form the *Nucleosome*. Core histones refer to a protein kernel of eight molecules: H2A, H2B, H3 and H4 (two of each) making a complex known as *octamer of histones*. Six nucleosomes linked by DNA fragments and other histones (*linker histones*) colis helicoidally to conform the so called 30-nm chromatin fiber. This fiber finally builds up chromosomes.

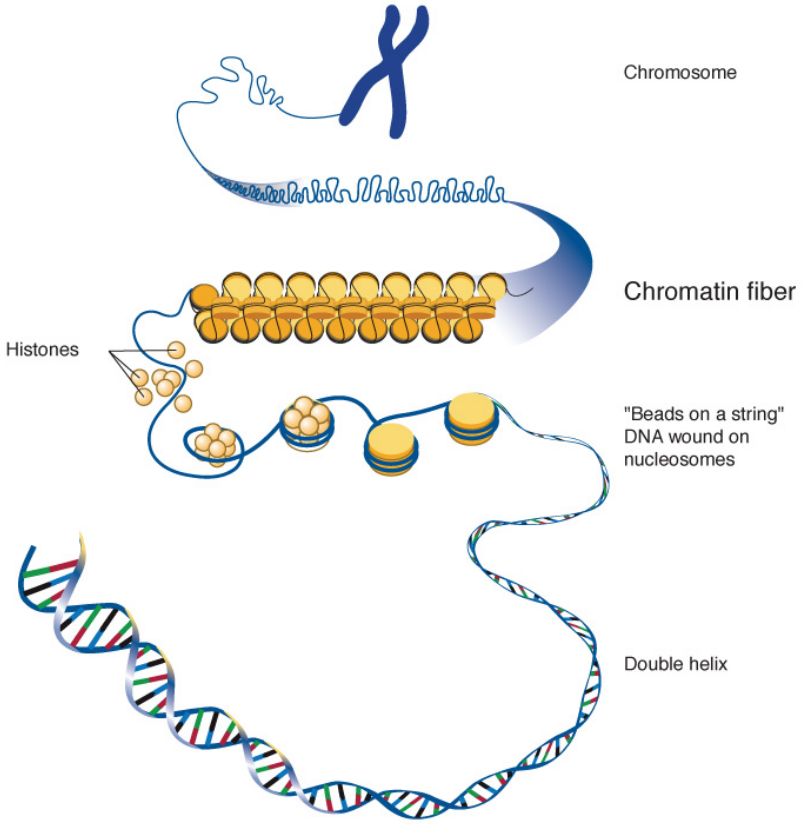


Figure 2 – Scheme of the organizational levels of the genetic material within the cell nucleus. The DNA packs or condense in the different structures by the action of the histones. source: https://www.genome.gov/glossary/index.cfm?id=32. All files from *Talking Glossary of Genetic Terms* are public domain.

2.1 Types of DNA and its biological function

There are about 6×10^9 base pairs in the nucleus of human cells. Nevertheless, not all the DNA molecules have the same chemical structure, despite their same chemical composition. The are three main forms of DNA in human beings: A-, B-, and Z-DNA [15]. B-DNA is the canonical and the most predominant form of DNA. It was firstly described in the famous work of Watson and Crick [14]. They discovered that guanine pairs with cytosine and adenine with thymine in such a way that the geometrical shapes of the resultant molecules are essentially equal.

This fact solves the question about how the helical structure of the genetic material, inferred by x-ray diffraction, can accommodate arbitrary sequences of bases despite the different sizes and shapes of these purines and pyrimidines. This scheme was supported by previous works about the composition of DNA, mainly those of Visher, E. Chargaff [16,17], Maurice Wilkins, Rosalind Franklin, and R. G, Gosling. The two chains of the helix run in opposite directions, this feature is what makes possible the successful replication of the DNA. As mentioned before, the nucleotides are composed by a base pair and two sugar-phosphate groups. These groups form two sequences known as the DNA backbone,

between which the two nitrogenous bases are disposed. Watson and Crick also determined that the axial distance between two adjacent nucleotides is about 3.4 Å and the helical structure repeats every 34 Å. Thus, there are 10 bp per turn so the twist angle is 36 °. The helix has 24 Å of diameter and its axis is nearly perpendicular to the plane defined by the bases.

When dehydrated DNA is put under x ray diffraction it shows a different structure than the B-form described before. This is called A-DNA. In both A- and B-DNA the twisting of the helix is right handed and the bases follow the Watson-Crick pattern. Nevertheless, in the A configuration the bp-plane is not perpendicular to the helix axis anymore. Such tilting (19 °) results in a wider and shorter helix when compared with the B-DNA case. Many of the differences between A- and B-DNA are due to the puckering of their sugar-phosphate groups. A-DNA seems to be related with DNA-drug as well as DNA-protein interactions. In addition, it has been looked up to its role in genome structure and function [18]. It is important to point out that A-DNA not only appears as dehydrated DNA, double stranded RNA and some RNA-DNA hybrids also show an A-like-structure.

The works of Alexander Rich and coworkers showed that there exists a third form of DNA [19, 20]. They studied the hexamer $d(CG)_3$ and concluded that it forms a double helical structure but, unlike the other configurations, the helix was left-handed. Furthermore, the sugar-phosphate backbone showed a zigzag arrangement, hence the name Z-DNA. This new form of DNA appears in certain physiological cellular processes and afterwards, decays to the B form [21]. Its biological function is still unknown but it have been suggested to play a role in the pathology of poxviruses [22]. Table 2 shows a comparison of the main features of the three DNA configurations presented and their corresponding helices.

Table 2 – Comparison of the main features of the A-, B- and Z-DNA conformations.

	DNA conformation		
	A	B	Z
Shape	Broadest	Intermediate	Narrowest
Rise per Base Pairs	2.3Å	3.4Å	3.8Å
Helix Diameter	25.5Å	23.7Å	18.4Å
Screw sense	Right-handed	Right-handed	Left-handed
Glycosidic bond	*anti*	*anti*	alternating *anti* and *syn*
Base pairs per helix turn	10.7	12	10
Pitch per Helix turn	25.3Å	35.4Å	45.6Å
Base pair tilt (Helix axis)	19°	1°	9°
Major Groove	Narrow and very deep	Wide and quite width	Flat
Minor Groove	Very broad and swallow	Narrow and quite deep	Very narrow and deep

2.2 DNA lesions and repair mechanisms

The breakage of the sugar-phosphate linkage of a strand is the simplest damage that direct impact of radiation can induce on the genetic material. This is termed as Single Strand Break (SSB). Usually, SSBs are easily repaired by enzymatic mechanisms that restore the initial DNA structure using the opposite strand as a template. It is commonly assumed that a Double Strand Break (DSB) is induced when two SSBs are produced in opposite strands (not necessarily in the same nucleotide but close enough to interact) . DSBs are lesions of major importance since they can lead to mutations and even to cell death. It has been observed that, for low dose irradiations, radiosensitivity and cell survival correlate with the number of initially-produced DSBs. For higher doses, unrepair or mis-repair of DSBs are related to cell survival. The two principal repair mechanisms for DSB are Non-Homologous End Joining (NHEJ) and Homologous Recombination (HR). The former acts on the blunt ended DNA fragments created after sugar-phosphate linkage breaking and the latter uses the sequence homology with an undamaged copy of the injured region so it has to wait until the late synthesis phase of the cell. The effect of lesions due to clustered DSBs have been studied as well [23–25]. Other types of lesions include base damage, protein-DNA cross-links and protein-protein cross-links among histones and non-histones proteins [26]. Base Excision Repair (BER), Nucleotide Excision Repair (NER), and Mismatch Repair (MR) are other repair mechanisms. They respond to base oxidation, alkylation and strand intercalation.

2.3 Chromosome Aberrations

There is consensus in the scientific community that DNA lesions may lead to chromosome aberrations (CA) [27]. Repair mechanisms, such as the classic NHEJ or its alternatives, are error-prone and become the main pathway to end up with chromosome or chromatid aberrations when there are DNA lesions [28]. Misrepaired DSBs appear at mitosis as terminal deletions or incomplete exchanges, leading to loss of genetic information. Such deletions and/or losses may induce cell death. The formation of an aberration will mainly depend on the repair pathway followed by the cell, the availability of certain proteins and the kinetics of the aberration formation. HR is fairly error-free but it appears just in the S and G2 phase whereas NHEJ is the dominant mechanism along the whole cell cycle.

Nowadays, it is understood that the NHEJ represents more than one molecular repair pathway. In Ref. [29] was showed that cells with deficit of the Ku heterodimer, DNA PKcs, and LigIV, proved to be efficient in DNA joining in the context of class switch recombination even though such proteins are essential for canonical NHEJ. This other path is named B-NHEJ and the canonical NHEJ is called C-NHEJ. B-NHEJ is

error-prone and it has been speculated that most of, if not all, the CA originate from misrepair in this pathway. Nevertheless, B-NHEJ does not account for the fast kinetics of the aberration formation and is not clear why a cell able to carry out C-NHEJ would switch to B-NHEJ.

Another mechanism of damage repair that has been considered is the microhomology-mediated-end-joining (MMEJ) which could be promoted by DNA joining mediated by CtIP (carboxy-terminal interaction protein) [30]. MMEJ could account for a portion of the slow kinetics aberrations suggesting that C-NHEJ is the responsible for the fast kinetic parcel of chromosome aberrations.

In all these CA formation paths, it is implied that basic notions of DSB-DSB interaction should be taken into account. For instance, how long two DSB ends would remain open and what is the likelihood for them to lose track of each other and not to rejoin.

3 Cell Death

There are two well defined periods in the cell proliferation cycle, namely, *mitosis* (M) and *DNA synthesis* (S). The gap between M and S is called the G1 phase and the gap between S and M, the G2 phase. Mitosis is the period where cells divide and, along with the G2 phase, is the most radiosensitive cycle stage. Cells are regarded as 'killed' when they lose their reproductive integrity due to radiation damage [10]. This definition of cell death has nothing to do with whether the cell physically disappears or not. There are several mechanisms by which loss of reproductive integrity may occur: Apoptosis, necrosis, mitotic catastrophe or induced senescence. In apoptosis, the cell metabolism is shut down, the permeability of the membrane increases and there is an inflammatory response among other characteristics. Apoptosis is also known as programmed cell death. In necrosis, for instance, plasma membranes appear discontinuous and DNA degrades. Senescence is the stopping of cell division, accompanied by shortening of the telomeres and poor cell-cell contact. At higher doses, mitotic catastrophe is more likely to occur. A cell may undergo mitotic catastrophe in the first few divisions, it is, basically, an unsuccessful division that can lead the cell's progeny to apoptosis.

The cell cycle may take from 10h to 40h with about 30% of the time spent in the G1 phase, 50% in the S period, G2 taking around 15% and M a 5% of the time. Radiosensitivity varies throughout the cell cycle, being the S period the one with the greatest radioresistance as long as HR and NHEJ primarily occur at this stage. In the G2 and M phases, cell are more susceptible to radiation damage because chromatin compaction and limited repair ability due to reduced enzyme access. G1 has an intermediate position regarding radiosensitivity.

3.1 On Cell Survival Curves, the LQ-model and the DRAT

Cells with its reproductive integrity intact are called *clonogenic*. A cell survival curve describes the relationship between the fraction of surviving cells (i.e those which are still clonogenic after irradiation) and the absorbed dose. Typically, this relationship is exponential. Thus, it is common to represent it in a log-linear plot. For densely ionizing radiations, this plot is almost a straight line while for sparsely ionizing radiation it starts linear but, after a certain dose, shows a shoulder and finally falls as a straight line for higher doses. Fig. 3 shows some examples of cell survival curves for both densely and sparsely ionizing radiation.

The *Linear-Quadratic Model* (LQ-model) has been the most successful procedure for fitting the cell survival fraction as a function of dose, at least for doses between about

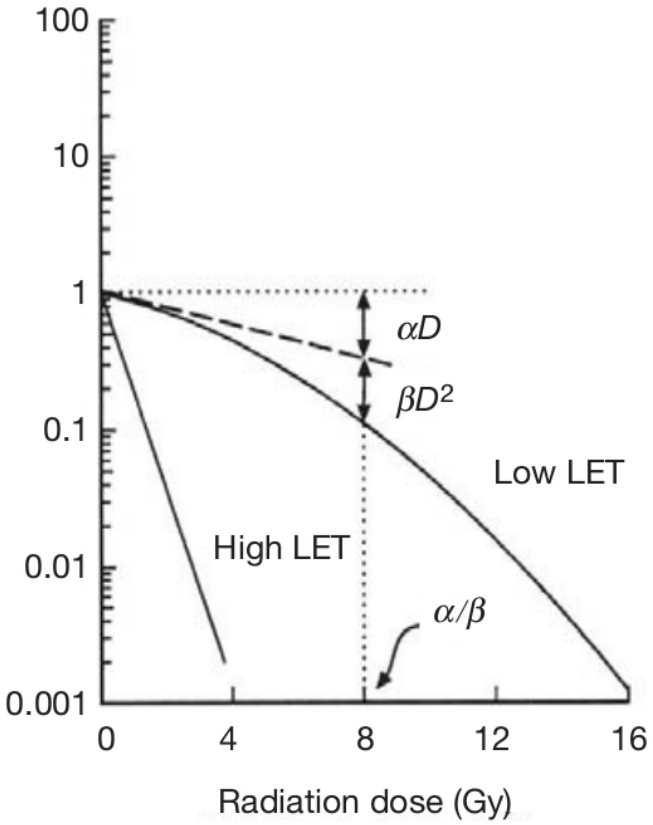


Figure 3 – Illustrative cell survival curves for both densely (High LET) and sparsely (Low LET) ionizing radiation. The α parameters measures the initial slope of the curve. Large values of α result in steeper curves. β defines the curvature of the curve, the smaller the β the lesser the curvature.

2 and 15 Gy [31, 32]. The experimental data is fitted with the function

$$S = e^{-(\alpha D + \beta D^2)} \tag{3.1}$$

where S is the survival fraction at dose D. The parameters α and β are then found using a standard regression method. The importance of the LQ-model relies on its ability to reproduce experimental data. It is used to determined the iso-effect dose when alternating fractionation regimes in cancer therapy. In addition, Tumour Control Probability (TCP) and Normal Tissue Complication Probability (NTCP) may be predicted by means of the LQ-model. Recently, it has been compared the LQ approach with the TCP in the study of the radiosensitivity of mixed cell colonies [33], for instance.

Historically, the LQ-model has appeared just functional, i.e. its parameters are introduced *ad hoc* in the practice. Nevertheless, both the αD and βD^2 terms admit an interpretation as first and second order mechanisms for damage induction, respectively. This interpretation comes from the mechanistic grounds attributed to the LQ-model based, mainly, on the Dual Radiation Action Theory (DRAT) developed by Kellerer and Rosi [34, 35].

Before explaining the DRAT, it is crucial to mention that as early as 1942 Lea and Catchside [36] proposed, based on their experiments, the formula $\alpha D + \beta D^2$ to model the number of chromatid and chromosome interchanges induced in cells irradiated with X-rays. They noticed that the number of such aberrations increases with dose more rapidly than the first power as well as with the dose rate. This leads to believe that there would be a non-linear mechanism behind the formation of chromosome aberrations. They went a step further, proposing a multi-parameter biophysical model to predict the number of chromosome interchanges per unit dose. In their model, the probability for the interchange of two DSBs separated a distance x is given by $e^{-x/h}$, where h is a characteristic distance of the order of microns. This was the first attempt to construct a target theory in radiobiology and many other approaches followed it in the next years. For instance, the "multi-target single hit" or the "single-target multi-hit" approaches. There is a short but rich discussion on target theory in Ref. [37].

DRAT

Kellerer and Rosi started from the Lea's target theory to develop their Dual Action Radiation Theory. The basic postulates of the theory are (1) ionizing radiation produces sub-lesions in the cell that are proportional to the dose and (2) the combination of two sub-lesions may result in a lesion that has a probability p to lead to cell death. As in the target theory, there is a characteristic distance within which the interaction between the sub-lesions is likely to occur. Three models for the probability $g(x)$ of sub-lesion interaction were proposed by Kellerer et al. in [35]

$$g(x) = \begin{cases} 1, & if\ x \leq h \\ 0, & if\ x > h \end{cases} \tag{3.2}$$

$$g(x) = e^{-(x/h)} \tag{3.3}$$

$$g(x) = e^{-(x/h)^2} \tag{3.4}$$

They also showed that putting $h = 0.4\mu m$, $h = 0.2\mu m$ and $h = 0.1\mu m$ respectively, similar results are obtained. In the DRAT, the survival fraction is of the form

$$S = e^{-k(\zeta D + D^2)} \tag{3.5}$$

where k is a proportionality constant that reflects a biological property of the system since it is related to the average number of lesions induced and to the specific energy z. ζ is defined as

$$\zeta = \frac{\int z^2 f(z) dz}{\int z f(z) dz} \tag{3.6}$$

In Eq. 3.6 $f(z)$ is the distribution of single-event specific energy (the sum of all energy deposits per unit mass). With this definition, ζ plays a role analog to the celebrated α/β ratio of the LQ-model.[1]

[1] The α/β ratio is the dose at which both terms in Eq. 3.1 contribute equally to the cell death.

In short, the DRAT states that the number of lethal lesions N_{LL} (those leading to cell death) shows a linear-quadratic dependence on the dose. The LQ-model takes this result and combines it with the assumption that one lethal lesion is enough to induce loss of reproductive integrity. Also, the likelihood of cell inactivation is assumed Poisson distributed so the surviving fraction is, under these conditions, $e^{-N_{LL}}$.

To better understand what do mean the terms *first* and *second* order mechanism mentioned before, the solution of the following system of differential equations that describes the DRAT [38], may be of help (details in Appendix B)

$$\dot{N}_{RL} = 2p\dot{D} - \mu N_{RL}(t) \tag{3.7}$$

$$\dot{N}_{LL} = \alpha\dot{D} + p\epsilon\dot{D}N_{RL}(t) \tag{3.8}$$

where N_{RL} is the mean number of repairable lesions, $\dot{D}$ is the dose rate (constant), p is the probability per unit absorbed dose to induce a lesion, μ is the rate of damage repair, α is regarded as the probability per unit absorbed dose to directly induce a lethal lesion and ϵ is the probability of two lesion interact. It is accepted that the α coefficient is related to the interaction of two DSBs produced by the same track so it is dose-rate independent.

The solution of the Eq. 3.7 describes how many lesions are available for interaction after a time t and is obtained subtracting the number of lesions repaired from the number of pairs of lesions induced[2]. The integral of the Eq. 3.8 represents the number of lethal lesions created plus the number repairable lesions that interacted becoming lethal as well. Every term in the integrals of Eqs. 3.7 and 3.8 has a probability to occur which is determined by the parameters described above. The solution for N_{LL} is

$$N_{LL} = \alpha D + \frac{2\beta D^2}{\mu^2 t^2}\left(\mu t - 1 + e^{-\mu t}\right) \tag{3.9}$$

where $\beta = p^2\epsilon$. For acute irradiation, as that carried out in high dose-rate radiotherapy treatments, the irradiation time is much lower than the repair time scale, that is $\mu t << 1$. Thus, the expansion to second order of the exponential in Eq. 3.9 leads to

$$N_{LL} = \alpha D + \beta D^2 \tag{3.10}$$

Now the α and β in this equation have a clear meaning. α is the probability per unit absorbed dose to directly induce a lethal lesion. This means that it is the effect of a single track, the action of one primary particle its associated secondary particles that are able to cause lethal damage. Thus, is a first order effect of the ionizing radiation. β was defined as $p^2\epsilon$, that is, the probability to have 2 sub-lesions (p^2) times the probability that they interact (ϵ). This can be interpreted as an inter-track mechanism of production

[2] The factor 2 could be cast to the second equation and even it could be absorbed into ϵ, but since interaction of pairs of lesions is being looking for, keep it explicitly is cleaner for the reasoning.

of damage since in principle, the two sub-lesion may not appear at the same time or have the same origin. To arrive at the LQ-model, it is assumed that this N_{LL} determines the probability of cell death via a Poisson distribution. So, the determination of the N_{LL} using the DRAT does not depend on the Poisson distribution assumption.

4 Monte Carlo Method

What is?

The *Monte Carlo* (MC) method is a numerical solution to an object-object interaction problem [39]. It is based on the sampling of random numbers in conjunction with probability density functions. From the physic's point of view, the MC method arises as a powerful tool for modelling the behavior of macroscopic (possibly complex) systems through simulating the basic dynamics of their microscopic components. In this sense, is an attempt to imitate the nature with its stochasticity. The idea is to perform repetitive calculations until they converge in order to reach the solution of the problem in question. Due to this repetitiveness, it is mandatory the use of computers to carry out the simulations to solve modern problems.

The fundamental interactions of a system are not always well known. In these cases, the researcher has to make assumptions based on the phenomenology and/or intuition. Here is where MC, experiments and theory work together for a better understanding of the phenomena. Sometimes, theory and assumptions provide small corrections to experiments. Other times, experiments and simulations (MC) guide the theoretical approaches to a problem. Finally, experiments, theory and MC altogether may lead to new insights and comprehension of nature.

In other cases, as electromagnetic interactions for instance, the fundamental interactions are well described and the mathematical features are sufficiently precise to allow detailed and highly accurate simulations. Complex trajectories depiction, track of histories and computation of local quantities are, among others, the most important applications of the MC method.

The fast growth of computer calculation speeds has played a crucial role in the introduction of MC codes to study more and more problems. Still, deterministic/analytic calculations are of the main importance since they provide deep insight and allow one to develop intuition about the behavior of the fields and particles studied. Nevertheless, there is a fair advantage of MC over deterministic calculations, namely, MC relies on minimal data storage and maximum floating-point operations while analytic methods usually demands maximum data with minimal floating-point operations. Modern computer architectures favor MC simulations because emphasize in interactions reducing storage. This avoids bottleneck problems of communication between CPU, cache and disk.

4.1 Basic Principles

4.1.1 RNG

In order to reproduce the stochasticity of nature, the Random Number Generator (RNG) is the closest thing one can think at since it can be associated to stochastic variables through Probability Density Functions (PDF) and Sampling Methods (SM). The way to do this is, first of all, to generate a (pseudo-)random number via a well-tested method (e.g. Linear Congruential RNGs, Fibonacci generators) [40]. The next step is to use the generated random number along with a sampling method (will be presented below) to choose the value of the stochastic variable in question. In this step is when the PDF or its inverse comes into scene.

Linear Congruential Random Number Generator - LCRNG
LCRNGs are described by the equation

$$u_{n+1} = (a * u_n + c) \mod m \tag{4.1}$$

Where the parameter a is known as the multiplier, m is usually of the order of the largest integer that can be represented in the machine and c is a constant that allows the generation of an exact zero. This c may be taken as zero, with the sequence obtained this way as good as in the non-zero c case. A good choice of a and m may depend on c. It is clear that dividing by m, the set $\{u_n\}$ is mapped into the interval [0,1].

This method has a basic weakness pointed out by Marsaglia [41] and it is that d-tuples of such numbers, show lattice patterns when considered as points in the d-space. That is, they lie in planes (Marsaglia's planes). Nevertheless, the trick followed by most RNG developers is to maximize the number of such planes in hyperspace.

The LCRNG is a well-studied method to produce not-so-long sequences of random numbers. The method can be improved to obtain long sequences using the 'subtract-with-borrow' approach, which is, essentially, the conjunction of two independent LCRNGs to produce the random number. Another approach may be the *Shuffling* in which one RNG is used to produce the random numbers and a second one is used to choose their order.

4.1.2 Sampling Methods

Given a PDF it is possible to compute its Cumulative Probability Function (CDF) by integration. Suppose a PDF $f(x)$ is defined over the interval [a,b] (not necessarily finite), then, its CDF is,

$$c(x) = \int_a^x f(x)\, dx \tag{4.2}$$

with $c(b) = 1$ because $f(x)$ is properly normalized.

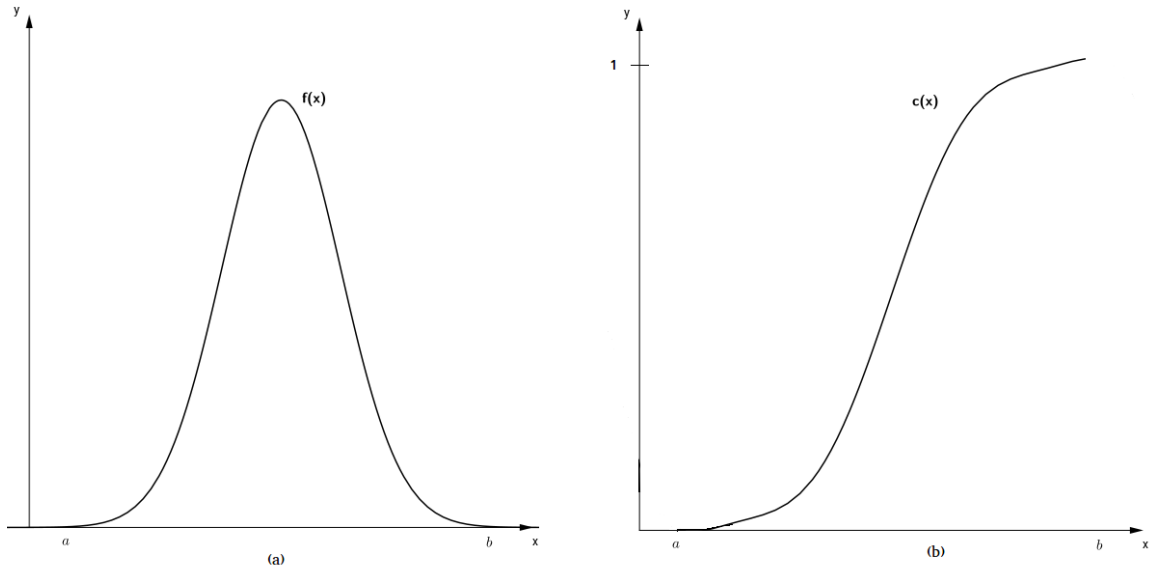


Figure 4 – (a) Typical Probability Density Function. (b) Cumulative Probability Function obtained by integration of the PDF represented in (a).

Direct Sampling Method

From the definition of the CPF or looking at the figure 4 it can be noted that $c(x)$ takes its values in the range [0,1]. Thus, it can be mapped into the normalized set $\{u_n\}$. This means that for every $u_n \in [0,1]$ there exist an $x \in [a,b]$ such that

$$u_n = c(x) \tag{4.3}$$

And, since $c(x)$ arises from a well-behaved (integrable and normalized) PDF, this equation is invertible numerically and/or analytically, so

$$x = c^{-1}(u_n) \tag{4.4}$$

Then, using eq. 4.4 the value of the stochastic variable x can be determined from the value of u_n. This variable will be distributed according to $f(x)$ provided the $\{u_n\}$ are produced uniformly over [0,1]. Consider, as an example, the following PDF and its CPF

$$\begin{aligned} f(x) &= \mu e^{-\mu x} \\ c(x) &= \mu \int_0^x e^{-\mu x} dx = 1 - e^{-\mu x} \end{aligned} \tag{4.5}$$

Let r be a uniformly generated random number in [0,1], then, inverting $c(x)$ in eq. 4.5 and noting that $1 - r$ and r are equally distributed, leads to

$$x = -\frac{1}{\mu} \ln(r) \tag{4.6}$$

Physically speaking, eq. 4.5 represents the probability density function for the distance traveled by a particle to the next interaction in a medium with linear attenuation coefficient μ. Eq. 4.6 is the way to sample x in a MC simulation.

Rejection sampling Method

For some CPFs it could be very difficult to compute $c^{-1}(x)$. The rejection method is another approach in which $f(x)$ is used instead of $c(x)$.

The first step is to determine an upper bound for $f(x)$. Of course $f_{max}(x_{max})$[1] is the best choice but overestimation would work (less efficiently) as well. Then, a new function is defined as $p(x) = f(x)/f_{max}(x_{max})$. This $p(x) \in [0, 1]$ when $f_{max}(x_{max})$ is used, or $p(x) \in [0, 1)$ if the upper bound is overestimated.

Next, with a random number u_1 (uniformly distributed over [0,1]), a value of x is picked up in the interval [a,b] doing $x = a + (b - a)u_1$. If some or both endpoints are infinite, it is possible to map the interval onto the interval [0,1] by means of a logarithmic function for instance.

Finally, a second random number $u_2 \in [0, 1]$ is uniformly generated and compared with $p(x)$. If $u_2 < p(x)$ then x is accepted, else, go back to step 2 without taking the x value.

The rejection method is schematized in figure 5. Provided an x (after u_1), it is rejected if $(x, p(x))$ lies in the shaded region. The choice of this technique over the direct method depends on whether wasting random numbers compensates saving computation time in complicated inversion process ($c^{-1}(x)$) or not.

Mixed Method

It was said that $c(x)$ could be so difficult to invert that rejection method arises as a better option to pursue. Nevertheless, if the rejection region is considerably larger than the region of acceptance, efficiency starts to be an important issue to care about. For such cases, a combination of both methods can be applied as follows:

1. Split $f(x)$ into two functions, that is, write

$$f(x) = p(x)q(x) \tag{4.7}$$

   The idea is to factor the 'spiky' part of $f(x)$ in, say, $p(x)$ and the mathematically complex in $q(x)$.

2. Normalize $p(x)$ to get $\tilde{p}(x)$ such that $\int_a^b \tilde{p}(x)dx = 1$

3. Normalize $q(x)$ to get $\tilde{q}(x)$ such that $\tilde{q}(x) \leq 1 \in [a, b]$

[1] x_{max} is not necessarily the endpoint of the interval, it is simply the value of x where $f(x)$ takes its maximum value.

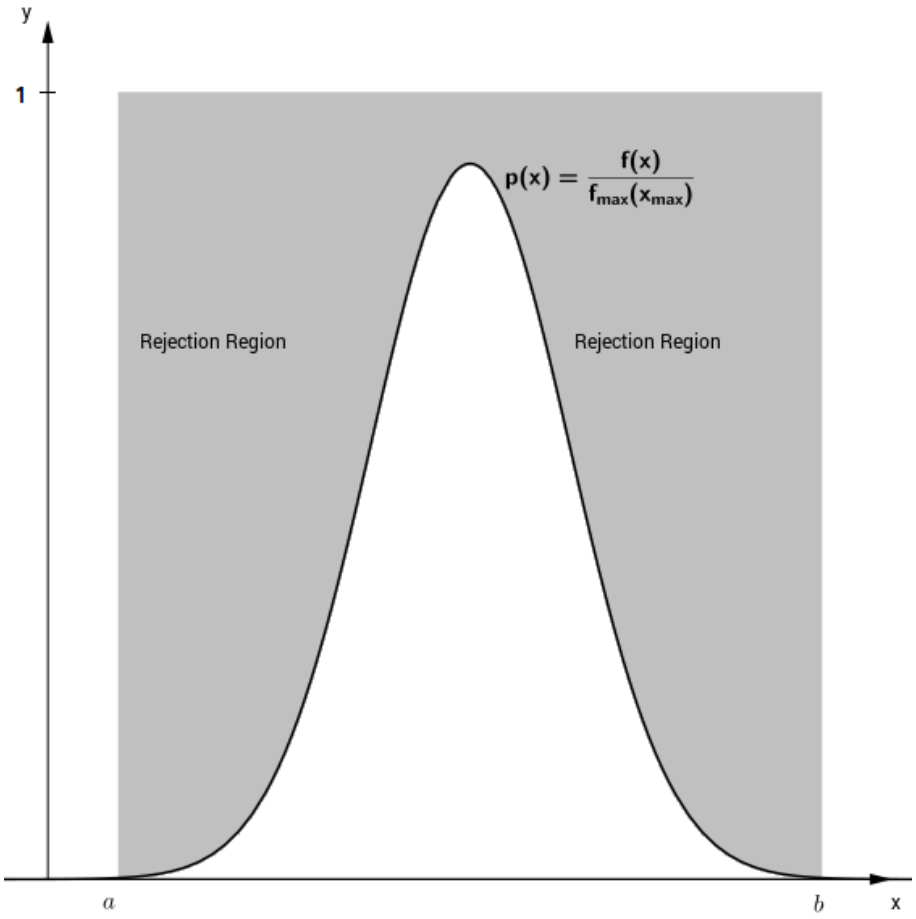


Figure 5 – The scaled PDF $p(x)$. Illustration of the Rejection Method with overestimation. All points in the shaded region are rejected.

4. Apply the direct method to $\tilde{p}(x)$ and pick a value of x.

5. Sample a random number r, uniformly distributed over [0,1], and compare it with $\tilde{q}(x)$ (the x found in the previous step). If $\tilde{q}(x) \leq r$, accept x, else, go back to step 4.

As an example, consider the following PDF with $x \in [0, \infty)$

$$f(x) = N \frac{2x}{(1+x^2)^2} e^{-x^2} \tag{4.8}$$

Where N is the normalization constant that makes $\int_0^\infty f(x)dx = 1$. Good choices of $p(x)$ and $q(x)$ are

$$\begin{aligned} p(x) &= \frac{2x}{(1+x^2)^2} \\ q(x) &= e^{-x^2} \end{aligned} \tag{4.9}$$

Now, the CDF for $p(x)$ is calculated to apply the direct method to this part of $f(x)$. Note that both $p(x)$ and $q(x)$ are normalized to 1.

$$u_1 = c(x) = 1 - \frac{1}{1+x^2} \tag{4.10}$$

so

$$x = \sqrt{\frac{u_1}{1-u_1}} \tag{4.11}$$

With this x, now it is possible to do the rejection test to

$$q(x) = q(u_1) = e^{\frac{u_1}{1-u_1}} \tag{4.12}$$

Now a second random number u_2 is sampled and, in order to accept the x value given by eq. 4.11, the following expression must be satisfied

$$e^{\frac{u_1}{1-u_1}} \leq u_2 \tag{4.13}$$

Multidimensional PDF

The methods just described above apply equally to PDFs depending on two or more variables. The only new feature is that marginal probabilities must be computed before actually doing the sampling. Marginal probabilities are obtained by integration over all variables except the one of interest. Once the marginal probability is determined, any of the SMs is used to choose the value of the variable. This value is put in the conditional probability defined by both the PDF and the marginal probability, and the same process is performed for the other variable(s).

4.2 Monte Carlo in Particle Transport

It was shown how stochastic variables can be sampled from their probability density functions by means of random numbers and a sampling method. The passage of a particle through a material medium is governed by the interactions that this entity may undergo with the components of the medium. What interaction will occur? After what distance the next interaction occurs? What energies and scattering angles are involved?. All of these are stochastic variables of the problem. The determination of such quantities may be done using the Monte Carlo technique, this is, the variables can be sampled from an appropriate PDF in every step of the transport. Those PDFs are constructed out of the interaction cross sections (CS) of the many processes involved.

Once the physics and the CS of an interacting system are known, the geometrical configuration is another feature to be defined properly in order to perform a MC simulation. Having clear this, the idea of a detailed simulation of a particle transport is, in first place, set the initial parameters (Incident energy, angle, processes to be considered, geometries and so on). Next, sample the distance to the next interaction, choose a physical process, pick up the scattering angle of the particle and check the conditions to continue the transport. After checking, store all the new relevant data and update the position and energy of the particle if necessary. This is the actual transportation.

The steps described in the above paragraph are not the whole story but an scheme of how MC transport works. There are many subtleties regarding the very nature of the particle and the way the transport has to be made. For instance, in the transport of charged particles, there are *classes* of transport schemes [42] in which histories (particle trajectories) are condensed according to a given approach. In the *Class I* scheme, for example, histories are condensed by fixing the path-lengths of the particles or even the energy losses (*Class I'*). *Class II* schemes are based upon small energy losses and deflections but allows independent sampling of catastrophic events (larger energy losses and deflections).

In computational radiobiology is very important to be able to transport radiation particles through biological media. This is why Monte Carlo became fundamental in the field. The MC approach provides a detailed insight into the events that occur when radiation interacts with matter, allowing the understanding of some of the effects of such particles in biological systems. Nowadays, there exist many MC codes to perform a simulation. Some of them are regarded as general purpose codes, i.e. they serve to transport several particles in virtually any medium in order to determine macroscopic dose distributions, among other applications. For example the codes PENELOPE [43], EGSnrc [44], MCNP [45], and GEANT4 [46]. On the other hand, there are specific purpose codes such as GEANT4-DNA [47] and PARTRAC [48]. These two latter codes are used in nanodosimetry and computational radiobiology. In the next part of the document, it will explained how the *GEANT4-DNA* code was used to simulate the passage of protons and alpha particles across cellular nuclei and the insight it provided into the early stage of radiation-induced DNA damage.

5 Work plan

The aim of this work is to develop geometrical models of the A- and Z-DNA configurations with atomistic resolution and combine them with the already implemented B-DNA model. These models will be used to construct a simplified geometrical model of the genetic material of a human cell nucleus. High LET ionizing particle tracks simulated by the Monte Carlo method will be superimposed over this model to study the early effects induced by these particles on DNA and how DNA conformation influences damage yields.

The work is divided into 4 stages, the first one being the design of the models. Once the geometrical models are developed, the radiation beams must be set up by defining the type and energy of the incident particles. Next, the *GEANT4-DNA* code is used to perform the simulation. The phase space file created in each simulation contains the energy deposition distribution and labels of the beam particles, secondary particles and physical processes involved in the transport. The third phase of the work plan consists in reading those phase space files and superimposing them onto the geometrical models to produce several new files that store the relevant information (strand break indexes and strand break maps) for the post-processing phase. At the post-processing stage, the strand breaks on the genetic material are counted and classified. In **??** (Results) it will be shown what kind of information can be extracted from the study of the strand break distribution.

5.1 DNA models

In chapter 2 was explained that most of the genetic material is inside the cell nucleus. The DNA is packaged into different structures through the action of histones. Starting from the very basic unit of genetic material, the *base pair (bp)*, those structures will be built up by means of linear transformations that preserve the geometrical/spatial restrictions of each DNA configuration (See Table 2).

Over the last 30 years, several models of the human genetic material have been developed in order to study, from the computational point of view, the effects of ionizing radiations on the DNA. The pioneer works of Charlton et al. [49] followed by the first atomistic model of Pomplum [2] to study the damage due to ^{125}I Auger electrons, have evolved to more detailed representations as, for instance, the work of Friedland et al. [50] in which six DNA organizational levels were taken into account. All these works, and other representatives [51–54], were based on the B-DNA conformation since this is the standard form of DNA found in our cell nuclei.

In the recent years, the biological roles of the A- and Z-DNA have began to be studied. Part of this work aims to elaborate geometrical models of the three aforementioned DNA conformations accounting for five organizational levels of the genetic material, namely, the base pair, the double helix, the nucleosome, the 10-nm and the 30-nm chromatin fibers. To this end, the position of each atom in the nucleotide pair will be used as the brick to construct the higher levels.

The specific positions of the atoms in the CG nucleotides were taken from the GLACTONE project web page [55]. The .pdb files downloaded contain a complete turn of the double helix for the A-, B-, and Z-DNA configurations. The positions of the base pair's atoms and its corresponding sugar-phosphate molecules were extracted from these files. In the case of Z-DNA, a GCGC sequence [1] is read since in this configuration two bps would represent a building block of the double helix.

The coordinates are centered along the z-axis (helix axis) and one Van der Waals radius is assigned to each atom according to [2] (See Table 6 in section 6.1). It is important to note that base pairs are not necessarily centered in xy-plane and this issue should be considered when determining the helix diameter in each configuration. The axial step is the distance between two contiguous bps and it is different from the bp height. This avoids excessive overlapping of the molecules. Each DNA configuration demands its own geometry for the double helix irrespective of the nucleotide used. For instance, an A-DNA double helix has 10.7 bps per turn while the B-DNA- and Z-DNA-type has 12 and 10, respectively. Also, helix diameters in each case vary.

5.1.1 The DNA helix

The number of base pairs per helix turn determines the rotation angle between two consecutive rungs. The axial step is the shift along the z(helix)-axis. Thus, to generate a right helix, the building block has to be copied, rotated and translated properly. That is, with the rotation wise, angle, and shift of the corresponding conformation. Figure 6 shows the helices obtained in the way explained, using the GLACTONE's coordinates. A base pair is a union of spheres, every one representing one atom. The atom's coordinates are referred to their centers and the radii are taken as the corresponding Van der Waals radius. The helix is the union of the base pairs created by rotation and shifting. With the helices ready, it is now possible to extract parameters of the model, such as the helix diameter and rise per bp. Also, the screw sense and orientation of the bps can be checked. The parameters extracted are shown in Table 3.

[1] GCGC means a G-C base pair right above a C-G one. Thus, the chemical bases are intercalated along the strand.

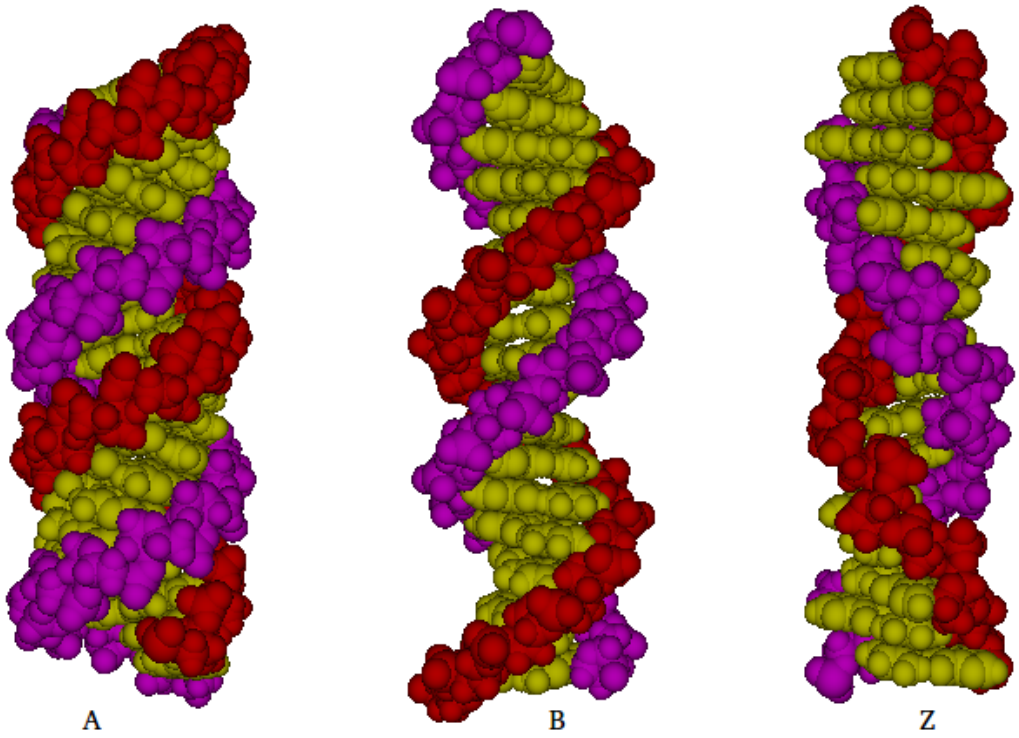


Figure 6 – DNA helices constructed out of the atom's coordinates extracted from the GLACTONE web page. A, B and Z conformations. Each atom is represented by a sphere with radius equal to one Van der Waals radius. C-G atoms are colored in yellow. Violet and red are used for the sugar-phosphate groups of each strand.

Table 3 – Helix diameters and helix rise per turn as determined from the model.

	DNA conformation		
	A	B	Z
Rise per Base Pairs	2.3Å	3.3Å	3.7Å
Helix Diameter	25.5Å	23.7Å	20.9Å
Screw sense	Right-handed	Right-handed	Left-handed

5.1.2 The Chromatin Fiber

The nucleosome is formed by a long strand of DNA coiled around an octamer of histones, forming a two-turns helix with pitch D_{hel}. This is accomplished nn the model by bending the helix around a sphere that would represent the core histones. The sphere diameter and the number of bps per nucleosome depend on the diameter and axial step of the helix and are restricted to the biological demand that the chromatin fiber should have nearly 30 nm of diameter. The chromatin fiber is a helix composed by six of these nucleosomes per turn, bounded by DNA linker fragments. These fragments are modeled as a sigmoid curve (perpendicularly to the nucleosome plane) strand of DNA that binds the top of a nucleosome with the first bp of the next one.

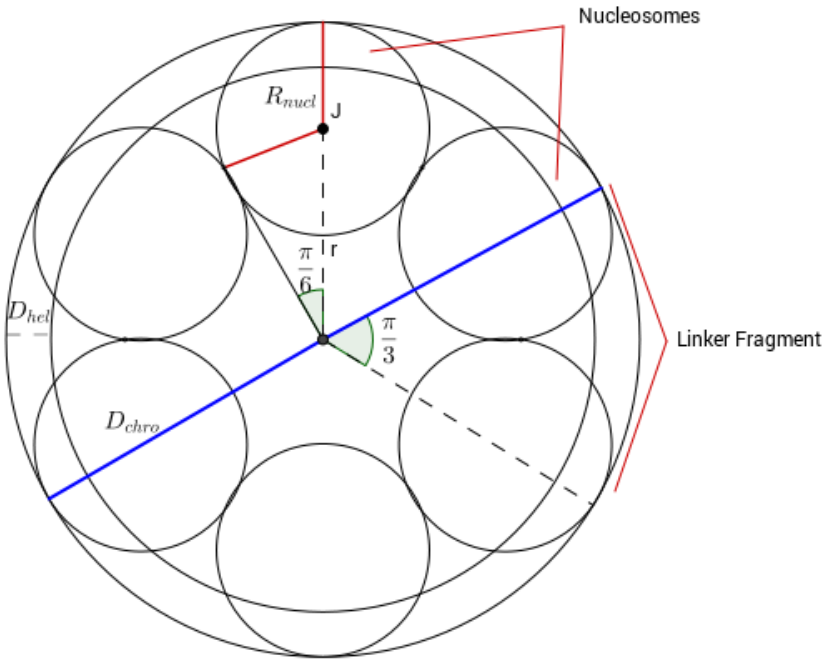


Figure 7 – Simplification of the Chromatin. Upper view of the 30-nm chromatin fiber with its main geometrical parameters.

B-DNA

To construct these structures, the optimization method developed by Bernal *et al.* [56] for the B-DNA was followed. The geometry and quantities related to this problem are shown in Fig. 7. The linear equation system to solve is,

$$\begin{aligned} r \sin \frac{\pi}{6} &= R_{nucl} \\ 2R_{nucl} + 2r &= D_{chro} \end{aligned} \tag{5.1}$$

Eq. 5.1 gives $R_{nucl} = D_{chro}/6$ which would lead to $R_{nucl} = 5\ nm$ for $D_{chro} = 30\ nm$ in the B-DNA conformation. Next, the number of bps in the nucleosome and linker are determined. The fixed parameter here, is the total number of bp per nucleosome, which is 200 in the B-DNA configuration [57]. The helix rise per bp is s and the length of the helix in both the nucleosome and the linker is l_{nucl} and l_{bind} respectively, so

$$200bp = \frac{l_{nucl} + l_{bind}}{s} \tag{5.2}$$

with

$$\begin{aligned} l_{nucl} &= 4\pi(R_{nucl} - R_{Bhel}) \\ l_{bind} &= \frac{\pi}{3}\left(\frac{D_{chro}}{2} - R_{Bhel}\right) \end{aligned} \tag{5.3}$$

Inserting $D_{chro} = 6R_{nucl}$ and solving Eq.5.2 for R_{nucl} it is found that

$$R_{nucl} = \frac{2}{10\pi}\left(200s + \frac{13\pi}{3}R_{Bhel}\right) = 5.23nm \tag{5.4}$$

The number of bps in the nucleosome and the linker are found using the parameters of the model (Table 3), putting this R_{nucl} back into Eq. 5.3 and dividing by s. Since these values should be integer, they have to be rounded, inducing a small correction over the value of R_{nucl}. The result is that there are 154 bp in the nucleosome and 46 bp in the linker fragment. The final value of R_{nucl} remains (the correction is too tiny) $5.23nm$. Thus, $D_{chro} = 31.44nm$. All this is consistent with the already mention requirement that the chromatin should have $\sim 30nm$.

Z-DNA

In chapter 2 was commented the main features about this conformation, being its zig-zag (hence the name) backbone the most relevant geometrical aspect to remark. Due to this characteristic, Z-DNA is quite stiff when compared with the other types of DNA. Then, it would be rare to find it in nucleosomes. The linker fragment is less curved than the nucleosome, thus, an hybrid model with B- and Z-DNA is constructed.

In this combination Eq. 5.1 do not hold anymore because the Z-helix will be placed in the linker and its axial step is greater thus, separating the nucleosomes. Eqs. 5.2 and 5.3 have a slight modification because helix radii and rises are different in each case. Keeping the optimized parameters [2] of 154bp in the nucleosome and 46 in the linker, The second equation of 5.3

$$6l_{bind} = 2\pi\left(\frac{D_{chro}}{2} - R_{Bhel}\right) \tag{5.5}$$

Can be easily solved. Dividing by $0.37nm/bp$ (Z-helix axial step) and noting that $l_{bind}/0.37 = 46bp$ the chromatin diameter is found for this model, $D_{chro} = 34.87nm$. Again, this value is of the order of the expected.

A-DNA

This is the least common form of DNA found in humans. Its main geometrical properties are: the greatest radius and smallest helix rise per bp among the three conformations treated here. In this model it wouldn't be impractical to keep the same parameters as in the previous models because 154 bp would form an excessively overlapped nucleosome and even worse, the binding fragments wouldn't be long enough to avoid that several nucleosomes overlap as well. In this way, is not possible to use Eq 5.2 since the total number of bps in the nucleosome, cannot be set *a priori*.

The situation is quite the same as in the homogeneous BB-model but with different parameters (D_{hel}, s) therefore, the solution goes quite the same as in the first model but 'scaled' according to the new parameters. The approach to follow is to keep the histone

[2] This make sense because the Z conformation occurs as a transition from B-DNA so, it is fair to say that neither the number of bp per nucleosome nor per linker changes.

radius of the BB-model, $R_{nucl} - D_{Bhel} = 4.05nm$, and solve eqs. 5.1 and 5.3 for the number of bps and the chromatin diameter. The idea behind this approach is that, irrespective of the kind of DNA present in the nucleus, the octamer has its own dimensions. As the standard form of DNA is the B type, structural features and dimensions in this configuration may be regarded as default.

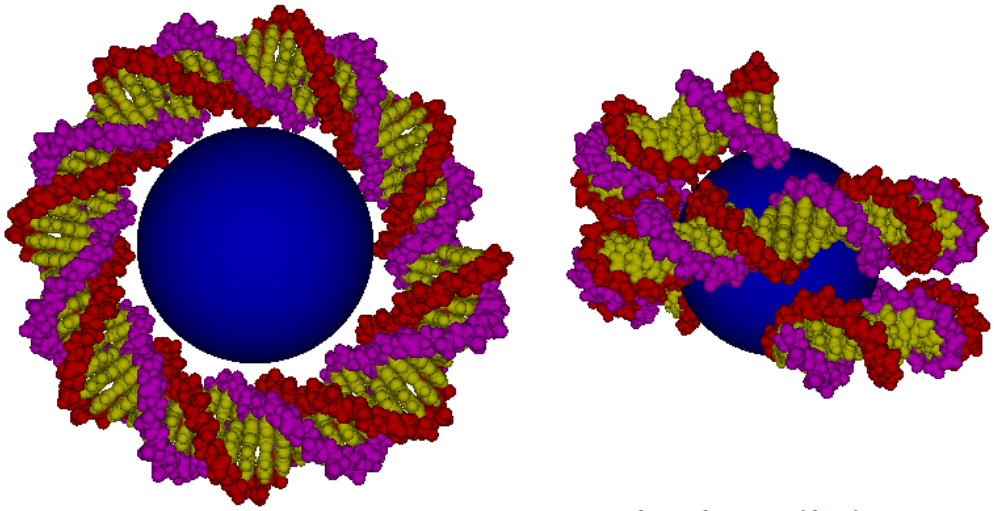
The nucleosome radius is the sum of the histone radius plus the helix diameter. This means that, in the current approach, $R_{nucl} = 4.05\ nm + D_{Ahel} = 5.41\ nm$. The chromatin diameter is (by Eq. 5.1) $D_{chro} = 32.46\ nm$. Knowing R_{nucl}, Eq. 5.3 divided by $s = 0.23\ nm/bp$ can be used directly to compute the number of bps in each structure. The final values are: 224 bp in the nucleosome, 70 bp in the linker fragment and the above R_{nucl} and D_{chro}.

Table 4 shows the results found in the optimization process. At a first glance, the fact that all chromatin diameters be of the order of $30nm$ gives some confidence in the models. Nevertheless, in **??** will be shown the results of the validation tests made to these constructions.

Table 4 – Summary of the geometrical parameters of the DNA structures found in the optimization process. In the model's name, the first letter stands for the type of DNA in the nucleosome and second refers to the linker fragment.

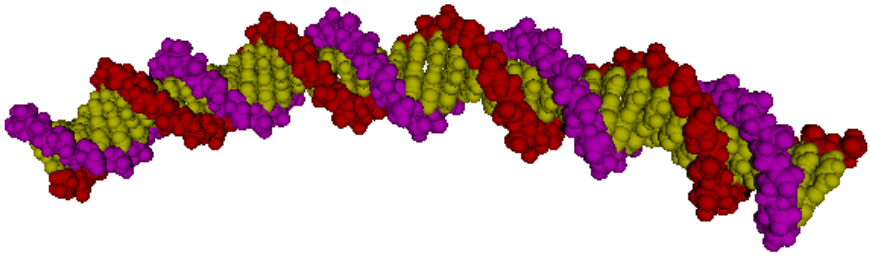
	Chromatin Model		
	AA	BB	BZ
Nucleosome Radius	5.41 nm	5.23 nm	5.23nm
Chromatin Diameter	32.46 nm	31.48 nm	34.87 nm
bps per nucleosome	224	154	154
bps per linker	70	46	46

Having all the geometrical parameters of the base pairs, helices, nucleosomes and linker fragments, the chromatin fiber can be constructed now. Each bp in the helix is created by rotation and translation of the building block coordinates, as a whole. This is done in the bp system of reference (SR). The new bps may belong to the nucleosome or to the linker fragment. In each case, there is a set of transformations that give its coordinate relative to its corresponding structure as well as relative to the chromatin system. What is meant to these different systems of reference is illustrated in Figs. 12 and 13. These are nothing but right-handed systems with origin in their corresponding centers (except for the bp in which the SR is given). Figs. 8-10 along with Fig. 6 show the resulting DNA structures as built up using the radii, diameters, axial steps and number of bps encountered for each configuration. In Appendix A is shown a POVray script to generate those figures. The important issue to note is that there is no need to specify other parameters than the aforementioned to get the designs.

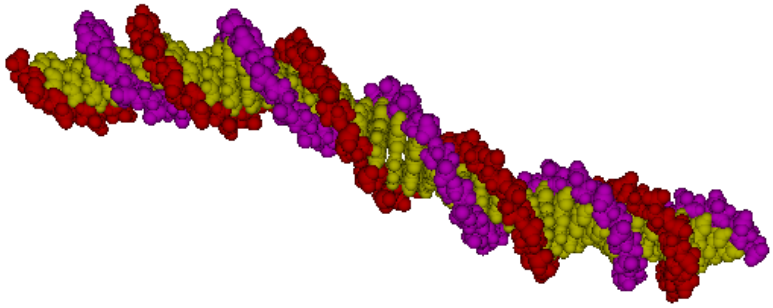


(a) Nucleosome Upper View (b) Nucleosome Side View

Figure 8 – BB-model of the nucleosome obtained with the parameters of tables 3 and 4.

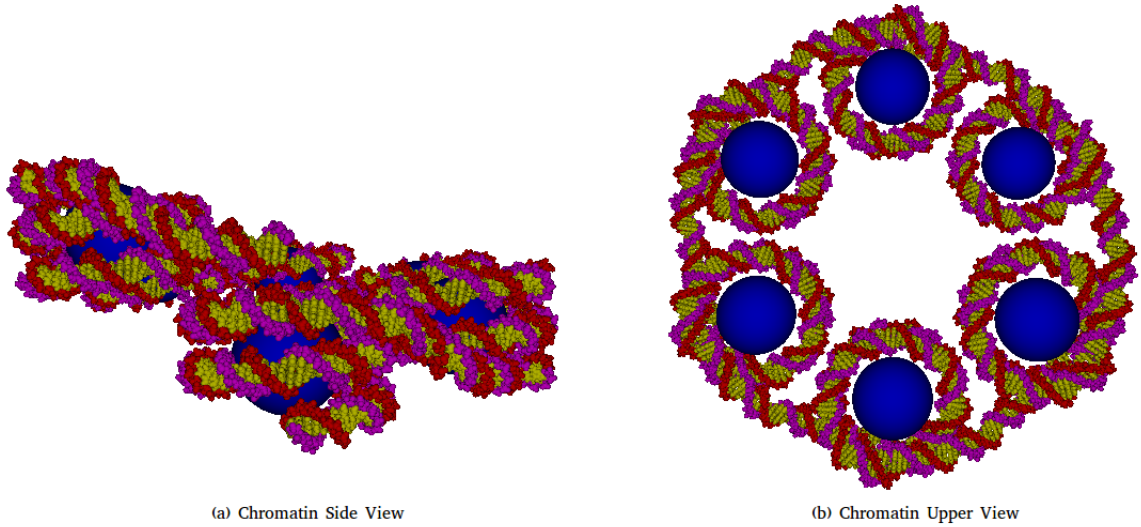


(a) Linker Fragment Upper View



(b) Linker Fragment Side View

Figure 9 – BB-model of the linker fragment. Each fragment links the top of a nucleosome with the bottom of the next one. The parameters of tables 3 and 4 permit to create this structure.



(a) Chromatin Side View (b) Chromatin Upper View

Figure 10 – BB-model of the chromatin. Six nucleosomes bound by linker fragments conform a turn of the helix. The helix pitch corresponds to $3D_{hel}$.

5.1.3 Cell Nucleus

To finally recreate a cell nucleus, several chromatin fibers are put together to form a squared lattice. There are about 6×10^9 bps in the human cell nucleus, thus, the amount of fibers is chosen such that it contains this number of bps. To estimate the number of fibers, the following equation was used,

$$(6 \times (N_{bp}^{nucl} + N_{bp}^{bind}) \times N_{chro}) \times N_{fib} \approx 6 \times 10^9 \tag{5.6}$$

where $N_{bp}^{nucl}, N_{bp}^{bind}$ is the number of bps in the nucleosome and linker respectively, N_{chro} stands for the number of chromatin levels in a single fiber and N_{fib} is the number of fibers in the nucleus. The nucleus was modeled as a straight cylinder with axis parallel to the chromatin z-axis. As the typical diameter of the nucleus is about 3 μm, the number of fibers in the xy-plane N_{fib} is chosen $10^2 \times 10^2$, that is, 100 fibers in each direction.

In the BB-model (standard) $N_{bp}^{nucl} + N_{bp}^{bind} = 200bp$. Using Eq. 5.6 and demanding those $3\mu m$ at least, $N_{chro} = 475$ is selected. This gives $5.7 \times 10^9 bps$ in the BB-nucleus. Since the BZ-model has the same number of bps per nucleosome + linker as the BB-model, the same choices of N_{chro} and N_{fib} remain. For the AA case the physical dimensions of the structure are considered as the fixed parameters again, thus, N_{chro} and N_{fib} are taken from the BB-model as well. This gives $8.4 \times 10^9 bps$ in the AA-nucleus. Despite this amount of bps do not correspond to reality, [3] all damage yield calculations will be done per bp per unit dose, removing the dependence on the number of bps.

[3] Actually, there are no AA-nuclei, but still, from the statistical point of view, it is worth to construct such model since is like if A-DNA were placed in all possible positions inside the nucleus.

5.2 *GEANT4-DNA* settings

Microdosimetry example

Among the examples included in the release 10.1 of the *GEANT4-DNA* code, there is the *microdosimetry* user code. This is a code to simulate a point-like source of protons that pass through a 1 mm^3 water phantom. There are specific classes in the code that define and control both the external and internal parameters of the simulations. External parameters mean those defined by the user (geometries, type of particles, incident energies, verbose). Internal parameters refer to default settings of the simulation such as the definition of particles, energy cutoff's, physical processes to simulate, initialization parameters, etc.

In the **DetectorConstruction.cc** class, the geometry of the detector is set defining the *World* and *Target* region. The *World* is the region in which particles do exist and propagate while the *Target* is the region in which they will be detected. In the *microdosimetry* example, the world has dimensions of 1 $mm \times 1\ mm \times 1\ mm$ and the target $1\ mm \times 1\ mm \times 0.05\ mm$, This means that the particle's information to be stored is that generated at a 0.05 mm thick slice of the cube and not in the whole space. The moral is that when constructing a detector there are two kinds of regions that have to be specified and the *DetectorConstruction.cc* class is where this has to be done. Also, the material of the detector can be defined in the same class and, if necessary, new materials may be created indicating its components and density.

The beam is set in the **PrimaryGeneratorAction.cc** class. Here, the initial position and direction of the beam particless (primaries) are defined. For the *microdosimetry* example, these are simply the point $(0\ mm, 0\ mm, -0.5\ mm)$ and the vector $(0, 0, 1)$ but any source geometry can be implemented. Every primary particle is called a *history*.

What and how the information is stored is coded in the **SteppingAction.cc** class. This file contains the lines that determine what variables (event position, energy deposited, energy transferred and the like) will be copied to the output file. Also, the very format of the output file may be defined here. In the **SteppingAction.cc** class are specified the labels of the particles and physical processes as well.

Starting from the *microdosimetry* example, the irradiation settings of the nuclei can be prepared. Instead of put those nuclei in the target region and transport the particles through the genetic material directly, the approach to follow is,

1. Choose or create the material that mimics the nucleus composition in the **DetectorConstruction** class .

2. Since nuclei are modeled as straight cylinders, define the *World* and *Target* regions with this geometry as well. The dimensions of each region must be consistent with

the dimensions discussed in subsection 5.1.3.

3. Modify the **PrimaryGeneratorAction** class to switch from a point-like source to a uniform squared beam.

4. Modify the **SteppingAction** class to get the output file in binary format. The .bin file must contain a table with,

   - **Flag History** The label of the primary particle corresponding to the event.
   - **Flag Particle** The type of particle followed.
   - **Flag Process** The label of the physical process suffered by the particle.
   - **Position** Coordinates of the event (Physical Process).
   - **Energy Deposited** Energy locally absorbed by the medium.
   - **Energy Transferred** Energy transferred by the particle to another particle or molecule, that is not deposited locally.
   - **Kinetic Energy**.

5. Set the type, number and initial energy of primary particles. These parameters are passed to the **PrimaryGeneratorAction** class through the external file *micro-dosimetry.mac*. For a given species and energy, the number of particles is selected aiming to deliver the same dose in all cases.

Set all these up, the application can be compiled and ran.

5.2.1 ROI geometry and properties

To simulate the nucleus composition, the detector is made up of water but with density scaled to 1.06 g/cm^3 in order to simulate the cell composition. This is accomplished by a simple redefinition of the already water material included in the *GEANT4*. This redefinition has direct impact in the sampling process because the distance between two consecutive interactions depends on μ/ρ (the mass attenuation coefficient) c.g. the first example of subsection 4.1.2.

The *World* has dimensions 10 $\mu m \times 10\ \mu m \times 25\ \mu m$ which is some bigger than the nucleus dimensions because electronic equilibrium must be guaranteed. That is, electrons may be created outside the nucleus but travel toward this, depositing energy hence, affecting the dose delivered to the nucleus. The *Target* is constructed with the same dimensions as the world because the *GEANT4* makes condensation of histories [4] outside the target region which means that the transport is not so detailed there. This is not

[4] This time, *history* refers to a single event suffer by the particle, in the jargon of MC.

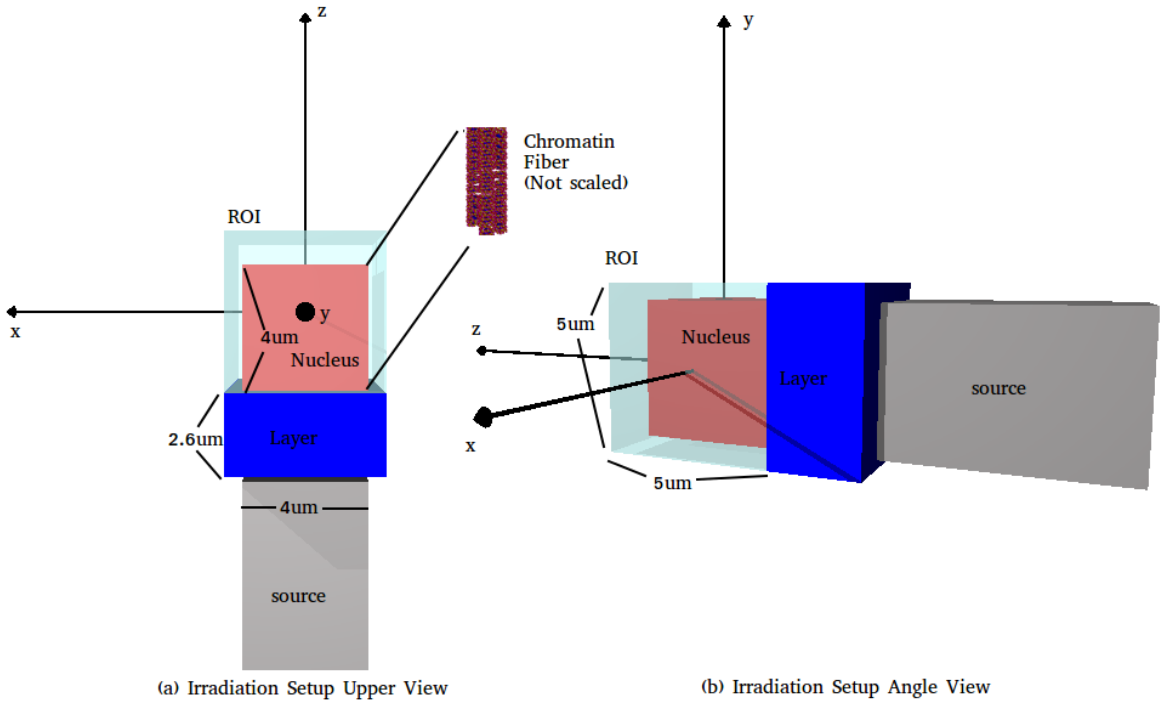


Figure 11 – Irradiation setup. The ROI and the layer are immersed in the *World* so, particles are transported step by step within this region. The *GEANT4* origin has been shifted $-7.4\mu m$ along the z axis to facilitate its visualization.

exactly a problem since condensation algorithms are well studied and widely used, but having such small region and a relatively powerful machine, is not necessary to do it.

In the **SteppingAction** class, the output file is prepared as explained in step 4 above. The events recorded are those inside a cube of dimensions 5 $\mu m \times 5\ \mu m \times 5\ \mu m$ coaxial to the target straight cylinder. This is the Region Of Interest (ROI). The bottom face (in the xy-plane) is just 2.6 μm above the bottom of the *World* cylinder, i.e. the cube is not centered along the z-axis. The 2.6 μm emulates the water equivalent layer commonly used in *in vitro* cell irradiation setups [58].

5.2.2 Beams

The nuclei was irradiated with protons and alpha particles of different energies. The chosen energies are $0.5, 1, 5, 7, 10\ MeV$ for H^+ and $2, 5, 7, 10\ MeV$ for He^{2+}. These energies are found around the Bragg peak no matter what the primary beam energy is. Several of these peaks are superimposed to obtain a relatively homogeneous biological dose distribution inside the tumour volume in hadrontherapy. In order to account for statistical fluctuations of the future quantities to be studied, thousands of primary particles were simulated for every radiation quality (see Table 5) to subsequently split the .bin files into minor files intended for statistical analysis (simulation batches). As mentioned above, there is an external file called *microdosimetry.mac* that passes the type and energy of

primary particles to the **PrimaryGeneratorAction** class. Thus, several .mac files and a splitting FORTRAN routines were written down to perform the simulations.

Table 5 – Number of particles of every quality simulated with the *Geant4-DNA*. The phase-space files were split afterwards.

Beam Quality		Number of Particles
H^+	0.5 MeV	9000
	1 MeV	1200
	5 MeV	3600
	7 MeV	5400
	10 MeV	58500
He^{2+}	2 MeV	840
	5 MeV	360
	7 MeV	450
	10 MeV	7590

The only modification in the **PrimaryGeneratorAction** class that was made is in the source geometry. The initial position of a primary is uniformly sampled over a 4 $\mu m \times 4$ μm square in the $z = -12.5$ μm plane, being z the symmetry axis. This means that the source is right before the water equivalent layer commented in subsection 5.2.1.

The irradiation setup is shown in Fig. 11. All the incident particles have initial momentum pointing along the z-direction. Before writting the output files, the coordinates are translated by $(0, 0, 9.9\mu m)$ in order to place the bottom of the nuclei in the xy-plane as in the geometrical models. This will be referred as the Nucleus System of Reference (NSR).

5.3 FindTheClosestAtom Algorithm

In the approach of section 5.1 seems like if the structures were to be constructed and the coordinates of millions of bps were to be stored. From the computational point of view this would be catastrophic since the location of every physical event to be simulated would have to be compared with those millions of coordinates in order to know what structure was impacted. To save computational time, the rationale of the *FindTheClosestAtom* (FCA) algorithm developed by Bernal *et al.* [56] was followed. Some modifications were done in the formalism. The algorithm is intended for reading a position in the nucleus system of reference and successively transform it into local structure-coordinate systems (i.e. Chromatin fiber-, Nucleosome- and bp-coordinates).

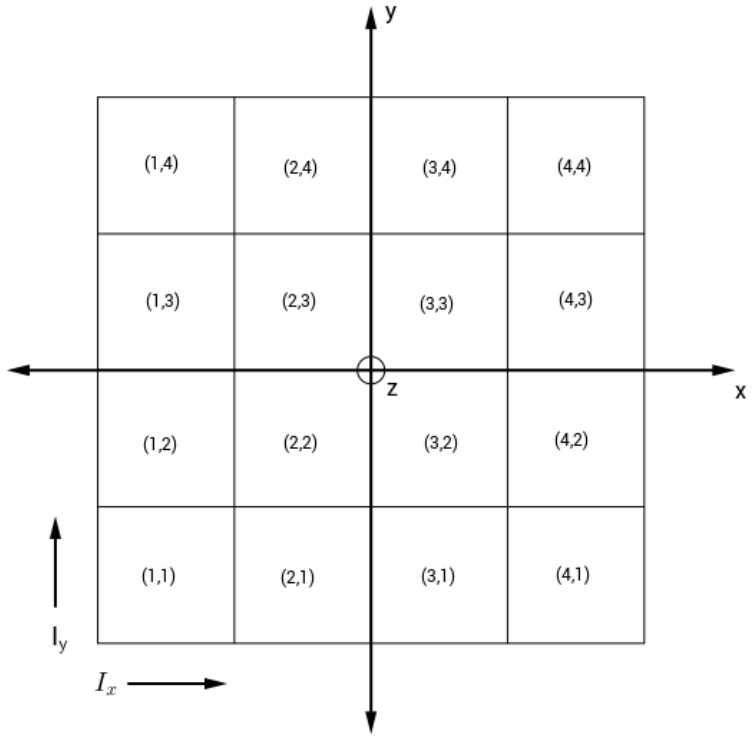


Figure 12 – Scheme of the nucleus grid and the index convention. Every square represents a chromatin fiber viewed from the top (Fig. 13). The coordinates shown would represent the Nucleus-coordinates. The actual grid has 100×100 chromatin fibers.

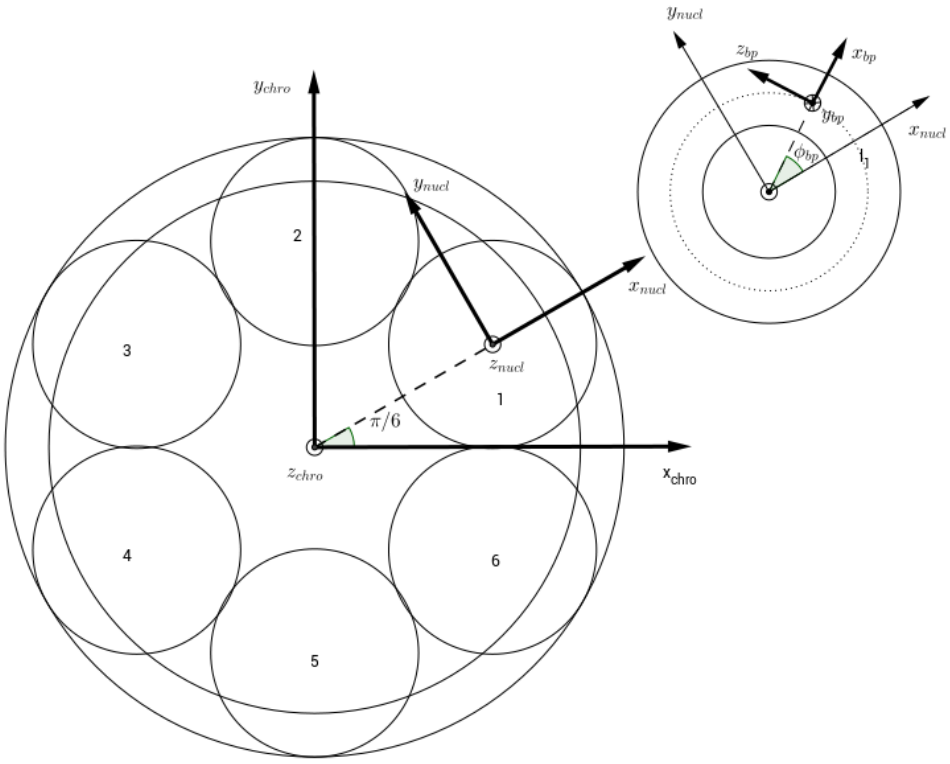


Figure 13 – Relation among the systems of reference of each DNA structure. The linker fragment's SR shares the x-axis with the nucleosome but has its origin in the origin of the chromatin's SR. Also, the nucleosomes are labeled according to the FCA algorithm convention.

Given a point $(x_{hit}, y_{hit}, z_{hit})$ in space (NSR), the steps to determine if the point lies within one corresponding Van der Waals radius of one or more atoms belonging to the sugar-phosphate group of the bp, are:

1. Check if the point lies inside the nucleus. As the nucleus is made up of 100×100 chromatin fibers, its dimensions are $100D_{chro} \times 100D_{chro} \times 475 \cdot 3 \cdot D_{hel}$, thus, the condition $(x_{hit}, y_{hit}, z_{hit}) \in [-50D_{chro}, 50D_{chro}] \times [-50D_{chro}, 50D_{chro}] \times [0, 475 \cdot 3 \cdot D_{hel}]$ must hold to further proceed.

2. Compute the *Chromatin Index*, that is, the pair $(I_x, I_y) \in [1, 100] \times [1, 100]$ as shown in Fig. 12. The index allows to express (x_{hit}, y_{hit}) in Chromatin-coordinates (x_{chro}, y_{chro}).

3. Determine to which nucleosome or linker (x_{chro}, y_{chro}) the point belongs to and find the corresponding *Nucleosome Index* I_{nucl} according to Fig. 13.

4. Having I_{nucl}, the vertical level in the chromatin fiber and z_{chro} can be determined .

5. Check in what structure, nucleosome or linker fragment, the point $(x_{chro}, y_{chro}, z_{chro})$ lies.

6. The local coordinates $(x_{nucl/bind}, y_{nucl/bind}, z_{nucl/bind})$ now can be computed.

7. Next, the index of the closest bp within the structure has to be found.

8. Transform the local-structure-coordinates into bp-coordinates and check, for every atom in the phosphodiester group, if the distance between the atom and the transformed point is less than one atom's Van der Waals radius.

In what follows, the coordinates mentioned are those of the hit after transformations and they should not be confused with the axis labels of Fig. 13, which were for reference purposes. Additionally, $INT(X)$ indicates the integer part of X.

5.3.1 Chromatin Index

Once the point have been verified to be inside the nucleus, the Chromatin Index (I_x, I_y) is determined as,

$$I_x = 1 + INT\left(\frac{x_{hit} - x_{min}}{D_{chro}}\right) \tag{5.7}$$

$$I_y = 1 + INT\left(\frac{y_{hit} - y_{min}}{D_{chro}}\right) \tag{5.8}$$

With x_{min} and y_{min} equal to $-50D_{chro}$ so the index goes according to the scheme of Fig. 12. Now, x_{chro} and y_{chro} are found, as well as the sextant (nucleosome) in which the hit resides.

$$x_{chro} = x_{hit} - (x_{min} + (I_x - 0.5)D_{chro}) \tag{5.9}$$

$$y_{chro} = y_{hit} - (y_{min} + (I_y - 0.5)D_{chro}) \tag{5.10}$$

$$\phi_{chro} = \arctan\left(\frac{x_{chro}}{y_{chro}}\right) \quad \in [0, 2\pi) \tag{5.11}$$

$$I_{nucl} = 1 + INT\left(\frac{3\phi_{chro}}{\pi}\right) \tag{5.12}$$

Eq. 5.12 means that the first nucleosome is in the sextant $[0, \pi/3)$, the second is in $[\pi/3, 2\pi/3)$, the third in $[2\pi/3, \pi)$, and so on. This convention was depicted in Fig.13. After one turn of the chromatin fiber, nucleosomes with the same index are shifted by $3D_{hel}$ in the z direction. Also, there is another z-shift of $3D_{hel}/6$ between two contiguous nucleosomes. This rise goes in the crescent sense of I_{nucl}. Thus, the vertical level I_z of the chromatin fiber depends on I_{nucl} as

$$I_z = 1 + INT\left(\frac{z_{hit} - (I_{nucl} - 1)3D_{hel}/6}{3D_{hel}}\right) \tag{5.13}$$

Furthermore, the z-coordinate of the bottom of the nucleosome can be obtained

$$z_{min}^{nucl} = (I_z - 1)3D_{hel} + (I_{nucl} - 1)3D_{hel}/6 \tag{5.14}$$

The origin of the Nucleosome-coordinates is displaced by $RC_{chro} = (D_{chro} - R_{nucl})$ along the x-chromatin axis and rotated by $\beta = (I_{nucl} - 0.5)\pi/3$ with respect to the z axis in the Chromatin-coordinates. Thus, the coordinate transformation of the hit reads,

$$x_{nucl} = x_{chro}\cos\beta + y_{chro}\sin\beta - RC_{chro} \tag{5.15}$$

$$y_{nucl} = -x_{chro}\sin\beta + y_{chro}\cos\beta \tag{5.16}$$

$$z_{nucl} = z_{hit} - z_{min}^{nucl} \tag{5.17}$$

Having both the Chromatin- and Nucleosome-coordinates of the hit, the next step is to figure out in which structure it lies taking into account that the impacted atoms are being looked for. This condition excludes all the hits lying inside histones or empty regions.

5.3.2 Nucleosome Index

If the hit impacted any bp in the nucleosome, it is mandatory that the next conditions hold

$$\left|\sqrt{x_{nucl}^2 + y_{nucl}^2} - RC_{nucl}\right| < D_{hel}/2 \tag{5.18}$$

$$0 < z_{nucl} < 3D_{hel} \tag{5.19}$$

Where $RC_{nucl} = R_{nucl} - D_{hel}/2$. If Eq. 5.18 or 5.19 is not satisfied, the nucleosome analysis does not proceed any further. If both conditions are verified, the angular position of the hit is calculated,

$$\phi_{nucl} = \arctan\left(\frac{x_{nucl}}{y_{nucl}}\right) \quad \in [0, 2\pi) \tag{5.20}$$

Since the nucleosome has two loops, there are two bps at a given angular position ϕ_{nucl}. One bp is at a height $z_1 = (1 + \phi_{nucl}/\pi)D_{hel}/2$ and the other one is at $z_2 = (3 + \phi_{nucl}/\pi)D_{hel}/2$.

Let us define $R_{hit} = \sqrt{x_{nucl}^2 + y_{nucl}^2}$. If

$$\sqrt{(R_{hit} - R_{nucl})^2 + (z_{hit} - z_1)^2} < D_{hel}/2 \tag{5.21}$$

Then, the closest bp to the hit is in the first loop and has index

$$I_{bp} = 1 + INT\left(N_{bp}^{nucl}\frac{\phi_{nucl}}{4\pi}\right) \tag{5.22}$$

The 4π comes from the fact that the N_{bp}^{nucl} bps are disposed along two loops. On the other hand, if Eq. 5.21 is fulfilled by z_2 instead of z_1, the bp index is

$$I_{bp} = 1 + INT\left(N_{bp}^{nucl}\frac{2\pi + \phi_{nucl}}{4\pi}\right), \tag{5.23}$$

so it is in the second (highest) loop. This I_{bp} and the other indexes and coordinates allow to do the *Base Pair Analysis* in subsection 5.3.4.

5.3.3 Linker Index

To considered that the hit impacts a bp, it has to reside within the linker shell defined as the cylindrical shell with radius $RC_{bind} = (D_{chro} - D_{hel})/2$ and thickness D_{hel}^{bind} [5]. Therefore, it is necessary to check if

$$\left|\sqrt{x_{chro}^2 + y_{chro}^2} - RC_{bind}\right| < D_{hel}^{bind}/2 \tag{5.24}$$

[5] In the hybrid model D_{hel} is the diameter of the B-DNA according to the considerations in the BZ-model construction.

holds. The binding fragments are contained in two sextants as the way they were defined. To correct this and finally assign a single index to the linker fragment, the angular position must be redefined as follows,
If ($\phi_{chro} \geq \pi/6$) then

$$\phi_{bind} = \phi_{chro} - \frac{\pi}{6}$$

Else

$$\phi_{bind} = \phi_{chro} - \frac{\pi}{6} + 2\pi$$

Now, the binding index I_{bind} and the vertical level, are obtained in the same way than in the nucleosome case.

$$I_{bind} = 1 + INT\left(\frac{3\phi_{bind}}{\pi}\right) \tag{5.25}$$

$$I_z^{bind} = 1 + INT\left(\frac{z_{hit} - (I_{bind} - 1)3D_{hel}/6}{3D_{hel}}\right) \tag{5.26}$$

$$\phi_0 = \frac{\pi}{3}\left(I_{bind} - 0.5\right) \tag{5.27}$$

The third equation just gives the initial angular position of the linker I_{bind}. Then, the index of the closest bp to the hit[6] is,

$$I_{bp} = 1 + INT\left(N_{bp}^{bind}\frac{\phi_{bind} - \phi_0 + \pi/6}{\pi/3}\right) \tag{5.28}$$

5.3.4 Base Pair Index

As the chromatin fiber is made up of a long strand of DNA that condenses around histones proteins, the calculation of the bp twist around the helix axis must take into account not only I_{bp} but all the bps before the corresponding nucleosome or linker. The other important issue to discuss is the possibility for a single hit to lie inside two sugar-phosphate group's atoms. This overlapping may occur among atoms of the same bp or even among atoms of different bps. The former case is regarded as one Single Strand Break (SSB), so that a hit impinging on several atoms of the same molecule is counted once. The latter case is called a SSB+ and both bp indexes have to be stored for future treatment.

The overlapping problem is puzzle out making the bp analysis for the closest bp as well as for the two nearest neighbors, i.e. for the $I_{bp} \pm 1$ indexes. This strategy has to do with the total bp index counting related to the twisting angle because the bp index has to be continuous[7]. The sequential bp index I_{seq} is obtained from I_{bp} as

[6] Despite the linker has a sigmoid shape in the plane parallel to z, the closest bp is chosen testing its distance to the hit in the xy-plane.

[7] The I_{bp} is calculated locally, thus, when the analysis is made over the first or the last bp of a given structure, one of the nearest neighbor belongs to another structure.

Nucleosome I_{bp}

$$I_{seq} = \underbrace{\Big(6(I_z - 1) + (I_{nucl} - 1)\Big)}_{\substack{\text{Number of blocks} \\ \text{before the actual nucleosome}}} \underbrace{\Big(N_{bp}^{nucl} + N_{bp}^{bind}\Big)}_{\text{bps in each block}} + I_{bp} \tag{5.29}$$

Linker I_{bp}

$$I_{seq} = \underbrace{\Big(6(I_z^{bind} - 1) + (I_{bind} - 1)\Big)}_{\substack{\text{Number of blocks} \\ \text{before the actual linker}}} \underbrace{\Big(N_{bp}^{nucl} + N_{bp}^{bind}\Big)}_{\text{bps in each block}} + N_{bp}^{nucl} + I_{bp} \tag{5.30}$$

A nucleosome and a linker fragment that have the same index in a given chromatin fiber form a block. Every block has $N_{block} = N_{bp}^{nucl} + N_{bp}^{bind}$ bps. Since the scanning includes several bps (one at the time), every structure index has to be computed from I_{seq} to account for moves across the structures. Given any of the $I_{seq}'s$ to be scanned, the indexes are recalculated by means of

$$I_z = 1 + INT\Big(\frac{I_{seq} - 1}{6N_{block}}\Big) \tag{5.31}$$

$$I_{block} = 1 + INT\Big(\frac{(I_{seq} - 1) - 6(I_z - 1)N_{block}}{N_{block}}\Big) \tag{5.32}$$

$$I_{bp} = I_{seq} - \underbrace{\Big(6(I_z - 1) + (I_{nucl} - 1)\Big)N_{block}}_{N_{bp0}} \tag{5.33}$$

This time I_{bp} runs from 1 to N_{block}. If $1 < I_{bp} < N_{bp}^{nucl}$, the bp is in the nucleosome. If $(N_{bp}^{nucl} + 1) < I_{bp} < N_{block}$, then the bp belongs to a linker. For each case, there is a separated bp analysis to consider but let us present some definitions first:

- Let γ_0 be the twist angle of the $N_{bp0} - th$ bp (see Eq. 5.33), then

$$\gamma_0 = \underbrace{2\pi\Big(\frac{N_{bp}^{nucl}}{n_{hel}^{nucl}} + \frac{N_{bp}^{bind}}{n_{hel}^{bind}}\Big)}_{\substack{\text{Angle twist} \\ \text{per block}}} \underbrace{\Big(\frac{N_{bp0}}{N_{block}}\Big)}_{\substack{\text{Number} \\ \text{of blocks}}} \tag{5.34}$$

  Where n_{hel}^{nucl} and n_{hel}^{bind} are the number of bps per helix turn in the nucleosome and linker respectively (Table 2).

- Let H_{bind} be the height of the linker fragment, that is, the distance from the top bp to the bottom bp in the plane parallel to z[8]. In this plane, it was said that the profile of the binding fragment has a sigmoid shape. Let β be the angular position of a point (of the linker) in the xy-plane. Then, this point is below the top bp by a distance $H_{bind}f(\beta, \mu)$, with

$$f(\beta, \mu) = 0.5(1 - s(1 - e^{-\mu|\beta|})) \quad \text{with} \quad s = \begin{cases} -1, & if\ \beta > 0 \\ 1, & otherwise \end{cases} \tag{5.35}$$

  where μ controls how rapidly the sigmoid curve falls.

[8] In chromatin-coordinates.

Nucleosome Analysis

The angular position ϕ_{bp}^{nucl} of the bp in nucleosome-coordinates is

$$\phi_{bp}^{nucl} = (I_{bp} - 0.5)\frac{N_{bp}^{nucl}}{4\pi} \tag{5.36}$$

Hence, the hit coordinates transform as

$$x_{bp}{}' = x_{nucl}\cos\phi_{bp}^{nucl} + y_{nucl}\sin\phi_{bp}^{nucl} - RC_{nucl} \tag{5.37}$$

$$y_{bp}{}' = -\left(z_{nucl} - \left(1 - \frac{\phi_{bp}^{nucl}}{\pi}\right)\frac{D_{hel}}{2}\right) \tag{5.38}$$

$$z_{bp}{}' = -x_{nucl}\sin\phi_{bp}^{nucl} + y_{nucl}\cos\phi_{bp}^{nucl} \tag{5.39}$$

Note that, as shown in Fig. 13, apart from the rotation among the two systems, there is a Y-Z flip too. Nevertheless, this position vector still has to be rotated around the z axis by

$$\gamma = \gamma_0 + 2\pi\left(\frac{I_{bp} - 1}{n_{hel}^{nucl}}\right) \tag{5.40}$$

To give the actual bp coordinates of the hit. Thus,

$$x_{bp} = x_{bp}{}'\cos\gamma + y_{bp}{}'\sin\gamma \tag{5.41}$$

$$y_{bp} = -x_{bp}{}'\sin\gamma + y_{bp}{}'\cos\gamma \tag{5.42}$$

$$z_{bp} = z_{bp}{}' \tag{5.43}$$

So far, the nucleus has been just an abstraction. However, the base pair does actually exist in the code. For each atom in the sugar-phosphate group with coordinates $(x_{atom}, y_{atom}, z_{atom})$ and radius R_{atom} the closest atom to the hit is searched. If the condition

$$\sqrt{(x_{bp} - x_{atom})^2 + (y_{bp} - y_{atom})^2 + (z_{bp} - z_{atom})^2} < R_{atom} \tag{5.44}$$

is met, a SSB is scored and the indexes $I_z, I_{block}, I_{bp}, I_{seq}$ and I_{atom} (every atom has an index inside the code) are stored. The first atom to reach the condition 5.44 activates a escape line so, the next bp is analyzed. If there is no next bp to scan, the algorithm end and the next event is allowed.

Linker Analysis

The approach to the linker analysis starts, again, computing the angular position of the closest bp, this time labeled as ϕ_{bp}^{bind}.

$$\phi_{bp}^{bind} = \Big(\underbrace{I_{bp} - N_{bp}^{nucl}}_{\substack{I_{bp}{}' \\ \text{local bp index}}} - 0.5\Big)\frac{\pi}{3N_{bp}^{bind}} \tag{5.45}$$

The binding fragment of a given block, starts to fall from the top of the nucleosome at a height

$$z_{top} = (I_z - 1)3D_{hel} + 2.5D_{hel} + (I_{block} - 1)3D_{hel}/6 \tag{5.46}$$

The height of the linker is $H_{bind} = 1.5D_{hel}$. The sigmoid curve defined in Eq. 5.35 is symmetrical with respect to the x-axis of the nucleosome [9] so the angle ϕ_{bp}^{bind} must be shifted by $-\pi/3$ to find β. Bringing this together, the z coordinate of the bp is

$$z_{bp}^{bind} = z_{top} - H_{bind} f(\beta = \phi_{bp}^{bind} - \pi/6, \mu = 20) \tag{5.47}$$

The angular position of the bp in chromatin-coordinates is

$$\phi_{bp} = (I_{block} - 0.5)\frac{\pi}{3} + \phi_{bp}^{bind} \tag{5.48}$$

Thence, the hit (intermediate) bp-coordinates are

$$x_{bp}{}' = x_{chro} \cos \phi_{bp}^{bind} + y_{chro} \sin \phi_{bp}^{bind} - RC_{bind} \tag{5.49}$$
$$y_{bp}{}' = -(z_{chro} - z_{bp}^{bind}) \tag{5.50}$$
$$z_{bp}{}' = -x_{chro} \sin \phi_{bp}^{bind} + y_{chro} \cos \phi_{bp}^{bind} \tag{5.51}$$

To finally get the hit bp-coordinates, the twist must be considered. That is, the rotation around the DNA helix axis by γ has to be performed

$$\gamma = \gamma_0 + 2\pi \frac{N_{bp}^{nucl}}{n_{hel}^{nucl}} + 2\pi \left(\frac{I_{bp}{}' - 1}{n_{hel}^{bind}} \right) \tag{5.52}$$
$$x_{bp} = x_{bp}{}' \cos \gamma + y_{bp}{}' \sin \gamma \tag{5.53}$$
$$y_{bp} = -x_{bp}{}' \sin \gamma + y_{bp}{}' \cos \gamma \tag{5.54}$$
$$z_{bp} = z_{bp}{}' \tag{5.55}$$

Now the Eq. 5.44 can be applied with the same spirit as in the nucleosome analysis.

[9] In the caption of Fig. 13 was stated that the linker and the nucleosome share the same x-axis

6 Validations

6.1 DNA models

The most common target for radiation-induced direct damage is the sugar-phosphate group and not the whole base pair [48]. For this reason, the effective volume of the sugar-phosphate group was determined. This target is defined as the union-volume of the sugar-phosphate group atoms, having into account the overlapping between two consecutive bps. The site-hit probability then was computed as the ratio of the total target volume to that of the whole nucleus, where energy depositions occur. Since the molecule was constructed as a union of spheres (atoms), its volume was calculated *À la Monte Carlo*. The effective volume in the helix configuration was computed as follows:

1. A box of dimensions $3nm \times 3nm \times 10nm$ was constructed.

2. One B-bp was placed inside the box and the all the atom's radii were set to 0 nm except those of the sugar-phosphate group. For these atoms, the radii were set to one corresponding Van der Waals' radius (Table 6).

Table 6 – Van der Waals radii of the sugar-phosphate group's atoms [2].

Atom	Radius[nm]
H	0.12
O	0.14
N	0.15
C	0.17
P	0.19

3. N points were randomly and uniformly sampled inside the box and the number of points n lying inside any atom was counted. Only the sugar-phosphate group's atoms were considered since their radii were not set to 0 nm. As in the FCA algorithm, there was an escape line once an atom was reached.

4. The union volume was computed as the ratio of n to N, times the volume of the box.

$$V_{eff} = \left(\frac{n}{N}\right)(90nm^3) \tag{6.1}$$

5. The process was repeated from step 3 ten times and the various $V_{eff}'s$ were averaged.

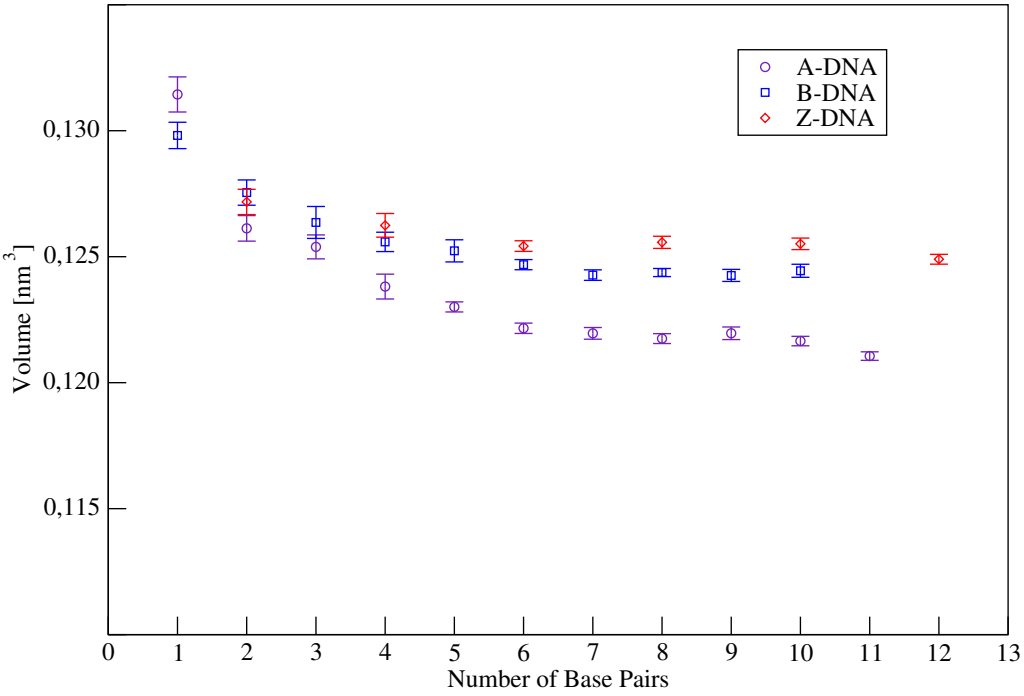


Figure 14 – Effective volume of the bps as determined by MC with $N = 10^6$ points sampled.

6. A new bp was put inside the box as forming the helix and the steps 3 to 5 were repeated. This time, averaging over the number of bps as well. This was made till the number of bps per helix turn was attained.

7. Steps 2 and the following were carried out for all the DNA conformations, one at the time.

Table 7 – Effective volume of bps in the helix configuration. Ending values of Figure 14.

DNA conf.	$V_{eff}[nm^3]$
A	0.122
B	0.124
Z	0.125

The effective volumes of each conformation are shown in Fig. 14 and summarized in Table 7. As the number of bps grows the effective volume converges to a definite value in each DNA model. This approach considers the overlapping in the helix as standard unlike the overlapping due to the bending around histones in the nucleosome, which appears as circumstantial. The effective volume is taken as a quality indicator because the probability of damage is defined as volume ratio. This ratio will be compared with the number of strand breaks to the number of events inside the nucleus ratio (see section 7.1) so, volumes consistent with the DNA structure are required. The values found showed to be in agreement with other values reported in the literature [50, 59]. The chromatin

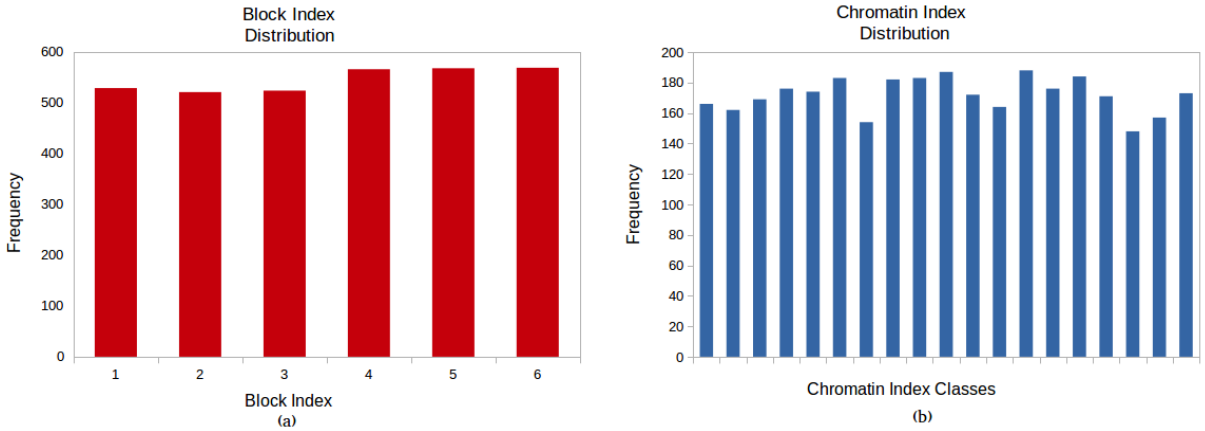


Figure 15 – (a) Block and (b) Chromatin Index distributions obtained in the tests of the FCA algorithm. In (b) the chromatin classes have size 25. An uniform distribution was expected since the points were shot at random.

parameters obtained by optimization also proved to be consistent with those reported in standard bibliography since they kept close to the $30nm$ chromatin diameter and number of bps per nucleosome.

Despite the squared nucleus model do not account for chromosomes neither in geometry nor in spatial domains, what will be investigated are just *Single* and *Double Strand Breaks*. The linear dimensions of such kind of damages are of about $3nm$, this means that, in principle, there is no dependence on the structure dimensions beyond the chromatin level which is ten times larger. Another argument in favour of the squared nucleus is that only physical damage (the very early stage of damage) unlike chemical or even biological damage will be studied, thus, chromosome differentiation is not needed.

6.2 FCA algorithm

The algorithm explained in section 5.3 was implemented in FORTRAN. All the parameters and bp coordinates were defined within the code as well as the geometrical transformations. The algorithm input is $(x_{chro}, y_{chro}, z_{hit})$ [1] and the output is the set $\{I_z, I_{block}, I_{bp}, I_{at}, I_{seq}\}$ where all but I_{seq} are 2-arrays in order to store possible SSB+'s. To check auto-consistency of the algorithm, 4×10^6 points uniformly distributed over the chromatin fiber dimensions, that is, $[-D_{chro}/2, D_{chro}/2] \times [-D_{chro}/2, D_{chro}/2] \times [0, 475 \cdot 3D_{hel}]$ were randomly sampled. Every time a structure, from chromatin down to phosphodiester's atom, was reached, its index was stored. As the points were shot at random, it is expected that the distribution of indexes be uniform if the algorithm has no biases. The histograms constructed from the output of the test are shown in Fig. 15. It is fair to say that the distribution can be consider uniform since variations in the frequencies are relatively small

[1] Note that it starts from item 3 of section 5.3

among the indexes. The chromatin level indexes are grouped in classes of size 25 to avoid an unreadable histogram of 475 classes.

Another test is performed calculating, by hand, the explicit position in the NSR of a given atom and using it as an input. Of course, the algorithm must retrieve the same atom index as the input. This procedure was carried out for several atoms having perfect matches every time.

7 The XX.for algorithm

The radiation-induced damage starts with the physical effect of the ionizing particle on the biological system. Such effects might be ionization, excitation or elastic scattering. For the different particles transported, that is, for both primaries and secondary particles, all these possible events were taking into account in simulations. They were included in the class **PhysicsList.cc** of the *microdosimetry* example. Implementation and validation of the physical models used in the *GEANT4-DNA* can be consulted in [60–63].

The aim of this work is to study the influence of the DNA conformation on the probability of DNA damage induced by ionizing particles, specifically by light ions. To this end, a precise definition of what is considered as *damage* should be cleared. As the nucleus was modeled as a cube filled with water of density 1.06 g/cm^3, the cross sections used in the simulation are those of the water material but properly scaled. This suggests to define the *damage* in terms of the effects of the ionizing particles in water. Since the physical models included in the *GEANT4-DNA* allow to follow electrons until very low energies, of the order of $eV's$, and there are excitations of the water molecule even at such low energies, the early stage of damage that will be characterized here is defined as any event transferring 8 eV or more in the site of a sugar-phosphate atom. Those 8 eV correspond to the minimum excitation energy of liquid water. It has been shown that electrons with energies as low as $\sim$ 0.1 eV can break the DNA [64], but it is common in this field to use threshold damage energies of the order of 10 eV. Electrons with even lower energies can produce damage as well, nevertheless, the MC transport of electrons with very low energies ($\sim$ 1 eV) has been questioned by some authors [65].

The .bin files containing the energy deposition phase-space are superimposed on the squared nucleus model through an algorithm that reads every row of the .bin file and, essentially, invokes the FCA routine to determine whether a damage occurs or not. Such algorithms are called **BB.for, BZ.for** and **AA.for**, according to the DNA model realized. They are intended for producing an output with the relevant information of the damages recorded along the reading of the phase-space files (the .bin files).

Once an event is read within the XX.for application, it is determined if it occurred inside the nucleus. If so, the energy deposited by the event is summed up recursively in order to determine the absorbed dose into the ROI. Also, if the energy transferred E_{tra} is greater than 8 eV, the chromatin index (I_x, I_y) and the coordinates (x_{chro}, y_{chro}) are computed. Then, the FCA routine is called and, if a target is impacted, the routine gets back the indexes $(I_z, I_{block}, I_{bp}, I_{at}, I_{seq})$ as explained in section 5.3. This is called a Single Strand Break (SSB). If there is no target impacted, the code returns 0 for all the indexes

and the XX.for app goes through another event. This is done for all the events in the .bin file with $E_{tra} \geq 8eV$. The number of primary particles simulated for each beam quality is enough to perform the calculation of the SSBs ten times. At the end of each batch, the dose delivered to the nucleus is computed as the total energy deposited divided by the mass of the nucleus.

$$Dose = \frac{Total\ Energy\ Deposited\ within\ the\ Nucleus}{(Nucleus\ Volume) \times (1.06 g/cm^3)} \tag{7.1}$$

The output of the XX.for are a .dat file as well as an onscreen resume. The latter shows the total number of SSBs, SSB+s, DSBs, the site hit probability compared with the theoretical one, the dose absorbed, the number of events for each batch as well as the mean value of every quantity (section 7.1). The .dat file contains the 4D-matrix $ISSB(i,j,k,l)$ that will be explained in chapter 8.

7.1 Strand Break Yields

As mentioned in section 6.1 the probability of damage is defined as the ratio of the total target volume to the nucleus volume. This is called *Site Hit Probability (SHP)*. Then,

$$V_{targets} = 6 \cdot N_{chro} \cdot N_{fib} \cdot \Big(\underbrace{N_{bp}^{nucl} \cdot V_{nucl}^{hel} + N_{bp}^{bind} \cdot V_{bind}^{hel}}_{Volume of the Block} \Big) \tag{7.2}$$

$$V_{nucleus} = \Big(N_{fib} \cdot D_{chro}^2 \Big) N_{chro} \cdot 3 \cdot D_{hel} \tag{7.3}$$

$$SHP = \frac{V_{targets}}{V_{nucleus}} \tag{7.4}$$

As can be seen, the SHP thus defined depends on V_{nucl}^{hel} and V_{bind}^{hel}, the sugar-phosphate group's volume in the nucleosome and linker respectively. As the nucleus was modeled as an homogeneous region filled with water, i.e. the medium is uniform, the total number of Strand Breaks (TSB) per batch divided by the number of events within the nucleus admits an interpretation as the geometrical probability defined by Eq. 7.4. Fig. 16 shows both the theoretical and the radiation-induced SHPs for all the qualities and models used in this work. In all models the two SHPs are close enough as to validate the idea behind considering the probability of damage in terms of geometrical parameters. The differences mainly come from the fact that not all the absorbed doses are exactly the same, also the sizes of the atoms were considered only at first approximation. Other feature that influences the site hit probability is the overlapping among the bps due to the helix bending within the structures. Such overlapping was realized by recording the double counting fraction in every calculation. The results showed that it could be of the order of 9% where one part was already considered in the determination of the effective volume of the target. Furthermore, if one admits the theoretical value to have an uncertainty

associated to the helix volume uncertainty and compares it with the mean probability of damage for protons and alphas particles, the agreement between the SHPs is improved.

Table 8 shows the averaged SHP per particle type and the theoretical SHP uncertainty. It can be noted that, within the uncertainties, the matching is excellent. On the one hand, this means that the probability of damage depends very weakly on the incident energy of the projectile. On the other hand, the results show that the chance for a given particle and energy to cause a strand break is determined by the effective volume of the target, at least in the LET range studied here.

Table 8 – Comparison of the averaged site hit probability obtained in the simulation with the predicted value.

Model	Protons	Alphas	Pred. Value
AA	0.051 ± 0.007	0.050 ± 0.003	0.05300 ± 0.00003
BB	0.040 ± 0.006	0.040 ± 0.002	0.04200 ± 0.00004
BZ	0.031 ± 0.004	0.032 ± 0.002	0.03400 ± 0.00003

Single lesions in the strand are repaired at a relatively high rate compared to the combination of two or more of such lesions. The biological effects start to manifest when SSBs interact to form a type of cluster damage (CL) [66], and this cluster damage is not well-repaired. The simplest form of CL is the Double Strand Break (DSB) which occurs when two SSBs in opposite strands are close enough to interact, breaking the DNA helix. The two ends may rejoin (repair) or not. The latter case could lead to more complex damages in the chromosome such as those discussed in chapter 2. In the XX.for algorithm, is followed the convention that a DSB is recorded if two SSB at opposite strands are separated by no more than ten bps.

The Total Strand Break Yield (TSBY) and Double Strand Break Yield (DSBY) are defined as the total number of SSB[1] and DSB divided by the absorbed dose and the number of base pairs in the nucleus, respectively. The yields are defined that way in order to analyze the influence of the DNA conformation solely, hence, a normalization per unit dose per bp is introduced. The behavior of the yields with the LET are shown in Figs. 17 and 18.

The general behavior of the TSBY is to remain quasi-constant (within the uncertainty) for a given DNA conformation in the LET range covered. As in the SHP case, this yield seems to be determined only by the target volume because the greater the effective volume, the greater the TSBY. The exception comes from the highest LET particles and may be related to the lower rate of interactions of such particles. In the work of Bernal et al. [67], there is a detailed discussion of the invariance of the Total Strand Break Yield for

[1] Even those leading to DSBs.

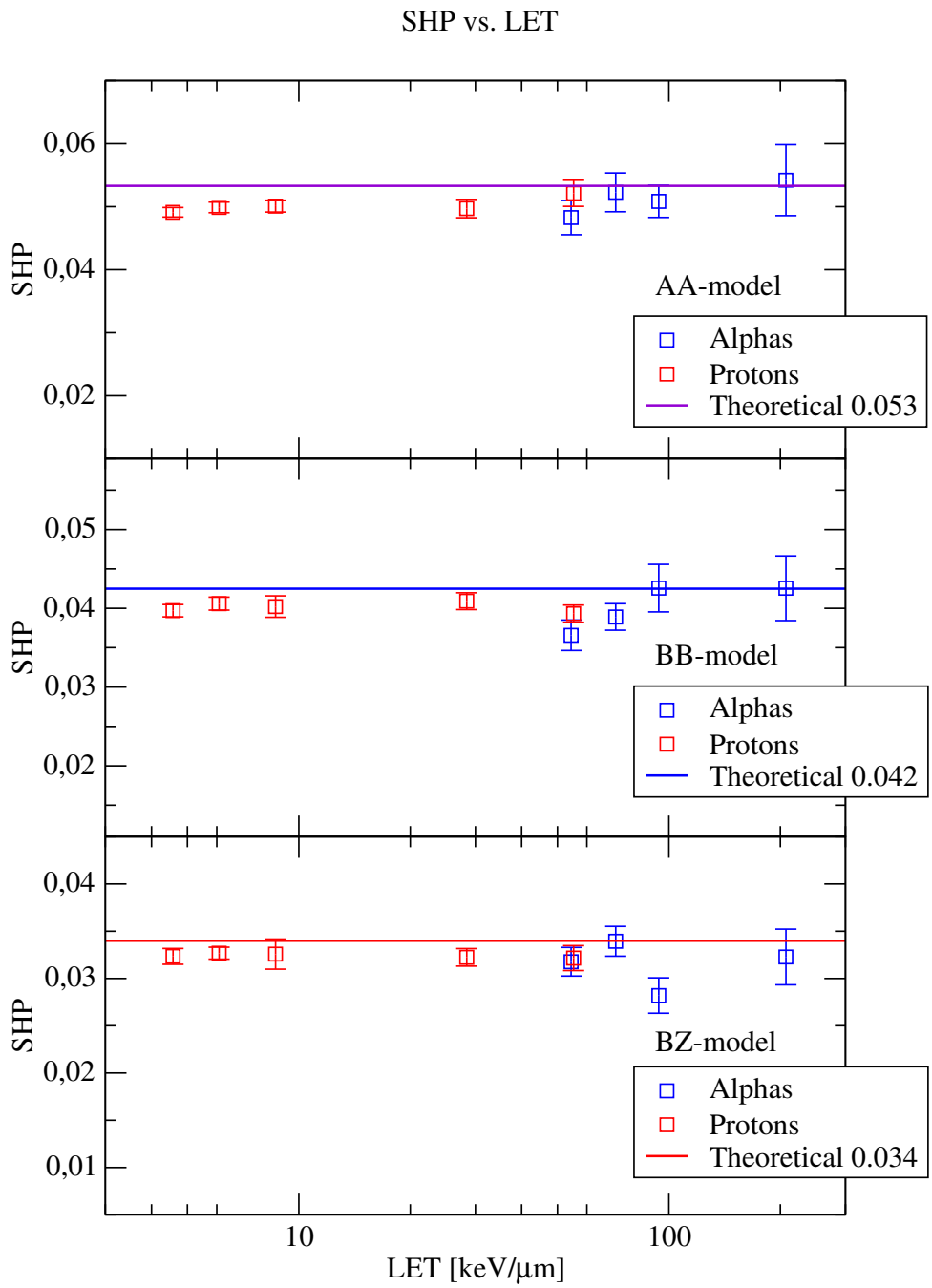


Figure 16 – Site Hit Probability (SHP) against LET. The points represent the ratio of the TSB to the number of events inside the nucleus, averaged over the 30 batches performed for every quality beam. The theoretical value is that of Eq. 7.4 for the V_{hel}^{nucl} and V_{hel}^{bind} according to the model.

the B-DNA. The physics is the same despite the differences in the model. In that work, the TSBY is related to the SHP through the number of events per unit dose induced by the beam. That number of events per dose remains constant as well. To explain why is so, the hypothesis proposed is that the energy depositions only depend on the target structure [68], thus, for a given dose, the number of inelastic events depends on the average energy deposited by the incident particle per event. For the qualities chosen in this work, the energy of the secondary electrons (energy lost by the primaries) produced is nearly the same ($\approx 45eV$). Hence, the number of events they induce thereafter would be nearly the same. Furthermore, the differential cross sections as a function of the secondary electron energy [69, 70] are similar for the light ions used in the simulations. In this context, both primary and secondary particles produce energy deposits very similar ($\sim 15eV$) [71].These considerations give support to the assumption that the number of energy deposition events per Gy is nearly constant.

On the other hand, the DSBY increases with LET as could be expected since this is a measure of the clustering of damage and the highest LETs produce closer SSBs. Also, it seems to depend on the linear density of targets because the DSBY is greater for the AA model in which the helix pitch is just $0.23nm$ ($4.3bp/nm$). The BB- and BZ- models share almost all the parameters so, they have similar DSBY. In the next section more will be said about the behavior of the DSB.

The SHP and yields analyzed here serve to compare the influence of the DNA conformations in the probability of damage, the physical models implemented in the *GEANT4* account for energies low enough to realize damages caused by excitations of the order of 8 eV. The quantities shown have been discussed from a relative point of view, i.e. are not intended for being absolute values.

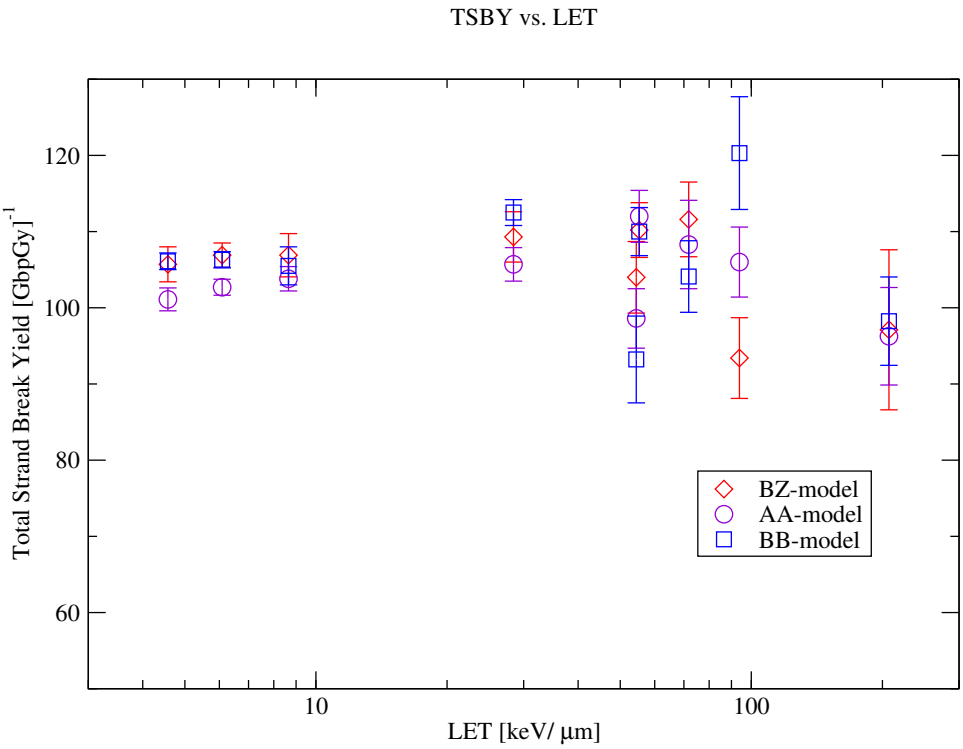


Figure 17 – Averaged Total Strand Break Yield as a function of LET for protons and alpha particles obtained with the simulations and the XX.for algorithm.

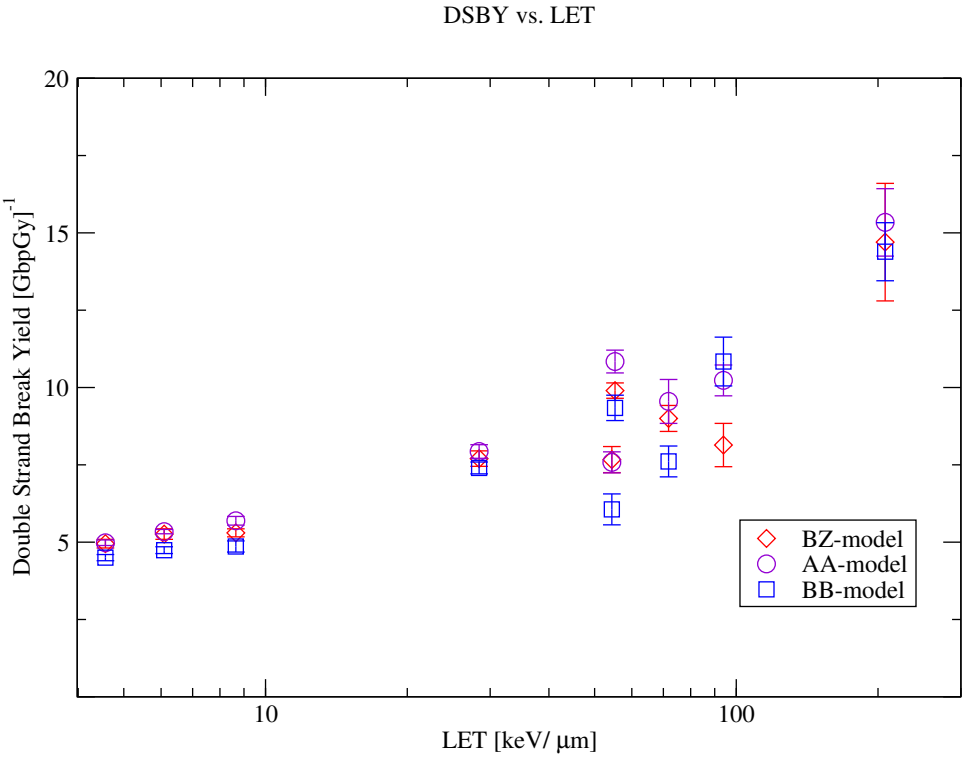


Figure 18 – Averaged Double Strand Break Yield as a function of LET for protons and alpha particles obtained with the simulations and the XX.for algorithm.

8 Numerical Insight into the DRAT

In section 3.1 were explained the main concerns about the Dual Radiation Action Theory and its link to the LQ-model on the basis of a mechanistic approach to the interpretation of both the α and β coefficients. It is commonly accepted to think in such coefficients as first and second order mechanisms of damage production. Solving Eq. 3.8 gave support to this interpretation since the α and β were related with the probability to induce a lethal lesion directly and by the combination of two sub-lesions respectively. The simulations allow to follow each primary particle and its secondaries independently, thus, the origin of every SSB can be tagged according to the type of particle and history that induces it. This means that both the intra-track and inter-track early effects of the ionizing radiation can be investigated via MC simulations. The DRAT's interpretation of the LQ parameters, the geometrical model and the MC simulations were combined with a home-made biophysical model of DNA damage in order to study the number of lethal lesions created in the nucleus.

8.1 Biophysical Model

The basic assumptions of the DRAT are that (1) ionizing radiation is able to produce sub-lesions on the DNA that (2) may interact to become lethal lesions (LL). The solution of Eq. 3.8 determines the number of such LLs induced in the genetic material as a function of the dose in terms of the parameters α and β. Recall that both parameters were defined as probabilities. In the biophysical model proposed here, (1) the DSBs are regarded as sub-lesions that (2) interact in pairs to form a lethal lesion whenever they are separated by no more than 240 nm, which is similar to that obtained by Kellerer and Rossi [34]. This is equivalent to pick up $\epsilon = g(x) = 1$ for $x < 0.24$ μm (see Eq. 3.2). Notice that nothing is assumed about the neutralization of the DSBs after interplay, thus, in this model, all sub-lesions are able to interact as many times as the condition 3.2 be fulfilled. In this way, the lethal lesions are systematically overestimated. Considering that neither neutralization nor reparation of LLs is being taking into account, overestimation of the lethal lesions do not represent a disadvantage of the model. In addition, just the mechanisms of production unlike the absolute number of LL are being investigated.

8.2 Simulations

Using the same irradiation setting of Fig. 11, a new set of simulations were carried out. The beam qualities were protons of 0.5 and 10 MeV and alpha particles of 2 and

10 MeV. 30 batches of one, two, three and so on particles were simulated for each projectile. The maximum number of primaries per batch P_{max} was selected aiming to deliver up to $\sim 15Gy$, therefore, it was different in each case. Once the .bin files were generated, the BB.for algorithm was used to determine the number, location and origin of the SSBs generated in every batch. Each .bin file was superimposed over the nucleus four times rotating the hit coordinates by $0°, 30°, 60°$ and $90°$ (with respect to the y-axis) successively, to mimic homogeneous irradiation conditions. In view of the nucleus was modeled as a squared cylinder with its lateral faces equivalent, it is believe that these four incident angles cover well all the possible irradiation directions, provided that the beam's cross section is greater than the nucleus'. In addition, the upper and bottom view of the nucleus are almost equivalent, being different just in the screw sense of the structures. The output of the BB.for[1] algorithm was stored in a .dat file that contains the 4D matrix $ISSB$.

8.2.1 ISSB matrix

The $ISSB(i, j, k, l)$ 4D-matrix in the .dat output file contains the record of the bps impacted in each batch keeping track of both the location and the origin of every SSB. The information was stored as follows

- **l-dimension:** Number of the batch.
- **k-dimension:** Domain index. The nucleus was divided in 49 domains of 14×14 chromatin fibers each[2]. Nevertheless, such domains, initially intended to be chromosomes, were not use in the analysis because *total* DSB was ultimately counted. Still, the indexes were kept once they proved to be useful in terms of post-processing time.
- **j-dimension:**

  I_{seq} The index of the bp impacted not just in the chromatin fiber coordinates as in Eqs. 5.29 and 5.30 but in the whole nucleus context.

  I_{str} Impacted Strand Index: 1 or 2 depending on which strand of the double helix was damaged.

  n_{hist} Number (within the batch) of the primary particle that originated the event.

  x_{hit} In Nucleus-coordinates.

  y_{hit} In Nucleus-coordinates.

  z_{hit} In Nucleus-coordinates.

[1] Notice that only the BB-model was studied because is the canonical form of DNA in the human cells.
[2] In this part of the work, the total number of fiber was changed to 98×98 to enable the domain division.

I_{proj} Type of projectile: 1 for Ions, 2 for electrons.

- **i-dimension:** Sequential counter of SSBs.

For a given batch and domain, the $ISSB(i, j, k_{fixed}, l_{fixed})$ displays a 2D-matrix of dimension $I_{max} \times 7$, where I_{max} is the number of SSBs scored in the given domain and batch[3]. Two special cells in the $ISSB$ were set apart to store the mean dose absorbed (per batch) in the nucleus and its uncertainty.

8.3 BreakAnalyzer Algorithm

A second code was written in FORTRAN in order to process the .dat file. The aim of the code is to analyze the first and second order mechanisms of the radiation-induce damage, hence the name *BreakAnalyzer*. The algorithm was designed to perform the analysis in two steps. In a first stage, the code reads the $ISSB$ matrix and compute the number of DSBs produced in each batch, for a given quality and incident angle. A DSB is scored when two SSBs at opposite strands are separated by at most 10 bps. Also, the origin and location of the DSB is determined. The information is stored internally in a 4D-matrix called $IDSB$ to subsequent processing (see Figure 19).

The position of a DSB was defined as the midpoint between the two SSBs that induced it. Each SSB has a n_{hist} associated, that is, a label to the primary particle that produced such event, hence, two $n_{hist}{}'s$ were tagged to the DSB. Moreover, the two breaks could be originated by either electrons, ions or an electron and an ion respectively, thus, the type of projectile was tagged to the DSB as well, multiplying the corresponding indexes I_{proj}.

The $IDSB$ matrix then was read in order to compute the total number of LLs formed in the nucleus, according to the biophysical model proposed. To this end, the distance between a given DSB and all the DSBs in the surroundings was calculated. The *surrounding DSBs* mean the DSBs scored in the neighbor chromosomes of the domain that contains the sub-lesion analyzed. If two DSBs were separated by $240nm$ or less, a LL was counted. Also, the 4 $n_{hist}{}'s$ were compared to determined if the lethal lesion was created by the same track (intra-track effect) or by the combination of two tracks (inter-track effect). Finally, the number of DSBs and LL induced are averaged the number of batches for each number of primaries (absorbed dose).

[3] Actually, the dimension of the i-th component of $ISSB$ was set to 20000 to avoid dynamic allocation bugs. However, the I_{max} can be easily extracted as a parameter so, in practice, the 2D matrix commented is what is read at the post-processing stage.

```
DO l=1,30                !Batch counter
  dsb=0                  !DSB counter
  DO k=1,49              !Chromosome/Domain index
    i=1                  !SSB counter within the
    IF(IMAX(k,l).GT.1)
      DO WHILE (i.LT.IMAX(k,l))
        hit1=ISSB(i,1,k,l)
        sthit1=ISSB(i,2,k,l)
        hit2=ISSB(i+1,1,k,l)
        sthit2=ISSB(i+1,2,k,l)
        IF(abs(hit1-hit2).LE.10.AND(sthit1.NE.sthit2)) THEN
          dsb=dsb+1

          !Location of the DSB (Iseq)
          IDSB(dsb,1,k,l)=(hit1+hit2)/2

          !Location of the DSB (x,y,z). Midpoint of the SSBs
          IDSB(dsb,2,k,l)=(ISSB(i,4,k,l)+ISSB(i+1,4,k,l))/2
          IDSB(dsb,3,k,l)=(ISSB(i,5,k,l)+ISSB(i+1,5,k,l))/2
          IDSB(dsb,4,k,l)=(ISSB(i,6,k,l)+ISSB(i+1,6,k,l))/2

          !Corresponding nhist of each SSB
          IDSB(dsb,5,k,l)=ISSB(i,3,k,l)
          IDSB(dsb,6,k,l)=ISSB(i+1,3,k,l)

          !Type of DSB (Electronic, Ionic, Mixed).
          !Multiplication of the projectile indexes
          IDSB(dsb,7,k,l)=ISSB(i,7,k,l)*ISSB(i+1,7,k,l)
          i=i+1          !To avoid SSB double counting
        END IF
        i=i+1            !Next SSB
      END DO             !End of SSBs
    END IF               !Next Chromosome
  END DO                 !End of Chromosomes, next batch
END DO                   !End of Batch
```

Figure 19 – Simplification of the first part of the *BreakAnalyzer.for* code. The number of sub-lesions (DSBs) is computed from the input.

8.3.1 Example

To illustrate the method just described, take the case of protons of 10 MeV. For this quality, 58500 particles were simulated. This value comes from 30 batches of 6 up to $P_{max} = 150$ histories each (in multiples of 6). All the particles were transported independently and stored in a single binary file called *p10.bin*. Before using the BB.for algorithm, the p10.bin file was split into 25 sub-files: p10**180**.bin, p10**360**.bin,..., p10**4500**.bin to perform the following analysis with a fixed mean dose per batch. P_{max} was estimated from the observation that 6 proton of $10MeV$ delivered $\sim 0.6Gy$ within the nucleus[4].

After the splitting, the BB.for was used to generate the 25 $ISSB$ 4D matrices corresponding to each sub-file. The outputs: p10**30**.dat, p10**60**.dat,..., p10**750**.dat obtained at this point, corresponded to normal irradiation i.e. at 0°. Next, the BB.for code was modified in order to rotate the hit coordinates by 30° and another set of .dat files was accomplished. The process was repeated for the 60° and 90° cases.

The next phase of the investigation consisted in to read the 25 .dat files of a given angle, with the BreakAnalyzer.for algorithm. As explained in the previous section, the BreakAnalyzer returns a table with the average number of DSBs and lethal lesions induced in the nucleus. Also, in the table is discriminated the number of LL created by intra-track and inter-track effects (columns N_α and N_β respectively). Table 9 shows the output of the 60° case. The data was printed in a .txt file called p10**60**.txt. The bold number stands for the corresponding angle of irradiation. Furthermore, in order to visualize the spatial distribution of the DSBs, another file was generated containing the position and type of projectile that produced such sub-lesions. This is the p10DSBmap.txt file. Such map was printed only for normal irradiation.

After the p10**angle**.txt tables were generated, the values of each row were averaged to obtain the homogeneous irradiation values. The uncertainties were calculated as in the method of propagation of errors, that is, taking the squared root of the sum of the squares. The final analysis of the first and second order mechanisms of damage production were made based on the averaged values (homogeneous irradiation). In the next subsection will be shown the results for all the particles and energies studied in this work.

[4] A specific simulation was done for this propose. In this case, the batches were chosen to contain multiples of $180(= 6 * 30)$ particles in order to get a better statistic.

Table 9 – Example of a typical BreakAnalyzer output. All uncertainties represent one standard deviation of the mean.

Dose(Gy)	unc.(Gy)	DSB	unc.	LL	unc.	N_α	unc.	N_β	unc.
0.69	0.03	10.23	0.59	1.90	0.33	1.67	0.29	0.23	0.12
1.43	0.05	23.23	1.12	5.57	0.62	4.10	0.45	1.47	0.32
2.21	0.06	34.60	1.19	8.83	0.92	6.30	0.54	2.53	0.62
2.85	0.06	43.77	1.51	11.63	1.03	7.90	0.87	3.73	0.36
3.63	0.07	56.80	1.89	16.40	1.21	9.93	0.82	6.47	0.68
4.34	0.06	67.30	1.86	20.80	1.45	12.50	0.91	8.30	0.87
5.06	0.07	78.40	1.73	24.90	1.82	12.63	0.98	12.27	1.17
5.75	0.06	87.60	2.34	27.57	1.81	12.37	0.94	15.20	1.23
6.43	0.11	99.20	2.62	35.50	2.10	17.57	1.12	17.93	1.35
7.27	0.10	112.30	2.31	43.07	2.18	18.57	1.05	24.50	1.54
7.79	0.10	117.77	1.89	45.17	2.78	18.43	1.03	26.73	2.02
8.60	0.10	133.77	2.62	59.70	2.61	24.70	1.38	35.00	1.67
9.19	0.10	144.30	2.79	64.53	2.68	24.20	1.33	40.33	1.99
10.19	0.10	162.00	2.44	76.67	3.50	28.20	1.16	48.47	2.81
10.91	0.11	167.90	3.25	83.37	4.59	27.70	1.67	55.67	3.59
11.45	0.08	178.60	2.89	89.03	3.26	29.00	1.14	60.03	2.65
12.10	0.13	186.13	3.72	99.13	5.08	30.80	1.72	68.33	3.93
12.76	0.13	192.33	3.57	99.77	4.13	32.10	1.80	67.67	3.03
13.61	0.12	212.53	4.06	122.43	5.40	36.20	1.84	86.23	4.18
14.09	0.11	219.87	3.28	128.10	4.97	35.83	1.45	92.27	4.10
14.92	0.12	230.97	3.15	140.80	5.33	36.43	1.35	104.37	4.75
15.81	0.15	246.90	3.68	153.93	5.52	40.70	1.80	113.23	4.36
16.21	0.13	251.67	3.31	163.80	5.63	41.67	1.66	122.13	4.88
17.04	0.14	260.53	4.02	170.43	7.36	41.17	2.21	129.27	6.53
17.85	0.15	276.93	5.12	193.60	9.09	45.20	1.73	148.40	7.88

8.4 First and Second Order Mechanisms Analysis

The same steps of the previous example were followed for the alpha particles of 2 and $10MeV$ as well as for the protons of $0.5MeV$. The mean number of lethal lesions due to intra- and inter-track effects, for every quality, are shown in Fig. 20. As can be seen in the plot, the number of LLs that were originated by the same track (blue points) increase linearly with the dose in all cases. The function $N_\alpha = \alpha D$ was used to fit the data series. The results are also shown in the plot. The α coefficient represents the number of lethal lesions induced per unit dose but also can be seen as a probability of damage production

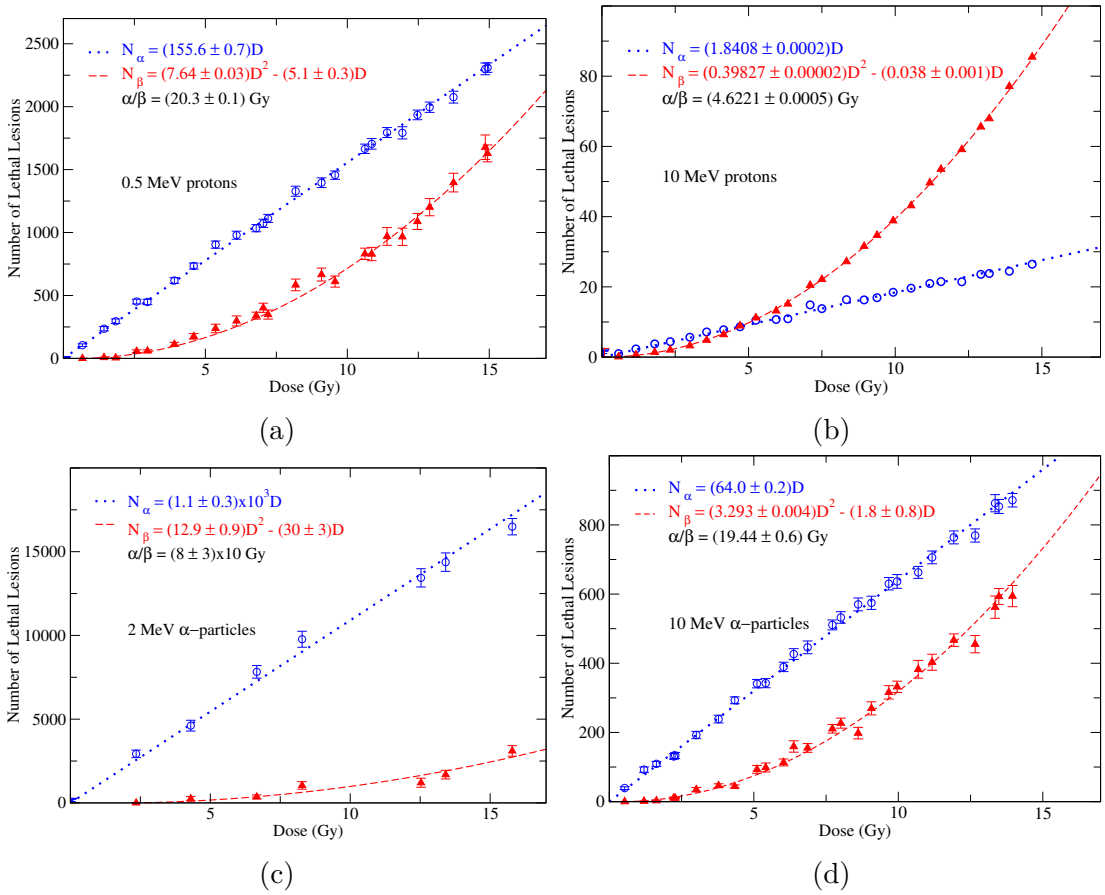


Figure 20 – Number of lethal lesions per cell induced by protons and α particles through both first (α) and second (β) order mechanisms. Lines represent the fitting of the results according to the shown formulas.

via first order mechanism as will be discussed soon.

The red points represent the number of lethal lesions produced by two different tracks i.e. by a second order mechanism. It is observed that the behavior of this fraction of the damage is quadratic with the dose. There is a dose below which there is no chance for the production of a LL by two independent tracks. This is just the one-track dose D_0. That is, the dose absorbed into the region of interest due to only one track. Furthermore, the number of LL must be zero at D=0 as well. Thus, the following formula is proposed for fitting the quadratic curve,

$$N_\beta = \beta D(D - D_0) \tag{8.1}$$

where β would represent the probability per unit squared dose for the induction of a LL by two independent tracks. D_0 would be the expected value of the specific energy deposited by a single track into the ROI. In other words, as this is a second order mechanism involving two tracks, no LL can be produced if the dose is similar to that deposited by a single track. According to the definition of D_0, two fitting approaches were carried out. In the first method, the value of D_0 was taken directly from the simulations and was

performed a one-parameter weighted regression to fit the data. In the second approach, a two-parametric weighted regression method was done. The variance of the data was used as the weighting factor in both cases. Table 10 shows a comparison of the parameters β and D_0 obtained by each technique.

Table 10 – One-track dose and quadratic parameter as determined by one- and two-parameter fitting procedures. D_0 is directly extracted from MC simulations in the one-parameter fitting approach.

Particle	One-parameter fitting		Two-parameter fitting	
	D_0 (Gy)	β (Gy^{-2})	D_0 (Gy)	β (Gy^{-2})
0.5 MeV protons	0.67 ± 0.04	7.64 ± 0.03	0.67 ± 0.01	7.6 ± 0.1
10 MeV protons	0.096 ± 0.002	0.39744 ± 0.00001	0.08 ± 0.09	0.397 ± 0.005
2.0 MeV α part.	2.4 ± 0.1	12.9 ± 0.9	2.3 ± 0.2	12.9 ± 0.9
10 MeV α part.	0.53 ± 0.04	3.293 ± 0.004	0.53 ± 0.01	3.36 ± 0.05

All results show very good consistency since the intervals for β and D_0 found by one or another way, overlap correspondingly. This means that the reasoning behind Eq. 8.1 has physical plausibility and suggests that the linear term in a common survival curve would also depend on the β parameter. Hence, the total number of LL follows the expression,

$$N_{LL} = (\alpha - \beta D_0)D + \beta D^2 \tag{8.2}$$

Figures 20a and 20b shows the number of LL as a function of dose for 0.5 and 10 MeV protons, respectively. The 0.5 MeV proton α/β ratio is greater than that for 10 MeV protons. Since there are first and second order mechanisms involved in the formation of this ratio, both of them have to be analyzed. On the one hand, the α parameter increases with LET as the spatial density of energy deposition events increases and so the probability for the induction of two DSB close enough to interact. On the other hand, the β parameter also increases with LET, as shown in these plots, but at a slower rate than α. Of course, as there are more DSB with the increment of LET, there is a tendency for β to increase. However, the number of primary tracks per unit area, or track planar fluence (Φ), needed to deliver a given dose decreases with LET. This fluence can be determined as follows

$$\Phi = \frac{D}{S/\rho} \tag{8.3}$$

where D is the dose and S/ρ, the mass-stopping power. The fluence decreases when LET (or S/ρ) increases. This effect could be the reason for the increase of β with

LET slower than for α, resulting in an increment of the α/β ratio. The increment of β was also reported by Jones in a work with neutron beams [72]. The data on ion beams compiled by Friedrich et al. seem to be too spread for a final conclusion [73].

Figures 20c and 20d displays the number of LL induced by both first and second order mechanisms for 2 and 10 MeV α particles, respectively. The fitting procedure explained for protons was followed in this case as well. Again, the higher the LET of the incident particles, the greater the α/β ratio. It is interesting to point out that 0.5 MeV protons and 10 MeV α particles have similar α/β ratios. At the center of the ROI, 10 MeV α particles and 0.5 MeV protons have a LET of about 58 keV/μm and 67 keV/μm, respectively. Notice that these protons have LET about 15 % higher than that of the α particles. This means that α particles would have greater α/β ratio, and so RBE, than protons with the same LET. This result is consistent with findings reported in many experimental works (see Fig. 4 in Ref. [73]). However, this is not the case for the DSB yield, where light particles show greater yields than heavy ones at the same LET. For the α/β ratio case, it should be accounted for that there is an inter-track effect not existent in a simple determination of DSB yields (see the discussion in the paragraph just above).

Conclusions

The most common approach in computational Radiobiology is to define targets and to count the number of these targets impacted by the ionizing particle. In this way, the DNA strand break yields are determined. Under this approach, site-hit probability is determined by the volume targets, that is, the volume of the sugar-phosphate groups. It was found in this work that this volume varies within ± 2% among the A-, B-, and Z-DNA configurations.

The total strand break yield is practically determined by the site-hit probability and depends weakly on the incident radiation quality. However, the double strand break yield strongly depends on the radiation quality since it is directly related to the capacity of the radiation to form energy deposition clusters. In addition, the DSB yield also depends on the DNA conformation since this yield can be seen as the result of a superposition of energy deposition clusters onto the DNA structure. Thus, the interplay between the radiation clustering capacity and DNA spatial packing influences this yield.

The numerical model developed in this work shows that there are first and second order mechanisms involved in the induction of cellular lethal lesions by ionizing particles. This hypothesis was proposed by Lea and Catchside [36] in 1942 and later verified experimentally and explained by mathematical modelling. This model led to the development of the Linear-Quadratic radiobiological model, characterized by the so-called α and β parameters. It has to be pointed out that the LQ models parameters are obtained by a fitting prodcedure so they are different from those derived from our model (physical parameters). It was found the LQ α parameter depends both on the physical α and β parameters determined here. This is due to the one-track dose effect explained before.

This model can be used as a platform for introducing other radiobiological effects such as the DNA damage repair, DSB diffusion and others. These effects would lead to the introduction of the parameter time into the model so it could be applied in continuous irradiation situations as those found in low dose rate brachytherapy of cancer. In addition, the possibility to study the lethal lesion induction process as the interaction of clusters of DSBs instead of two DSBs may be explored.

It is expected that this biophysical model could be applied in clinical situations in order to estimate the relative biological effectiveness (RBE) of several radiation qualities, mainly those used in hadrontherapy and low energy photon irradiations. These are situations where RBE is still on debate.

Bibliography

1 DESOUKY, O.; DING, N.; ZHOU, G. Targeted and non-targeted effects of ionizing radiation. *Journal of Radiation Research and Applied Sciences*, v. 8, n. 2, p. 247–254, April 2015. Cited 2 times in pages 9 and 17.

2 POMPLUN, E. A new dna target model for track structure calculations and its first application to i-125 auger electrons. *Int. J. Radiat. Biol.*, v. 59, p. 625–642, 1991. Cited 4 times in pages 11, 39, 40, and 59.

3 HO, P. S.; CARTER, M. *DNA Structure: Alphabet Soup for the Cellular Soul, DNA Replication-Current Advances*. InTech, 2011. Available on: <http://www.intechopen-.com/books/dna-replication-current-advances/dna-structure-alphabet-soup-for-the-cellular-soul>. Cited in page 15.

4 KUZIN, A. M. On the role of dna in the radiation damage of the cell. *Int. J. Radiat. Biol.*, v. 6, p. 201–209, 1963. Cited in page 16.

5 KADHIM, M. et al. Transmission of chromosomal instability after plutonium alpha-particle irradiation. *Nature*, v. 355, p. 738–740, 1992. Cited in page 16.

6 NAGASAWA, H.; LITTLE, J. Induction of sister chromatid exchanges by extremely low doses of alpha-particles. *Cancer Res.*, v. 52, p. 6394â6396, 1992. Cited in page 16.

7 JACOB, P. et al. Possible expressions of radiation-induced genomic instability, bystander effects or low-dose hyper-sensitivity in cancer epidemiology. *Mutat. Res.*, v. 687, p. 34–39, 2010. Cited in page 16.

8 BAVERSTOCK, K. Why do we need a new paradigm in radiobiology? *Mutat. Res.*, v. 687, p. 3–6, 2010. Cited in page 16.

9 UNSCEAR2008. *SOURCES AND EFFECTS OF IONIZING RADIATION*. New York, 2010. Cited in page 16.

10 IAEA. *RADIATION BIOLOGY: A HANDBOOK FOR TEACHERS AND STUDENTS*. Austria: IAEA Publishing Section, 2010. (42, I). Cited 2 times in pages 19 and 27.

11 IAEA. *RADIATION ONCOLOGY PHYSICS: A HANDBOOK FOR TEACHERS AND STUDENTS*. Vienna: IAEA Publishing Section, 2005. Cited in page 20.

12 FRIEDRICH, T. et al. Systematic analysis of rbe and related quantities using a database of cell survival experiments with ion beam irradiation. *J. Radiat. Res.*, v. 54, p. 494–514, 2013. Cited in page 20.

13 GRIFFITHS, A. J. et al. *An Introduction to Genetic Analysis*. 7th. ed. New York: W. H. Freeman, 2000. Cited in page 22.

14 WATSON, J. D.; CRICK, F. H. C. Molecular structure of nucleic acids. *Nature*, v. 171, p. 737–738, 1953. Cited 2 times in pages 22 and 23.

15 DICKERSON, R. E. et al. The anatomy of a-, b- and z-dna. *Science*, v. 216, p. 475–485, 1982. Cited in page 23.

16 VISCHER, E.; CHARGAFF, E. The separation and quantitative estimation of purines and pyrimidines in minute amounts. *J. Biol. Chem*, v. 176, p. 703–714, 1948. Cited in page 23.

17 CHARGAFF, E. Chemical specificity of nucleic acids and mechanism of their enzymatic degradation. *Experientia*, v. 6, p. 201â209, 1950. Cited in page 23.

18 KNEE, K. M. et al. Spectroscopic and molecular dynamics evidence for a sequential mechanism for the a-to-b transition in dna. *Biophys. J.*, v. 95, n. 7, p. 257–272, 2008. Cited in page 24.

19 WANG, A. H. et al. Molecular structure of a left-handed double helical dna fragment at atomic resolution. *Nature*, v. 282, p. 680–686, 1979. Cited in page 24.

20 THAMANN, T. J. et al. High salt form of poly(dg-dc)-poly(dg-dc) is left-handed z-dna: Raman spectra of crystals and solutions. *Nucleic Acids Res.*, v. 9, p. 5443–5457, 1981. Cited in page 24.

21 HA, S. C. et al. Crystal structure of a junction between b-dna and z-dna reveals two extruded bases. *Nature*, v. 479, n. 10, p. 1183â1186, 2005. Cited in page 24.

22 RICH, A.; ZHANG, S. Z-dna: the long road to biological function. *Nat. Rev. Genet.*, v. 4, p. 566–572, July 2003. Cited in page 24.

23 WARD, J. F. Radiation mutagenesis: The initial dna lesions responsible. *Radiat. Res.*, v. 142, n. 3, p. 362–368, June 1995. Cited in page 25.

24 GOODHEAD, D. T. Initial events in the cellular effects of ionizing radiations: clustered damage in dna. *Int. J. Radiat. Biol.*, v. 65, p. 7–17, 1994. Cited in page 25.

25 SCHIPLER, A.; ILIAKIS, G. Dna double-strand-break complexity levels and their possible contributions to the probability for error-prone processing and repair pathway choice. *Nucleic Acids Res.*, v. 41, p. 7589â7605, 2013. Cited in page 25.

26 FRANKENBERG-SCHWAGER, M. *The Early Effects of Radiation on DNA*. Berlin: Springer-Verlag, 1991. (H, v. 54). Induction, repair and biological relevance of radiation-induced DNA lesions in eukaryotic cells. Cited in page 25.

27 DURANTE, M. et al. From dna damage to chromosome aberrations: Joining the break. *Mutat. Res.*, v. 756, p. 5–13, August 2013. Cited in page 25.

28 NAGASAWA, H. et al. Some unsolved problems and unresolved issues in radiation cytogenetics: A review and new data on roles of homologous recombination and non-homologous end joining. *Mutation Research Genetic Toxicology and Environment Mutagenesis*, v. 701, n. 1, p. 12–22, August 2010. Cited in page 25.

29 VERKAIK, N. S. et al. Different types of v(d)j recombination and end-joining defects in dna double-strand break repair mutant mammalian cells. *Eurpean Journal of Immunology*, v. 32, n. 3, p. 701–709, 2002. Cited in page 25.

30 YOU, Z.; BAILIS, J. M. Dna damage and decisions: Ctip coordinates dna repair and cell cycle checkpoints. *Trends Cell Biol.*, v. 20, n. 7, p. 402–409, May 2010. Cited in page 26.

31 BRENNER, D. J. Point: The linear-quadratic model is an appropriate methodology for determining iso-effective doses at large doses per fraction. *Semin. Radiat. Oncol.*, v. 18, n. 4, p. 234–239, 2008. Cited in page 28.

32 FOWLER, J. F. Linear quadratics is alive and well: In regard to park et al. (int j radiat oncol biol phys 2008;70:847–852). *International Journal of Radiation Oncology*Biology*Physics*, v. 72, n. 3, p. 957–, 11 2008. Available on: <http://www.sciencedirect.com/science/article/pii/S0360301608029738>. Cited in page 28.

33 STAVREV, P. et al. On differences in radiosensitivity estimation: Tcp experiments versus survival curves. a theoretical study. *Physics in Medicine and Biology*, v. 60, n. 15, p. N293, 2015. Available on: <http://stacks.iop.org/0031-9155/60/i=15/a=N293>. Cited in page 28.

34 KELLERER, A. M.; ROSSI, H. H. The action of dual radiation action. *Current Topics in Radiation Rsearch Quarterly*, v. 8, p. 85–158, 1972. Cited 2 times in pages 28 and 69.

35 KELLERER, A. M.; ROSSI, H. H. A generalized formulation of dual radiation action. *Radiation Research*, Radiation Research Society, v. 75, n. 3, p. 471–488, 1978. ISSN 00337587, 19385404. Available on: <http://www.jstor.org/stable/3574835>. Cited 2 times in pages 28 and 29.

36 LEA, D. E.; CATCHESIDE, D. G. The mechanism of the induction by radiation of chromosome aberrations intradescantia. *Journal of Genetics*, v. 44, n. 2, p. 216–245, 1942. ISSN 0022-1333. Available on: <http://dx.doi.org/10.1007/BF02982830>. Cited 2 times in pages 29 and 78.

37 NOMIYA, T. Discussions on target theory: Past and present. *Journal of Radiation Research*, v. 54, n. 6, p. 1161–1163, November 2013. Cited in page 29.

38 MARIANA, G.; MARCO, C. Mechanistic formulation of a lineal-quadratic-lineal (lql) model: Split-dose experiments and exponentially decaying sources. *Medical Physics*, v. 37, n. 8, p. 4173–4181, August 2010. Cited 2 times in pages 30 and 95.

39 FUNDAMENTALS of the Monte Carlo method for neutral and charged particle transport. [S.l.]: The University of Michigan, 2000. Cited in page 32.

40 JAMES, F. A review of pseudorandom number generators. *Comput. Phys. Commun.*, v. 60, p. 329–344, 1990. Cited in page 33.

41 MARSAGLIA, G. Random numbers fall mainly in the planes. *Natural Academy of Science*, v. 61, p. 25–28, 1968. Cited in page 33.

42 BERGER, M. J. *Monte Carlo calculation of the penetration and diffusion of fast charged particles*. [S.l.]: Academic Press: New York, 1963. Cited in page 38.

43 SALVAT, F.; FERNANDEZ-VAREA, J. M.; SEMPAU, J. *A Code System for Monte Carlo Simulation of Electron and Photon Transport*. Issy-les-Moulineaux, France., 2008. Cited in page 38.

44 KAWRAKOW, I.; ROGERS, D. *The EGSnrc code system: Monte Carlo simulation of electron and photon transport*. Otawa, Canada, 2000. Cited in page 38.

45 TEAM., X.-. M. C. *MCNP a general Monte Carlo n-particle transport code, version 5, Volume I: Overview and Theory, 3 vols.* [S.l.], 2003. Cited in page 38.

46 AGOSTINELLI, S. et al. Geant4-a simulation toolkit. *Nuclear instruments and methods in physics research section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, v. 506, n. 3, p. 250–303, 2003. Cited in page 38.

47 INCERTI, S. et al. The geant4-dna project. *Int. J. Model. Simul. Sci. Comput.*, v. 1, n. 2, p. 157–178, 2010. Cited in page 38.

48 FRIEDLAND, W. et al. Track structures, dna targets and radiation effects in the biophysical monte carlo simulation code partrac. *Mutat Res Fundam Mol Mech Mutagen*, p. 28–40, 01 2011. Available on: <http://pubget.com/paper/21281649>. Cited 2 times in pages 38 and 59.

49 CHARLTON, D. E.; NIKJOO, H.; HUMM, J. L. Calculation of initial yields of single- and double-strand breaks in cell nuclei from electrons, protons and alpha particles. *International Journal of Radiation Biology*, Informa Clin Med, v. 56, n. 1, p. 1–19, 2012/09/11 1989. Available on: <http://dx.doi.org/10.1080/09553008914551141>. Cited in page 39.

50 FRIEDLAND, W. et al. Simulation of dna fragment distributions after irradiation with photons. *Radiat. Environ. Biophys.*, v. 38, p. 39–47, 1999. Cited 2 times in pages 39 and 60.

51 OTTOLENGHI, A.; MERZAGORA, M.; PARETZKE, H. G. Dna complex lesions induced by proton and alpha particles: track structure characterisitcs determining linear energy transfer and particle type dependence. *Radiat. Environ. Biophys.*, v. 36, p. 97–103, 1997. Cited in page 39.

52 MOISEENKO, V. V. et al. The cellular environment in computer simulations of radiation-induced damage to dna. *Radiation and Environmental Biophysics*, Springer Berlin / Heidelberg, v. 37, p. 167–172, 1998. ISSN 0301-634X. 10.1007/s004110050112. Available on: <http://dx.doi.org/10.1007/s004110050112>. Cited in page 39.

53 BRIDEN, P. E.; HOLT, P. D.; SIMMONS, J. A. The tracks structures of ionizing particles and their application to radiation biophysics. *Radiat. Environ. Biophys.*, v. 38, p. 175–184, 1999. Cited in page 39.

54 PONOMAREV, A. L. et al. A polymer, random walk model for the size-distribution of large dna fragments after high linear energy transfer radiation. *Radiation and Environmental Biophysics*, Springer Berlin / Heidelberg, v. 39, p. 111–120, 2000. ISSN 0301-634X. 10.1007/s004119900040. Available on: <http://dx.doi.org/10.1007-/s004119900040>. Cited in page 39.

55 GLACTONE. *The GLACTONE project.* Available on: <http://chemistry.gsu.edu-/Glactone/PDB/DNA_RNA/dna.html>. Cited in page 40.

56 BERNAL, M. A. et al. An atomistic geometrical model of the b-dna configuration for dna-radiation interaction simulations. *Computer Physics Communications*, v. 184, n. 12, p. 2840–2847, 2013. Available on: <http://www.sciencedirect.com/science/article-/pii/S0010465513002439>. Cited 2 times in pages 42 and 50.

57 RICHMOND, T. J.; DAVEY, C. A. The structure of dna in the nucleosome core. *Nature*, v. 423, p. 145–150, 2003. Cited in page 42.

58 FRIEDLAND, W. et al. Simulation of dna damage after proton irradiation. *Radiation Research*, The Radiation Research Society, v. 159, n. 3, p. 401–410, 2012/02/14 2003. Available on: <http://www.rrjournal.org/doi/abs/10.1667/0033-7587%282003%29159%5B0401%3ASODDAP%5D2.0.CO%3B2>. Cited in page 49.

59 FRIEDLAND, W. et al. Monte carlo simulation of the production of short dna fragments by low-linear energy transfer radiation usin high-order dna molecules. *Radiat. Res.*, v. 150, p. 170–182, 1998. Cited in page 60.

60 CHAUVIE, S. et al. Geant4 physics processes for microdosimetry simulation: Design foundation and implementation of the first set of models. *Nuclear Science, IEEE Transactions on*, v. 54, n. 6, p. 2619–2628, Dec. 2007. Available on: <10.1109/TNS.2007-.910425>. Cited in page 63.

61 WRIGHT, D. *GEANT4 Physics Reference Manual*. 9.1. ed. [S.l.], 2007. Cited in page 63.

62 INCERTI, S. et al. Comparison of geant4 very low energy cross section models with experimental data in water. *Medical Physics*, v. 37, n. 9, p. 4692–4708, 2010. Cited in page 63.

63 BERNAL, M. A. et al. Track structure modeling in liquid water: A review of the geant4-dna very low energy extension of the geant4 monte carlo simulation toolkit. *Physica Medica: European Journal of Medical Physics*, Elsevier, v. 31, n. 8, p. 861–874, 2016/02/19 2015. Available on: <http://dx.doi.org/10.1016/j.ejmp.2015.10.087>. Cited in page 63.

64 PANAJOTOVIC, R. et al. Effective cross sections for production of single-strand breaks in plasmid dna by 0.1 to 4.7 ev electrons. *Radiation Research*, Radiation Research Society, v. 165, n. 4, p. 452–459, 09 2006. Available on: <http://dx.doi.org/10.1667-/RR3521.1>. Cited in page 63.

65 THOMSON, R. M.; KAWRAKOW, I. On the monte carlo simulation of electron transport in the sub-1 kev energy range. *Medical Physics*, v. 38, n. 8, p. 4531, 2011. Cited in page 63.

66 HADA, M.; GEORGAKILAS, A. G. Formation of clustered dna damage after high-let irradiation: a review. *J. Radiats Res.*, v. 49, n. 3, p. 203–210, 2008. Cited in page 65.

67 BERNAL, M. A. et al. The invariance of the total direct dna strand break yield. *Medical Physics*, AAPM, v. 38, n. 7, p. 4147–4153, 07 2011. Available on: <http://link.aip.org/link/?MPH/38/4147/1>. Cited in page 65.

68 ICRU36. *Microdosimetry*. Bethesda, MD, 1983. Cited in page 67.

69 WILSON, W. E. et al. Differential cross sections for ionization of methane, ammonia and water vapour by high velocity ions. *J Chem Phys*, v. 80, p. 5601–8, 1984. Cited in page 67.

70 BERNAL, M. A.; LIENDO, J. A. Inelastic-collision cross sections for the interactions of totally stripped h, he and c ions with liquid water. *Nucl. Instrum. Methods Phys. Res. B.*, v. 262, n. 1, p. 1–6, 2007. Cited in page 67.

71 BERNAL, M. A. Evaluation of the mean energy deposit during the impact of charged particles on liquid water. *Physics in Medicine and Biology*, v. 57, n. 7, p. 1745–1757, 2012. Cited in page 67.

72 JONES, B. The apparent increase in the beta-parameter of the linear quadratic model with increased linear energy transfer during fast neutron irradiation. *The British Journal of Radiology*, The British Institute of Radiology, v. 83, n. 989, p. 433–436, 2015/11/15 2010. Available on: <http://dx.doi.org/10.1259/bjr/68792966>. Cited in page 77.

73 FRIEDRICH, T. et al. Systematic analysis of rbe and related quantities using a database of cell survival experiments with ion beam irradiation. *Journal of Radiation Research*, v. 54, n. 3, p. 494–514, 12 2012. Available on: <http://jrr-.oxfordjournals.org/content/early/2012/12/23/jrr.rrs114.abstract>. Cited in page 77.

APPENDIX A – Creating DNA Structures

This is a POVray script to create the DNA structures as was explained in section 5.1. The important issue to note is that only the position of the bp atoms along with the optimized parameters, are enough to construct the different organizational levels up to the chromatin fiber.

The coordinates and parameters showed here, correspond to the C-G base pair in the B-DNA conformation. All quantities are given in Å. Notice the rotation senses, they are left-handed-like because the POV-coordinates so are.

```
#include "colors.inc"
// Camera
#declare V_camera=<0,2000,0>;
background { color White }
camera{
    location V_camera
    look_at <0,360,0>
    right x*320/240
    }
light_source{V_camera,color White }
//Atoms
 #declare  V_center=array[64]
 #declare  V_colors=array[64]
 #declare  F_radius=array[64]

 #declare V_center[1]=<-9.18,-.75,-2.46>;
 #declare V_colors[1]=<1,0,0>;
 #declare F_radius[1]=1.90;
 #declare V_center[2]=<-10.50,-.05,-2.24>;
 #declare V_colors[2]=<1,0,0>;
 #declare F_radius[2]=1.40;
 #declare V_center[3]=<-8.98,-1.99,-1.67>;
 #declare V_colors[3]=<1,0,0>;
 #declare F_radius[3]=1.40;
 #declare V_center[4]=<-7.96,.24,-2.21>;
```

```
#declare V_colors[4]=<1,0,0>;
#declare F_radius[4]=1.40;
#declare V_center[5]=<-7.74,1.32,-3.15>;
#declare V_colors[5]=<1,0,0>;
#declare F_radius[5]=1.70;
#declare V_center[6]=<-7.20,2.53,-2.42>;
#declare V_colors[6]=<1,0,0>;
#declare F_radius[6]=1.70;
#declare V_center[7]=<-5.76,2.35,-2.21>;
#declare V_colors[7]=<1,0,0>;
#declare F_radius[7]=1.40;
#declare V_center[8]=<-7.76,2.78,-1.02>;
#declare V_colors[8]=<1,0,0>;
#declare F_radius[8]=1.70;
#declare V_center[9]=<-7.64,4.15,-.63>;
#declare V_colors[9]=<1,0,0>;
#declare F_radius[9]=1.40;
#declare V_center[10]=<-6.85,1.93,-.14>;
#declare V_colors[10]=<1,0,0>;
#declare F_radius[10]=1.70;
#declare V_center[11]=<-5.51,2.06,-.85>;
#declare V_colors[11]=<1,0,0>;
#declare F_radius[11]=1.70;
#declare V_center[12]=<-4.69,.82,-.80>;
#declare V_colors[12]=<1,1,0>;
#declare F_radius[12]=1.50;
#declare V_center[13]=<-3.31,.96,-.65>;
#declare V_colors[13]=<1,1,0>;
#declare F_radius[13]=1.70;
#declare V_center[14]=<-2.83,2.09,-.56>;
#declare V_colors[14]=<1,1,0>;
#declare F_radius[14]=1.40;
#declare V_center[15]=<-2.57,-.36,-.60>;
#declare V_colors[15]=<1,1,0>;
#declare F_radius[15]=1.50;
#declare V_center[16]=<-3.10,-1.38,-.70>;
#declare V_colors[16]=<1,1,0>;
#declare F_radius[16]=1.70;
#declare V_center[17]=<-2.30,-2.43,-.65>;
```

```
#declare V_colors[17]=<1,1,0>;
#declare F_radius[17]=1.50;
#declare V_center[18]=<-4.51,-1.54,-.85>;
#declare V_colors[18]=<1,1,0>;
#declare F_radius[18]=1.70;
#declare V_center[19]=<-5.26,-.41,-.90>;
#declare V_colors[19]=<1,1,0>;
#declare F_radius[19]=1.70;
#declare V_center[20]=<-8.69,1.58,-3.64>;
#declare V_colors[20]=<1,0,0>;
#declare F_radius[20]=1.20;
#declare V_center[21]=<-7.01,1.00,-3.91>;
#declare V_colors[21]=<1,0,0>;
#declare F_radius[21]=1.20;
#declare V_center[22]=<-7.39,3.42,-3.04>;
#declare V_colors[22]=<1,0,0>;
#declare F_radius[22]=1.20;
#declare V_center[23]=<-4.94,2.88,-.39>;
#declare V_colors[23]=<1,0,0>;
#declare F_radius[23]=1.20;
#declare V_center[24]=<-7.19,.89,-.11>;
#declare V_colors[24]=<1,0,0>;
#declare F_radius[24]=1.20;
#declare V_center[25]=<-6.80,2.33,.88>;
#declare V_colors[25]=<1,0,0>;
#declare F_radius[25]=1.20;
#declare V_center[26]=<-8.81,2.46,-.95>;
#declare V_colors[26]=<1,0,0>;
#declare F_radius[26]=1.20;
#declare V_center[27]=<-6.34,-.49,-1.02>;
#declare V_colors[27]=<1,1,0>;
#declare F_radius[27]=1.20;
#declare V_center[28]=<-4.96,-2.53,-.92>;
#declare V_colors[28]=<1,1,0>;
#declare F_radius[28]=1.20;
#declare V_center[29]=<-1.29,-2.31,-.53>;
#declare V_colors[29]=<1,1,0>;
#declare F_radius[29]=1.20;
#declare V_center[30]=<-2.70,-3.38,-.73>;
```

```
#declare V_colors[30]=<1,1,0>;
#declare F_radius[30]=1.20;
#declare V_center[31]=<8.52,-1.99,1.70>;
#declare V_colors[31]=<1,0,1>;
#declare F_radius[31]=1.90;
#declare V_center[32]=<9.89,-1.47,1.48>;
#declare V_colors[32]=<1,0,1>;
#declare F_radius[32]=1.40;
#declare V_center[33]=<8.15,-3.19,.91>;
#declare V_colors[33]=<1,0,1>;
#declare F_radius[33]=1.40;
#declare V_center[34]=<7.45,-.84,1.45>;
#declare V_colors[34]=<1,0,1>;
#declare F_radius[34]=1.40;
#declare V_center[35]=<7.38,.27,2.39>;
#declare V_colors[35]=<1,0,1>;
#declare F_radius[35]=1.70;
#declare V_center[36]=<7.02,1.54,1.66>;
#declare V_colors[36]=<1,0,1>;
#declare F_radius[36]=1.70;
#declare V_center[37]=<5.57,1.56,1.45>;
#declare V_colors[37]=<1,0,1>;
#declare F_radius[37]=1.40;
#declare V_center[38]=<7.61,1.71,.26>;
#declare V_colors[38]=<1,0,1>;
#declare F_radius[38]=1.70;
#declare V_center[39]=<7.68,3.08,-.13>;
#declare V_colors[39]=<1,0,1>;
#declare F_radius[39]=1.40;
#declare V_center[40]=<6.59,.99,-.62>;
#declare V_colors[40]=<1,0,1>;
#declare F_radius[40]=1.70;
#declare V_center[41]=<5.28,1.30,.09>;
#declare V_colors[41]=<1,0,1>;
#declare F_radius[41]=1.70;
#declare V_center[42]=<4.30,.19,.04>;
#declare V_colors[42]=<1,1,0>;
#declare F_radius[42]=1.50;
#declare V_center[43]=<4.51,-1.17,.12>;
```

```
#declare V_colors[43]=<1,1,0>;
#declare F_radius[43]=1.70;
#declare V_center[44]=<3.42,-1.89,.04>;
#declare V_colors[44]=<1,1,0>;
#declare F_radius[44]=1.50;
#declare V_center[45]=<2.40,-.94,-.10>;
#declare V_colors[45]=<1,1,0>;
#declare F_radius[45]=1.70;
#declare V_center[46]=<1.00,-1.12,-.23>;
#declare V_colors[46]=<1,1,0>;
#declare F_radius[46]=1.70;
#declare V_center[47]=<.35,-2.16,-.25>;
#declare V_colors[47]=<1,1,0>;
#declare F_radius[47]=1.40;
#declare V_center[48]=<.34,.12,-.35>;
#declare V_colors[48]=<1,1,0>;
#declare F_radius[48]=1.50;
#declare V_center[49]=<.96,1.35,-.35>;
#declare V_colors[49]=<1,1,0>;
#declare F_radius[49]=1.70;
#declare V_center[50]=<.16,2.41,-.48>;
#declare V_colors[50]=<1,1,0>;
#declare F_radius[50]=1.50;
#declare V_center[51]=<2.27,1.52,-.22>;
#declare V_colors[51]=<1,1,0>;
#declare F_radius[51]=1.50;
#declare V_center[52]=<2.93,.34,-.10>;
#declare V_colors[52]=<1,1,0>;
#declare F_radius[52]=1.70;
#declare V_center[53]=<8.36,.39,2.88>;
#declare V_colors[53]=<1,0,1>;
#declare F_radius[53]=1.20;
#declare V_center[54]=<6.62,.05,3.15>;
#declare V_colors[54]=<1,0,1>;
#declare F_radius[54]=1.20;
#declare V_center[55]=<7.33,2.39,2.28>;
#declare V_colors[55]=<1,0,1>;
#declare F_radius[55]=1.20;
#declare V_center[56]=<4.83,2.19,-.37>;
```

```
#declare V_colors[56]=<1,0,1>;
#declare F_radius[56]=1.20;
#declare V_center[57]=<6.78,-.09,-.65>;
#declare V_colors[57]=<1,0,1>;
#declare F_radius[57]=1.20;
#declare V_center[58]=<6.60,1.40,-1.64>;
#declare V_colors[58]=<1,0,1>;
#declare F_radius[58]=1.20;
#declare V_center[59]=<8.60,1.24,.19>;
#declare V_colors[59]=<1,0,1>;
#declare F_radius[59]=1.20;
#declare V_center[60]=<5.50,-1.61,.24>;
#declare V_colors[60]=<1,1,0>;
#declare F_radius[60]=1.20;
#declare V_center[61]=<-.65,.10,-.44>;
#declare V_colors[61]=<1,1,0>;
#declare F_radius[61]=1.20;
#declare V_center[62]=<.56,3.36,-.48>;
#declare V_colors[62]=<1,1,0>;
#declare F_radius[62]=1.20;
#declare V_center[63]=<-.86,2.28,-.58>;
#declare V_colors[63]=<1,1,0>;
#declare F_radius[63]=1.20;

//bp construction
#declare i=1;
#declare bp=object{
    union{
       #while(i<64)
           sphere{
              V_center[i], F_radius[i]
              pigment{rgb V_colors[i]}
           }
           #declare i=i+1;
       #end
       }
    }
```

```
// Helix Construction
#declare pitch=3.3; //axial step
#declare nbpturn=10;
#declare angz=360/nbpturn;
#declare helix=object{
    union{
       #declare ibp=0;
       #while(ibp<18)
           object{
              bp
              #declare a=angz*ibp;
              #declare p=pitch*ibp;
              rotate -a*z
              translate <0,0,p>
              }
           #declare ibp=ibp+1;
       #end
       }
    }

// Nucleosome Construction
#declare rnucl=52.3;
#declare Dhel=23.7;
#declare bpnucl=154;
#declare nucl=object{
    union{
       sphere{
          #declare rhist=rnucl-1.1*Dhel;
          <0,Dhel,0>
          rhist
          pigment{rgb <0,0,1>}
          }
       #declare ibp=0;
       #declare rc=rnucl-Dhel/2;
       #while(ibp<154)
           object{
              bp
              #declare a=angz*ibp;
              rotate -a*z
```

```
            #declare b=(360/(bpnucl/2))*(ibp+0.5);
            rotate  -b*y
            #declare p=(2*Dhel)*(ibp/bpnucl);
            translate<rc*cos(b*pi/180),p,rc*sin(b*pi/180)>
            }
        #declare ibp=ibp+1;
    #end
    }
}

// Linker Fragment Construction
#declare Dchro=314.8;
#declare rbind=Dchro/2-Dhel;
#declare a0=bpnucl*360/nbpturn; //bp angle at the end of the nucleosome
#declare bpbind=46;
#declare bind=object{
    union{
        #declare ibp=0;
        #declare mu=20/pi;
        #declare H=1.5*Dhel;
        #while(ibp<46)
            object{
                bp
                #declare a=a0+angz*(ibp);
                rotate -a*z
                #declare b=(ibp+0.5)*60/bpbind;
                rotate -b*y
                #declare beta=pi/6-b*pi/180;
                #if (beta>0)
                    #declare s=1;
                #else
                    #declare s=-1;
                #end
                #declare f=0.5*(1-s*(1-exp(-mu*abs(beta))));
                translate <rbind*cos(b*pi/180),
                           2.5*Dhel-H*f,
                           rbind*sin(b*pi/180)>
                }
            #declare ibp=ibp+1;
```

```
        #end
        }
    }

// Block=Nucleosome + Binding Fragment
#declare block=object{
    union{
        #declare r=Dchro/2-rnucl;
        object{
            nucl
            translate <r,Dhel/2,0> //place the base of the nucleosome
                                   //in the xz plane
            }
        object{bind}
        }
    }

// Chromatin Construction (10 levels)
#declare chro=object{
    union{
        #declare inucl=0;
        #declare r=Dchro/2-rnucl;
        #while(inucl<60)
            #declare a=60*(inucl+0.5);
            object{
               block
               rotate -a*y
               translate <0,inucl*(3*Dhel)/6,0>
               }
            #declare inucl=inucl+1;
        #end
        }
    }

// Nucleus construction (5x5 mesh)
#declare nucleus=object{
    union{
        #declare ix=-2;
        #while(ix<3)
```

```
            #declare iy=-2;
            #while(iy<3)
                object{
                   chro
                   translate <Dchro*ix,0,Dchro*iy>
                   }
                #declare iy=iy+1;
            #end
            #declare ix=ix+1;
        #end
        }
    }

nucleus    //Visualization
```

APPENDIX B – Solution of the DRAT differential equation system

The following differential equation system is taken from [38] and is solved for the particular case of acute irradiation. That is, for $\mu << t^{-1}$ and constant $\dot{D}$.

$$\dot{N}_{RL}(t) = 2p\dot{D} - \mu N_{RL}(t) \tag{B.1}$$

$$\dot{N}_{LL}(t) = \alpha\dot{D} + p\epsilon\dot{D}N_{RL}(t) \tag{B.2}$$

Eq. B.1 can be cast as,

$$e^{\mu t}\dot{N}_{RL}(t) + \mu e^{\mu t}N_{RL}(t) = 2p\dot{D}e^{\mu t} \tag{B.3}$$

$$\frac{d}{dt}[e^{\mu t}N_{RL}(t)] = 2p\dot{D}e^{\mu t} \tag{B.4}$$

where $e^{\mu t}$ is the integrating factor of the equation. Since $\dot{D}$ is constant, the Eq. B.4 is easily integrated from 0 to t. Thus,

$$e^{\mu t}N_{RL}(t) = 2p\dot{D}\mu^{-1}(e^{\mu t} - 1) \tag{B.5}$$

or,

$$N_{RL}(t) = 2p\dot{D}\mu^{-1}(1 - e^{-\mu t}) \tag{B.6}$$

Now this $N_{RL}(t)$ can be inserted into Eq. B.2 and the integration is performed. This leads to, (c.g Eq 3.9)

$$\dot{N}_{LL}(t) = \alpha\dot{D} + 2p^2\mu^{-1}\epsilon\dot{D}^2(1 - e^{-\mu t}) \tag{B.7}$$

$$N_{LL}(t) - \cancelto{0}{N_{LL}(0)} = \alpha\dot{D}t + 2p^2\mu^{-1}\epsilon\dot{D}^2(t + \mu^{-1}(e^{-\mu t} - 1)) \tag{B.8}$$

but $\mu t << 1$, so, $e^{-\mu t} - 1 \approx -\mu t + \frac{1}{2}\mu^2 t^2$. Thus,

$$N_{LL}(t) = \alpha\dot{D}t + 2p^2\epsilon\dot{D}^2\mu^{-1}\left(\frac{1}{2}\mu t^2\right) = \alpha\dot{D}t + p^2\epsilon\dot{D}^2t^2 \tag{B.9}$$

which, finally leads to Eq 3.10 by putting $\beta = p^2\epsilon$ and $D = \dot{D}t$.