



Benemérita Universidad Autónoma de Puebla
Facultad de Ciencias Físico-Matemáticas

**Simulación Monte Carlo de plásmidos súper
enrollados de ADN para el estudio del daño
directo e indirecto de radiación ionizante de
baja LET.**

Tesis Presentada a la

Facultad de Ciencias Físico-Matemáticas

como requisito para la obtención del grado de

Maestría

por

Jorge Naoki Dominguez Kondo

asesorado por

Eduardo Moreno Barbosa & José Asunción
Ramos Mendez

Puebla, Pue.

Noviembre de 2018

Simulación Monte Carlo de
plásmidos súper enrollados de
ADN para el estudio del daño
directo e indirecto de
radiación ionizante de baja
LET.

Tesis

Eduardo Moreno Barbosa

José Asunción Ramos Mendez

Jorge Naoki Dominguez Kondo



Título: Simulación Monte Carlo de plásmidos súper enrollados de ADN para el estudio del daño directo e indirecto de radiación ionizante de baja LET..

Estudiante: Jorge Naoki Dominguez Kondo

COMITÉ

Javier M. Hernández
Presidente

Alexandra Deriabina
Secretario

Jorge Velázquez Castro
Vocal

Eduardo Moreno Barbosa
Asesor I

Jose A. Ramos Méndez
Asesor II

Para

Todas las personas presentes en mi vida y
aquellas que ya no están, por hacerme la
persona que soy hoy.

Índice general

1. Introducción	1
2. Teoría	9
2.1. Simulación de Plásmidos de ADN	9
2.1.1. Plásmidos Súper Enrollados de ADN	9
2.1.2. El Formato PDB para Generar un Modelo Atómico para Plásmidos de ADN	16
2.1.3. Clasificación de Rompimientos de ADN	18
2.1.4. Códigos Monte Carlo <i>Track Structure</i>	19
2.2. Simulación de la Etapa Química Heterogénea de la Radiólisis del Agua . . .	25
2.2.1. Método Paso a Paso	26
2.2.2. Tiempos de Reacción Independiente	27
2.2.3. Uso de Disolventes Químicos	29
2.3. Interacción Forzada de Rayos γ	30
3. Metodología	31
3.1. Plásmidos Super Enrollados y Estructura del Capítulo	31
3.2. Primera Simulación	33

3.2.1. Geometría	33
3.2.2. Irradiación	35
3.2.3. <i>Scorers</i>	36
3.3. Segunda Simulación	37
3.3.1. Geometría	37
3.3.2. Irradiación	38
3.3.3. <i>Scorers</i>	38
3.4. Tercera Simulación	38
3.4.1. Geometría	38
3.4.2. Irradiación	39
3.4.3. <i>Scorers</i>	39
4. Resultados y Discusiones	41
4.1. Plásmidos de Súper Enrollados	41
4.2. Modelo de ADN	45
4.3. Primera Simulación	48
4.4. Segunda Simulación	49
4.5. Tercera Simulación	53
4.6. Prueba de Configuración del ADN	58
5. Conclusiones	59
A. Monte Carlo	61
A.1. Historia del Método Monte Carlo	61

A.2. Básicos del Método Monte Carlo	63
A.3. Técnicas de Muestreo	63
A.4. Generadores de Números Aleatorios	64
A.5. Muestreo Aleatorio	65
A.6. Incertidumbres del Método Monte Carlo	66
B. Tablas de Reacciones Químicas	69

Resumen

Propósito: El uso de modelos geométricos más detallados de ADN, como plásmidos súper enrollados, integrados a los códigos de Monte Carlo track structure constituyen una herramienta poderosa para el estudio, desde primeros principios, del efecto biológico de la interacción de la radiación ionizante. El objetivo de este trabajo de tesis fue el desarrollar una herramienta Monte Carlo para la simulación de plásmidos súper enrollados de ADN. Los modelos se integraron al software track structure TOPAS-nBio (basado en Geant4-DNA) para el estudio del daño producido por los depósitos de energía causados por la radiación ionizante y por las reacciones químicas entre los componentes del ADN y las especies químicas producidas por la subsecuente radiólisis del agua.

Metodología: Se utilizaron los modelos de ADN existentes en TOPAS-nBio además de las conformaciones super enrolladas generadas en este trabajo. Dichas conformaciones correspondieron al plásmido pBR322 (4362 bases pares de longitud) y fueron creadas a partir del método de Monte Carlo. Para esto, se aplicó un proceso iterativo de deformaciones a un polígono regular de 145 lados (cada lado con 10.2 nm de longitud), rechazando conformaciones a partir de la energía elástica total y la presencia de nudos, hasta alcanzar el equilibrio termodinámico. Dicho criterio fue complementado utilizando el radio de giro del sistema. Se realizaron estudios del efecto de la temperatura en la generación de plásmidos y se reportó la energía y el número de torcimiento en función del número de iteraciones.

Subsecuentemente se desarrolló un modelo geométrico de elipsoides para representar a los nucleótidos de cada base par de ADN. Este modelo se alineó a las posiciones de los átomos de cada nucleótido obtenidas del protein data bank. El modelo combinado representó

una base par, y este fue colocado repetidamente a lo largo de la trayectoria descrita por la conformación super enrollada, a cada 0.34 nm de separación en incrementos de 36°. Este modelo geométrico fue comparado con los modelos de ADN existentes en TOPAS-nBio creados a base de segmentos de cilindros. Para esto, se realizaron simulaciones track structure utilizando espacios fase generados con protones de 10, 20 y 30 MeV incidiendo sobre plásmidos circulares de 2000 bases par de longitud, utilizando el modelo de elipsoides y el modelo de segmentos cilíndricos. Los resultados del daño directo al ADN fueron cuantificados mediante el número de rompimientos simples (SSB) y dobles (DSB) de ADN; posteriormente, se compararon con datos experimentales de la literatura.

Por otro lado, se realizaron simulaciones de radiólisis del agua utilizando, por primera vez, ^{60}Co como fuente de irradiación. Para esto, se validó la técnica de reducción de varianza interacción forzada para fotones a través de la comparación del valor G (número de especies químicas por cada 100 eV de energía depositada) en función del tiempo, con simulaciones establecidas utilizando tracks cortos de $100\mu\text{m}$ de longitud de protones de 300 MeV ($0.3\text{ keV}/\mu\text{m}$) y datos experimentales de la literatura. Adicionalmente, se realizaron simulaciones complementarias en presencia 1 mM de TRIS y 0.1 mM de EDTA a 1 atm de presión. La realización de las reacciones en la etapa química de la radiólisis se llevó a cabo utilizando el método de tiempos de reacción independiente (IRT) de TOPAS-nBio.

Por último, se extendió el algoritmo IRT de TOPAS-nBio para incluir la información del track geométrico de los plásmidos desarrollados y para permitir la simulación de tracks físico-químicos acumulados de diferentes historias, esto permitió la simulación de SSB y DSB a diferentes dosis de radiación prescrita. Utilizando estas modificaciones se realizó una simulación que consistió de seis plásmidos pBR322 a 293 K colocados aleatoriamente en un cubo de agua de 988 nm^3 , irradiados uniformemente por fotones de ^{60}Co (con interacción forzada para Compton escalado a $0.57\mu\text{m}$) para dosis prescritas de 6 a 100 Gy. El sistema incluyó 1 mM de TRIS, 0.1 mM de ETDA y 1.3 mM de O₂. Se consideró la generación de daño indirecto cuando una reacción del radical $\bullet\text{OH}$ ocurrió con el radical C4 de algún nucleótido usando un 50% de probabilidad. Se reportaron los SSB y DSB normalizados por el número de plásmidos en función de la dosis prescrita y se compararon con datos calculados y experimentales de la literatura.

Resultados: La herramienta Monte Carlo, desarrollada en este trabajo de tesis, pa-

ra generar plásmidos super enrollados tomó en promedio 8 hrs, utilizando un algoritmo de disminución exponencial de la temperatura, para generar un plásmido pBR322 hasta el equilibrio termodinámico con un número de torcimiento consistente con los reportados previamente en la literatura. El uso de un algoritmo de disminución exponencial de temperatura resultó en la forma más eficiente de producir plásmidos con conformaciones equivalentes a las conformaciones de temperatura constante.

El efecto en la cuantificación de SSB y DSB del uso del modelo de elipsoides comparado con el de segmentos cilíndricos fue una clara diferencia de hasta el -38 %, con un error estadístico menor del 1 %, atribuido a la diferencia de volumen y a la distribución espacial de cada nucleótido.

El uso de ^{60}Co directamente en el cálculo de los valores G produjo resultados equivalentes a las simulaciones establecidas que utilizan tracks cortos de protones. Se encontraron diferencias dentro del 0.85 % para el caso del hidroxilo, 2.45 % para el electrón solvado y de 11.43 % para el hidrogeno, mientras que ambas simulaciones tuvieron buena concordancia, dentro de la esparsidad de los datos experimentales. Esto confirma a la técnica interacción forzada como una técnica de reducción de varianza para Monte Carlo track structure. La simulación de solventes químicos redujo adecuadamente el número de especies químicas hacia el estado estable de la etapa química.

Finalmente, para la simulación del sistema de seis plásmidos se logró disminuir el error estadístico de la literatura. Para SSB el error máximo se redujo de 48 % a 0.5 % y para los DSB se infiere una disminución de al menos el mismo factor; no fue posible dar el número exacto por la falta de dicho dato en la literatura reportada. Comparado con los datos experimentales, se encontró una diferencia de $6 \% \pm 1 \%$ para DSB y $117.5 \% \pm 2 \%$ para SSB, mejorando lo reportado en otros trabajos. El grado de super enrollamiento de los plásmidos influyó en el número de SSB, siendo reducido en un $8.5 \% \pm 2.1 \%$ a mayor número de torcimiento.

Conclusiones: Se desarrolló una herramienta Monte Carlo para la generación de plásmidos super enrollados de ADN y se extendieron funcionalidades de TOPAS-nBio. Esto permitió mejorar la comparación con datos experimentales de la literatura, para la determinación del daño al ADN usando modelos geométricos más detallados y condiciones

más cercanas a las condiciones experimentales.

Estructura de la tesis Este trabajo está separado en cuatro capítulos y un apéndice, en el primer capítulo se comienza con una introducción general a la importancia de los plásmidos y por qué son tan utilizados en la investigación. Seguido de una breve descripción de las capacidades de los códigos Monte Carlo de track structure para el estudio de fenómenos en geometrías en escalas de nanómetros.

En el segundo capítulo, se describe la teoría de simulación de plásmidos en su conformación súper enrollada basado en métodos descritos en la literatura; se describe la física de los procesos involucrados en el código Geant4-DNA de Monte Carlo track structure y TOPAS-nBio, los algoritmos de reacción difusión de especies químicas basado en el método de paso a paso y en el tiempos de reacción independientes (IRT); por último, se describe una técnica de reducción de varianza para la interacción forzada de fotones en agua.

En el tercer capítulo se describe la metodología describiendo la implementación del algoritmo de generación de plásmidos y tres etapas que involucran la irradiación de modelos geométricos de ADN con su comparación experimental. Por último, una sección dedicada a los resultados y su discusión, para finalizar con una conclusión en el quinto capítulo.

El apéndice incluye las bases del método Monte Carlo, y una recopilación de tablas de reacciones químicas para el algoritmo de reacción y difusión de especies químicas.

Capítulo 1

Introducción

Es sabido que una alta cantidad de dosis de radiación ionizante es capaz de causar daño en el cuerpo humano. Es por esto que áreas de investigación de la física, química y biología se han dedicado a estudiar los mecanismos de interacción de la radiación con el tejido biológico. Algunos campos de estudio se centran en mejorar las técnicas de radioprotección, por ejemplo, la industria aeroespacial o la protección radiológica en hospitales; otros campos de estudio se enfocan en estudiar los efectos de la radiación en volúmenes bien definidos, que da lugar a la dosimetría en física médica para el tratamiento de cánceres y tumores.

Sin embargo, para entender de manera más precisa el fenómeno de la interacción de la radiación con el tejido biológico, es necesario utilizar escalas más pequeñas como la celular ($10\ \mu\text{m}$ de diámetro), hasta llegar a escalas comparables al diámetro del ADN ($2.3\ \text{nm}$), pero esto dificulta el estudio de esta área. A estas escalas, la modelación utilizando métodos de Monte Carlo es esencial para entender e interpretar las mediciones experimentales.

El ácido desoxirribonucleico (ADN) se descubrió en 1868 por el médico suizo Johannes Friedrich Miescher y posteriormente la estructura del ADN en 1953 por Rosalind Franklin, James D. Watson y Francis Crick [60]. Watson lo describió como “el causante de todos los procesos biológicos por los que pasan las células” [27]; desde el procesar alimento, hasta repararse y dividirse. Por esta razón, el ADN es el pilar sobre el cual se sustenta la vida y por lo tanto, si éste deja de ser útil a la célula que lo contiene, ésta es incapaz de

funcionar y muere. Independientemente de los procesos biológicos que causan daño al ADN, se demostró que la radiación ionizante también puede causar daño significativo, como se discutirá más adelante. Este daño puede culminar con mutaciones o una falla irreparable que eventualmente llevan a la muerte celular.

En el modelo de doble hélice de Franklin-Watson-Crick, a lo largo de la doble hélice existe un arreglo de parejas de nucleótidos, cada pareja es llamada base par (bp). Existen cuatro tipos de nucleótidos: Adenina, Timina, Citosina y Guanina. Cada nucleótido tiene una pareja predeterminada, por lo que, al conocer la cadena de una hélice (hebras, o strands en inglés), la información de la otra es implícita. Estas parejas son de la siguiente forma: Adenina - Timina, Timina - Adenina, Citosina - Guanina, Guanina - Citosina. Estas combinaciones suelen abreviarse por la primera letra: AT, TA, CG, GC.

Una cadena de ADN puede tener alguna de tres configuraciones distintas, dependiendo de las secuencias y la dirección del súper enrollamiento, estas pueden ser ADN-A, ADN-B y ADN-Z, ver Figura 1.1.

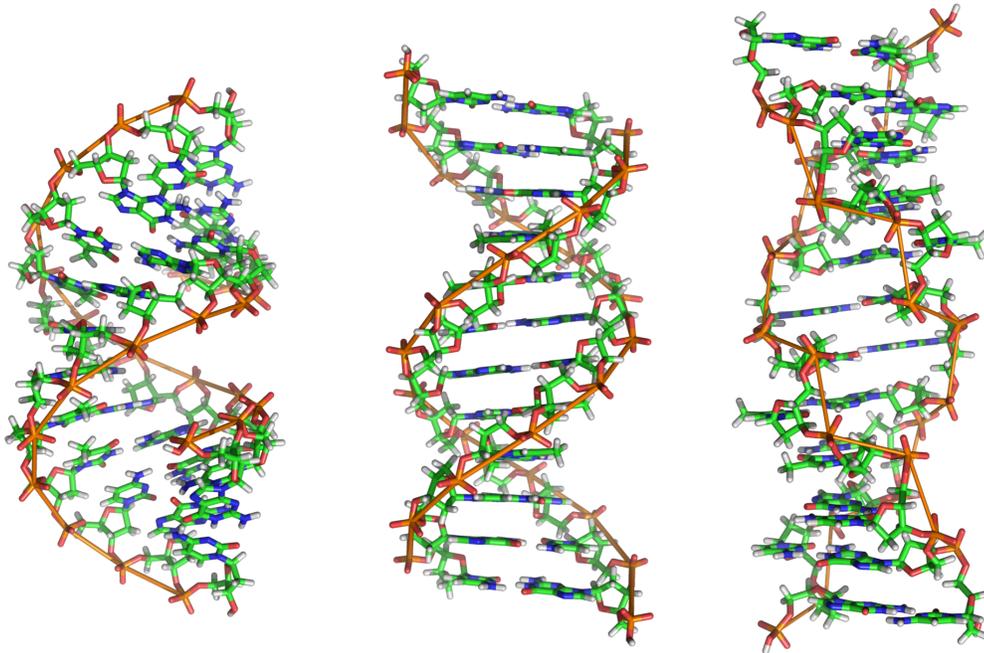


Figura 1.1: De izquierda a derecha, tenemos una cadena de ADN-A, ADN-B y ADN-Z. Estas cadenas están compuestas de diez bases pares, aproximadamente 3.4 nm de longitud total. Autor de la imagen Richard Wheeler (Zephyris) Wikipedia.

La configuración de ADN-B es la más común bajo condiciones fisiológicas, por lo que

esta configuración se utilizó en este proyecto de tesis. De aquí en adelante, en este trabajo cuando se mencionen las configuraciones de ADN para los modelos y simulaciones se estará refiriendo al ADN-B.

El ADN se encuentra en su mayoría dentro de una estructura en el núcleo celular llamada cromatina, esta es la estructura principal del ADN de las células. Sin embargo, también puede agruparse en otra conformación fuera del núcleo celular, este, al ser independiente, se puede replicar cuantas veces sea necesario y su efecto en la célula es el dar inmunidad ante ciertas condiciones (por ejemplo, a los antibióticos), a estas conformaciones de ADN se les conoce como plásmidos [8], Figura 1.2.

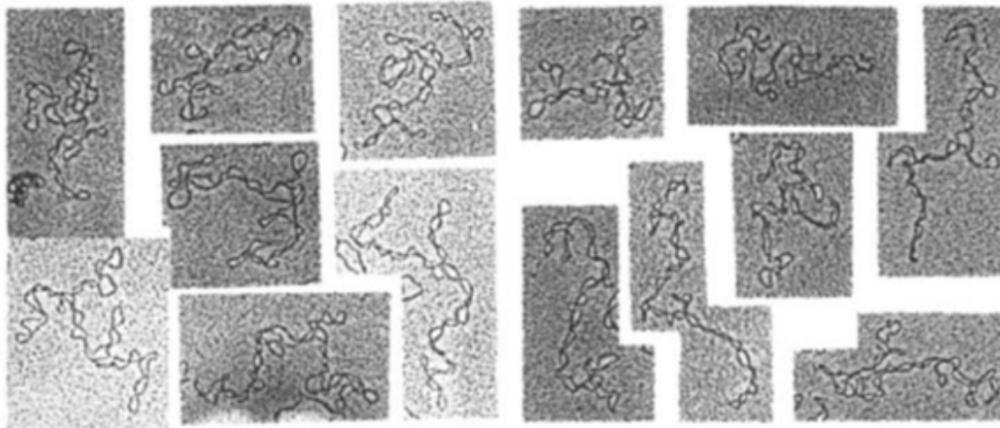


Figura 1.2: Micrografía electrónica de dos plásmidos con diferentes proyecciones, cada plásmido está compuesto por 7 kbps. Imagen tomada de Cozzarelli, 1994 [14].

Los plásmidos son estructuras de ADN que se pueden encontrar en diferentes conformaciones:

- Circulares abiertos.
- Súper enrollados
- Lineales.

Los plásmidos súper enrollados son los más comunes in vivo y se ha reportado que cuando estos se dañan pueden convertirse en circulares abiertos o lineales.

Los plásmidos varían en tamaño, pueden ser pequeños (2.3 kbps en el caso del plásmido pUB110) o enormes como los llamados mega plásmidos (1354.2 kbps en el caso del plásmido

pSymA). Los plásmidos tienen la cualidad de ser modificables artificialmente sin dañarlos, esto permite producir plásmidos sintéticos y reproducibles. Entonces, se puede diseñar plásmidos que induzcan comportamientos específicos en las células, como por ejemplo, la capacidad de alimentarse de cierta sustancia, la de agruparse en regiones específicas en una caja de Petri, o hacerlas inmunes a ciertos químicos para crear colonias celulares sin contaminación [8].

Sin embargo, solo se han encontrado plásmidos mayoritariamente en bacterias, aunque también se han visto en procariotas y eucariotas. Y debido a que suponen un consumo de energía para el huésped, este conservará a los plásmidos si el beneficio es equivalente al coste energético de manutención. Múltiples plásmidos han desarrollado un sistema en el cual mantienen como rehén a su huésped, generando toxinas de larga duración y suministrando un antídoto de corta duración, haciendo que la célula no pueda rechazar al plásmido.

Los plásmidos también son fáciles de duplicar, una vez que un biólogo ha modificado uno, solo lo tiene que introducir en un huésped y este lo duplicará automáticamente. También son heredables, cuando una célula se va a dividir, entre más concentración de plásmidos tenga, es más probable que ambas células terminen con al menos un plásmido.

Muchos de los procesos de replicación y reparación del ADN fueron estudiados a partir de los plásmidos, y se ha observado que los plásmidos pueden morir si sus hebras se rompen hasta tal punto que no puedan repararse. Los procesos por los cuales el ADN se daña siguen siendo un área de estudio hoy en día, sin embargo, es sabido que la radiación ionizante es un causante de daño en las hebras de ADN. En el estudio de daño por radiación, algunos investigadores [11] [26] [38] [56] han estudiado los rompimientos causados en el ADN utilizando métodos de Monte Carlo comparado con datos experimentales. El artículo detonador de tales estudios corresponde al de Charlton de 1989 [11]. En ese artículo, se reporta una simulación Monte Carlo de *track structure* de la irradiación de hebras de ADN. La finalidad es muestrear las posiciones pre-calculadas de los eventos de ionización producidas por partículas mono-energéticas cargadas en agua utilizando una representación geométrica de ADN. Como modelo geométrico de ADN se utiliza una geometría simplista. Cada base par es representada por tres secciones de cilindros diferentes, cuyo volumen se aproxima al de una base par de ADN (2.3 nm de diámetro y 0.34 nm de longitud), con la finalidad de contabilizar la energía depositada dentro de cada cilindro. Se asume que un

rompimiento simple de ADN ocurre si la energía depositada acumulada en un elemento geométrico (sección de cilindro) es mayor a un valor umbral y se estima el número de rompimientos dobles en función de la calidad de radiación (carga y energía). Al comparar con datos experimentales, se establece que el valor umbral para producir un rompimiento simple es 17.5 eV. En otro estudio, McNamara y colaboradores incluyen mayor detalle experimental en la simulación Monte Carlo al utilizar un anillo (representando un plásmido de ADN) con el mismo modelo del ADN de Charlton, el haz de radiación con los mismos parámetros experimentales y contabilizando directamente la energía depositada en cada volumen contabilizando así el número de rompimientos en un plásmido individual. Esta es la primera vez que condiciones experimentales son simuladas explícitamente en lugar de idealizadas representaciones [38]. Tomita, por otro lado, decide enfocarse en el número de rompimientos como consecuencia de la interacción de especies químicas ya que el 98 % del daño en el ADN es causado por este tipo específico de proceso. En su método, sin embargo, utiliza plásmidos de ADN dentro de una geometría de $988 \times 988 \times 988 \text{ nm}^3$ simulando el espectro de electrones generado por rayos gammas emitidas por una fuente de Cobalto-60. El mecanismo por el cual se generan las especies químicas es también efecto de la radiación ionizante al depositar energía en el medio, a este efecto se le suele llamar radiólisis y está muy bien estudiado en el caso del agua, material que Tomita usa para su medio.

La interpretación de los datos experimentales se realiza mediante simulaciones Monte Carlo de transporte de la radiación para ver el daño causado por depósitos de energía en las hebras de ADN y el daño por las reacciones químicas entre los átomos del ADN y las especies químicas generadas durante el proceso de irradiación. En años anteriores las limitaciones en la capacidad de cómputo de los ordenadores y la falta de programas especializados han dificultado el avance de esta área, esto último, obligando a cada grupo de trabajo a desarrollar su propio método desde cero.

Para estas escalas (nanómetros) es necesario usar códigos de *track structure*, en vez de los de historia condensada (o de propósito general) utilizados para el cálculo de dosis en pacientes. Este tipo de códigos condensan la información de muchas interacciones en una sola disminuyendo el tiempo de las simulaciones en gran medida, obteniendo excelentes resultados para aplicaciones generales como sistemas de planeamiento en hospitales o en el diseño de detectores para radiación. Sin embargo, los códigos de historia condensada no

proveen de la suficiente resolución a escalas comparables a la dimensión del diámetro del ADN, por lo que no son aplicables.

La lista de códigos Monte Carlo de track structure disponibles al día de hoy es variada, entre ellos están los códigos: ETRAB, MOCA8, CPA100, KURBUC, PITS, PARTRAC, RETRACKS por nombrar unos cuantos [45], estos códigos sin embargo están restringidos a los grupos que los desarrollaron.

Por otro lado, dentro de los códigos de propósito general, Geant4 [1] sobresale por ser software libre, con la capacidad de simular muchos tipos de partículas dentro del rango de energías de algunos eV hasta TeV. Geant4 es un conjunto de librerías de C++ con todo lo necesario para hacer una simulación de transporte de la radiación en la materia, y aunque no está especializado en el track structure, su extensión Geant4-DNA si lo está. Geant4-DNA está contenidos dentro de Geant4, el usuario puede acceder a Geant4-DNA definiendo las listas de procesos físicos correspondientes.

Geant4/Geant4-DNA tienen un problema que hace que los usuarios eviten usarlo al principio y es su agresiva curva de aprendizaje. Geant4 explota al máximo posible C++, así que el usuario debe ser experto en este lenguaje de programación y en física del transporte de radiación con la materia; más aún, para aplicaciones en radioterapia, el usuario tiene que ser también experto en física médica.

Por esta razón un grupo de investigadores de la Universidad de California San Francisco, el Laboratorio Nacional SLAC en la Universidad de Stanford y el Hospital General de Massachusetts de la escuela de medicina de la Universidad de Harvard decidió crear TOPAS [47]. Este software es un programa desarrollado por completo en Geant4 que ofrece un lenguaje más sencillo para crear las simulaciones para radioterapia (convencional y de hadrones), en lo que se denominan “archivos de parámetros”. TOPAS fue creado con la idea de ser amigable e intuitivo para el entendimiento de los parámetros utilizados. Sin embargo, no es de *track structure* y al igual de Geant4 tiene su propia versión de Geant4-DNA, TOPAS-nBio [38] [50].

Para aplicaciones más complejas de TOPAS (por ejemplo, que requieran de geometrías inexistentes en la suite de TOPAS), el usuario puede usar lo que llaman “extensiones”, que son códigos escritos con Geant4 para crear geometrías complejas, fuentes de partículas

primarias específicas, volúmenes de detección, entre otros. Y una vez que estas extensiones son escritas, es posible hacer uso de ellas en los archivos de parámetros, que solo requerirá de anexar algunas líneas.

En este trabajo de tesis se usará TOPAS-nBio, creando una serie de extensiones en las cuales se desarrollará la geometría de plásmidos acoplados a modelos del ADN para la simulación del daño directo y posiciones atómicas para el daño indirecto, la fuente de radiación y el volumen de detección. Todo el código desarrollado será incluido en TOPAS-nBio facilitando el llevar a cabo una simulación de este tipo.

Debido a lo extenso del trabajo de tesis, este se dividirá tres etapas. Cada una orientada a poner a prueba diferentes aspectos de la simulación final. En la primera etapa desarrollarán diferentes modelos del ADN; en la segunda etapa se verificará el algoritmo de reacción-difusión de especies químicas de TOPAS-nBio al calcular los valores G (número de especies químicas por cada 100 eV de energía depositada) utilizando fotones de ^{60}Co ; en la tercera etapa, se simularán plásmidos de ADN bajo condiciones experimentales para contabilizar el daño directo e indirecto a estos mismos.

Capítulo 2

Teoría

Las componentes indispensables para llevar a cabo una simulación Monte Carlo del transporte de la radiación son: geometría, fuente de radiación, procesos y modelos físicos, y métodos para registro de información (scorers o tallies en inglés). Componentes opcionales pueden incluir el post procesamiento y análisis de los datos, etcétera. La descripción de los fundamentos del método de Monte Carlo se encuentra en el apéndice I. En este capítulo se describen las bases teóricas de la simulación Monte Carlo de plásmidos súper enrollados de ADN. Ya que el objetivo de esta tesis es utilizar tales geometrías acopladas a simulaciones track-structure que incluyan la radiólisis del agua, se incluye también una descripción general del proceso de simulación de radiólisis con Geant4-DNA y un método acelerado incluido recientemente en TOPAS-nBio. Finalmente, se describe la clasificación de rompimientos de ADN y brevemente una técnica de reducción de varianza.

2.1. Simulación de Plásmidos de ADN

2.1.1. Plásmidos Súper Enrollados de ADN

La topología de los plásmidos súper enrollados se generó utilizando el método adaptativo de Vologodskii (Konstantin V. Klenin, Alexander V. Vologodskii, Vadim V. Anshelevich, 1991). En este método, la simulación comienza con una figura geométrica regular de

N lados iguales. Cada lado tiene una longitud l de 0.962 nm que corresponde a un décimo de la longitud estadística de Kuhn [18]. Si el número de lados N es muy grande, entonces esta figura geométrica adoptará la forma de un círculo. Esta configuración corresponde a la estructura de plásmido circular abierto idealizado como se ve en la Figura 2.1.

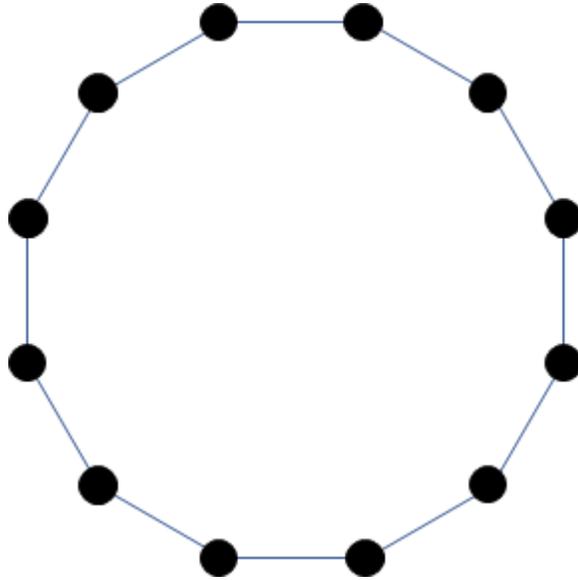


Figura 2.1: Un ejemplo de figura regular inicial para un plásmido idealizado de 12 lados, también se le considera el plásmido más simple. Los círculos negros son solamente incluidos para mostrar a las posiciones de los vértices.

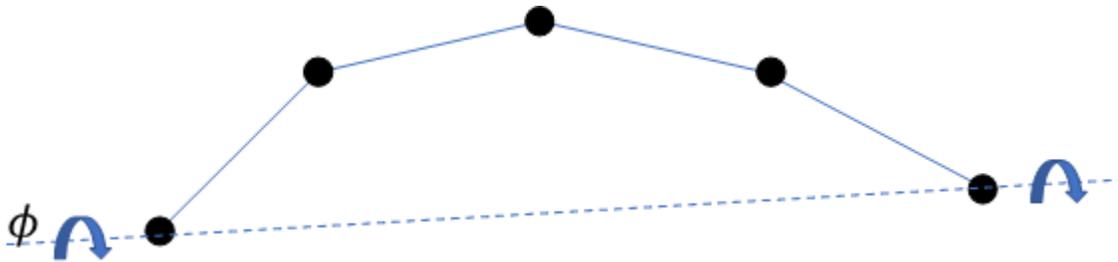


Figura 2.2: Esquema de deformación por rotación de cigüeñal, donde cada punto oscuro representa un vértice de la conformación actual, las líneas sólidas son las aristas imaginarias del plásmido (imaginarias porque solo se trabaja con los vértices en todo momento). La línea punteada es el eje que une los vértices seleccionados para la rotación.

A partir de una configuración de plásmido circular abierto, se aplican dos tipos de deformaciones (movimientos de los vértices diferentes) llamadas “rotación de cigüeñal” (crankshaft) y “reptación”. El movimiento de reptación es opcional, sin embargo sirve para acelerar la formación del plásmido [33] [59] [59]. La rotación de cigüeñal se muestra en la Figura

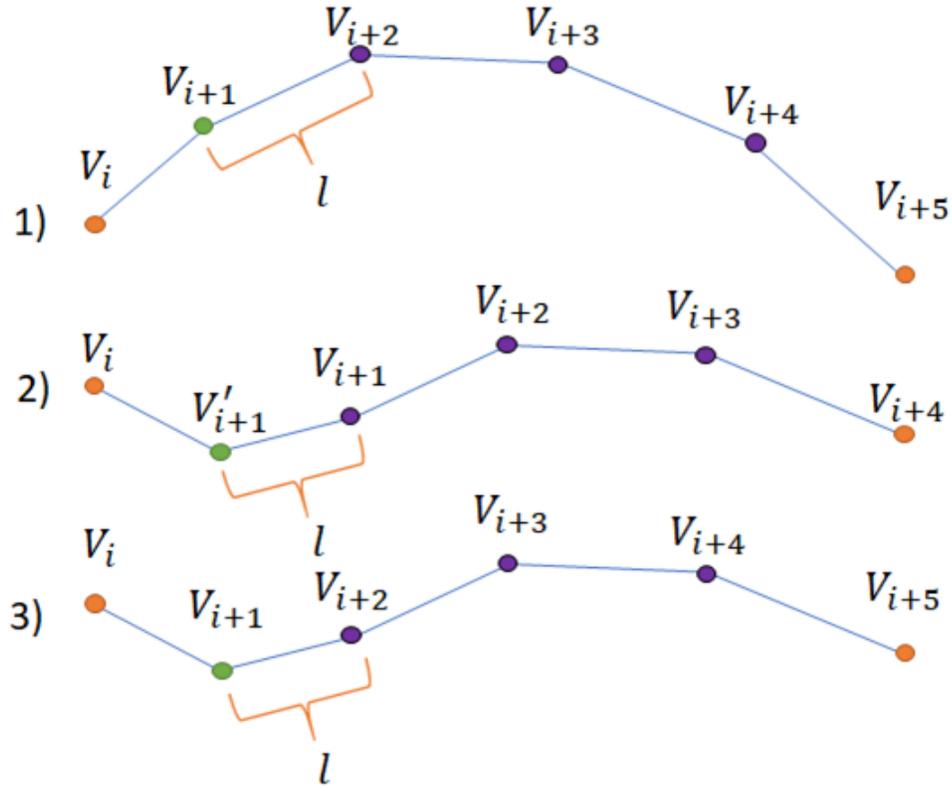


Figura 2.3: Esquema del algoritmo del movimiento de reptación, dividido en 3 etapas, la primera etapa es la selección del sub-segmento, la segunda etapa es la reubicación de los vértices y el cálculo del nuevo punto, la última etapa es el reacomodar los vértices para que coincidan sus índices, los puntos púrpura son los segmentos que conservaran su forma, los puntos rojos son los límites que no se moverán y el punto verde es el punto que se sustituirá.

2.2. Esta deformación consiste en seleccionar dos vértices o pivotes de manera aleatoria y rotar todos los vértices contenidos entre estos pivotes. El ángulo de rotación es muestreado aleatoriamente entre $[-\phi, \phi]$, donde ϕ es un ángulo que se ajusta después de cada iteración para que la probabilidad de éxito (es decir, aceptar una nueva configuración) de la deformación se mantenga lo más cercano a 50%. La rotación debe ser con respecto al eje que une ambos puntos previamente seleccionados.

El segundo movimiento, es la reptación, similar al movimiento que las serpientes utilizan para desplazarse por el suelo, es una idea más reciente y fue implementado para “dejar que el plásmido libere tensión”. La idea general, es la de tomar dos sub-segmentos de recta diferentes e intercambiarlos, posteriormente, se reacomodan los vértices para que las distancias entre vértices se mantengan constantes. En este trabajo de tesis se optó por una

alternativa a este tipo de movimiento, la cual resultó más eficiente computacionalmente. Como se muestra en la Figura 2.3, esta deformación consiste en tomar un sub-segmento de recta, que este contenido entre dos puntos elegidos aleatoriamente (V_i, V_{i+a}) donde a es un número entero entre tres y ocho (en el ejemplo de la Figura 2.3, $a = 5$), una vez que se tienen estos puntos, se pasará a mover todos los vértices entre estos dos, de tal manera que la forma del segmento original se conserve. Al hacer esto, dos vértices terminaran encimados, y se tendrá un hueco en donde antes estaba el primer vértice, por lo tanto, se debe calcular un nuevo punto entre los vértices i e $i + 1$. Este punto deberá estar a la misma distancia l de cada vértice antes mencionado. La posición espacial que cumple tal condición se encuentra en todos los puntos a lo largo de un aro perpendicular al vector que une ambos vértices. Este punto se elige de manera aleatoria. El último paso es reacomodar los índices, para incorporar este nuevo punto, y eliminar el vértice que estaba encimado. En el ejemplo de la Figura 2.3, a simple vista, parecería que el vértice que se sustituyo fue el $i + 1$, sin embargo, este vértice se conservó en $i + 2$, por lo que el vértice que se eliminó para hacer espacio al nuevo es el $i + 5$, es decir, el que se encimó en el paso 2.

Después de aplicar cada movimiento de deformación, se calcula la energía elástica total de la configuración mediante la ecuación:

$$E = E_b + E_t + E_p \quad (2.1)$$

donde E_b es la energía de doblamiento, E_t es la energía torsional y E_p es la energía potencial. La energía potencial considera, si existen, las interacciones eléctricas con los átomos del medio y fuerza gravitacional. En este trabajo solo se utilizó un potencial gravitatorio si algún punto está por debajo del plano $z = 0$.

La energía de doblamiento, causada por el estrés de la conformación, se mide a partir de la diferencia en el ángulo que forman dos segmentos contiguos de recta, una conformación sin estrés sería una línea recta (i.e. el ángulo que forman dos rectas contiguas es cero), se calcula de la siguiente manera:

$$E_b = K_B T \alpha \sum_{i=1}^N \Theta_i^2 \quad (2.2)$$

donde K_B es la constante de Boltzmann ($1,38064850 \times 10^{-23} J/K$), T es la temperatura del plásmido, en este caso, una constante igual a 293 K, α es la constante de doblamiento con un valor de 2.403 y sin unidades, Θ_i la diferencia de ángulos entre los segmentos $(i - 1$ e $i)$ y $(i, i + 1)$ en radianes.

La energía torsional, causada por el estrés individual de cada segmento de recta de la conformación, está dado por:

$$E_t = \left(\frac{2\pi C}{L}\right)(\Delta L_K - W_r) \quad (2.3)$$

En esta ecuación, L es la longitud total de la configuración, es decir, $L = N * l$. C es la rigidez torsional con un valor de $3 \times 10^{-28} Jm$, ΔL_K es la diferencia del número de enlace del ADN cerrado, se calcula como:

$$\Delta L_K = \frac{\sigma L}{Z_0} \quad (2.4)$$

donde σ es la densidad de súper hélice, con un valor de $-0,05a - 0,07$, muchos de los autores mencionados a lo largo de esta sección utilizan $-0,06$. Z_0 es la altura de la hélice de ADN, considerando una conformación de ADN relajado, este valor es de 3.375 nm. W_r es el número de torcimiento, este factor indica que tan torcida esta la configuración y tiene múltiples formas de calcularse, en este trabajo se utilizó el método 2b de Klenin y Langowski [30] el cual se explica a continuación.

Matemáticamente, el número de torcimiento de una cadena es una doble integral:

$$W_r = \frac{1}{4\pi} \int_C \int_C \frac{(dr_2 \times dr_1) r_{12}}{r_{12}^3} \quad (2.5)$$

donde C es una curva, r_1 y r_2 son los puntos sobre esta curva. Esta ecuación puede resolverse analíticamente, aunque esto no es la mejor opción en cuanto a coste computacional. La solución de esta integral se utiliza en el método 1a del artículo de Klenin, sin embargo, como se mencionó anteriormente, es el método 2b el que se utiliza en este trabajo. Para esto, primero se separa el número de torcimiento en dos cantidades independientes

de la siguiente forma:

$$W_r = W_{r_z} - T_w \quad (2.6)$$

donde W_{r_z} es el torcimiento direccional con respecto al plano $z = 0$ y T_w es el giro total de la conformación. El cálculo de estas variables se explica a continuación:

$$T_w = \frac{1}{2\pi} \sum_{i=1}^N [\cos^{-1}(a_{i-1}p_i) - \cos^{-1}(p_i a_i)] \text{sign}(p_i)_z \quad (2.7)$$

donde:

$$a_i = \frac{k \times s_i}{|k \times s_i|} \quad (2.8)$$

$$p_i = \frac{s_{i-1} \times s_i}{|s_{i-1} \times s_i|} \quad (2.9)$$

En este caso k es el vector unitario en el eje z y s_i es el vector que une los vértices V_i y V_{i+1} .

$$W_{r_z} = \sum_{i=2}^N \sum_{j < i} w_{ij} \quad (2.10)$$

donde:

$$w_{ij} = \text{sign}((s_j \times s_i)(r_j - r_i)) \quad (2.11)$$

La ecuación 11 es cierta solo si los segmentos i y j de la conformación se cruzan en el plano XY, de no ser así, w_{ij} es igual a cero (r_i son los vértices de la conformación).

A partir de estos cálculos se puede obtener la energía elástica total de una conformación. La conformación final se puede obtener a través del algoritmo de Metrópolis al alcanzar el estado de equilibrio del sistema. El algoritmo de Metrópolis en este contexto inicia con una deformación paulatina, en un proceso iterativo, de la cadena de vértices del plásmido. Después de cada deformación, se calcula la energía total de la conformación. Si se cumplen ciertas condiciones con respecto a la energía de una nueva conformación, ésta se acepta, en caso contrario, ésta se rechaza y se vuelve a intentar. Las condiciones para aceptar una nueva configuración son las siguientes:

Se acepta la configuración si:

- La nueva energía es menor que la energía de la configuración anterior: $E_{i+1} < E_i$.
- En caso contrario se acepta si: $\exp((E_i - E_{i+1})/K_B T_M) > \rho$.

donde ρ es un número aleatorio uniformemente distribuido entre $[0, 1]$ y T_M es la temperatura Monte Carlo, es decir, la temperatura que puede variar para obtener diferentes configuraciones.

Además de esta condición, debido a que estamos hablando de una conformación que está relacionada con una cadena de ADN súper enrollado, se tienen que tomar las siguientes consideraciones:

1. La distancia entre dos segmentos de recta debe ser mayor a $d = 3,5nm$, para evitar que dos segmentos de ADN se encimen o que tengan interacciones electromagnéticas.
2. Se debe hacer un chequeo de nudos, utilizando el polinomio de Alexander [2] para evitar que la configuración pase de súper enrollada a nudo.

Para el chequeo de nudos, se utilizó el programa de Harris y Harvey [23]. Este código está basado en el método de J. W. Alexander, quien utiliza un polinomio (ahora llamado polinomio de Alexander) para determinar la configuración de una curva. El polinomio de Alexander analiza una configuración en tres dimensiones usando una matriz de $K \times K + 2$. Donde K es el número de cruces de un segmento con otro. Cada elemento de dicha matriz representa un tipo de cruce, ya sea izquierdo o derecho. Dependiendo del número del cruce (previamente etiquetado) y la región del nudo, se va llenando esta matriz [2] [24]. Al final se eliminan dos columnas contiguas cualquiera (generalmente las últimas dos) y se calcula el determinante de esta misma, obteniendo así el polinomio de Alexander, la diferencia entre los distintos polinomios de Alexander que pueden resultar al eliminar diferentes columnas difieren de los otros posibles en más o menos potencias del grado del polinomio (múltiplos de la variable independiente, usualmente $\pm t$), como demuestra Alexander [2]. Al final se reduce el polinomio, se considera un nudo si el polinomio de Alexander es diferente a 1, si es igual a 1 se dice que no es un nudo. Esta explicación omite detalles importantes de la teoría de nudos por lo que se invita al lector a revisar las referencias para ver el método completo [2] [23] [24].

Un plásmido toma alrededor de 10,000,000 (diez millones) de iteraciones para formarse. El criterio típico que dicta el fin de la simulación es cuando la energía alcanza el equilibrio térmico; aunque hay autores que utilizan también el número de torcimiento. En este trabajo, sin embargo, se utiliza el método de Bryant y Fujimoto [22]. En este método se utiliza una cantidad llamada radio de giramiento, que está dado por:

$$R_g = \left(\sum_{i=1}^N \frac{R_i^2}{N} \right)^{1/2} \quad (2.12)$$

donde R_i es la distancia del centro de masas al vértice i de la configuración actual. De acuerdo con Fujimoto, esta cantidad es la última en entrar en equilibrio, por lo que usarla garantiza, con mayor probabilidad, un plásmido completamente terminado.

El chequeo de equilibrio se realiza de la siguiente forma, se calcula R_g cada M millones de iteraciones, si la diferencia entre dos chequeos es menor a la desviación estándar calculada en ese mismo intervalo, entonces se acepta el equilibrio y se entrega el plásmido terminado, de otra manera, se realizan nuevamente M millones de iteraciones. M esta entre 4 y 6 millones de iteraciones. El algoritmo completo para la generación de un plásmido se puede ver en la Figura 2.4 donde se muestra el diagrama de flujo para el programa de generación de plásmidos. El usuario puede optar por dar sus propios parámetros o usar los parámetros por defecto del plásmido. Los parámetros por defecto generan un plásmido pBR322 a 293K y $\sigma = -0,06$, alternativamente los valores de z_0 y d se pueden cambiar.

2.1.2. El Formato PDB para Generar un Modelo Atómico para Plásmidos de ADN

El Protein Data Bank (PDB, en español banco de datos de proteínas) es una recopilación de información sobre la estructura tridimensional de las proteínas y ácidos nucleicos quienes, además de proporcionar un portal para navegar los archivos tanto propios como por los investigadores de dichas áreas, también creo su propio formato para unificar toda la información en sus archivos y hoy en día, gracias a su extenso uso, se tienen herramientas de visualización en 3D de dichos formatos. Los archivos PDB no solo deben contener información espacial de las proteínas sino también, información extra relevante (tempe-

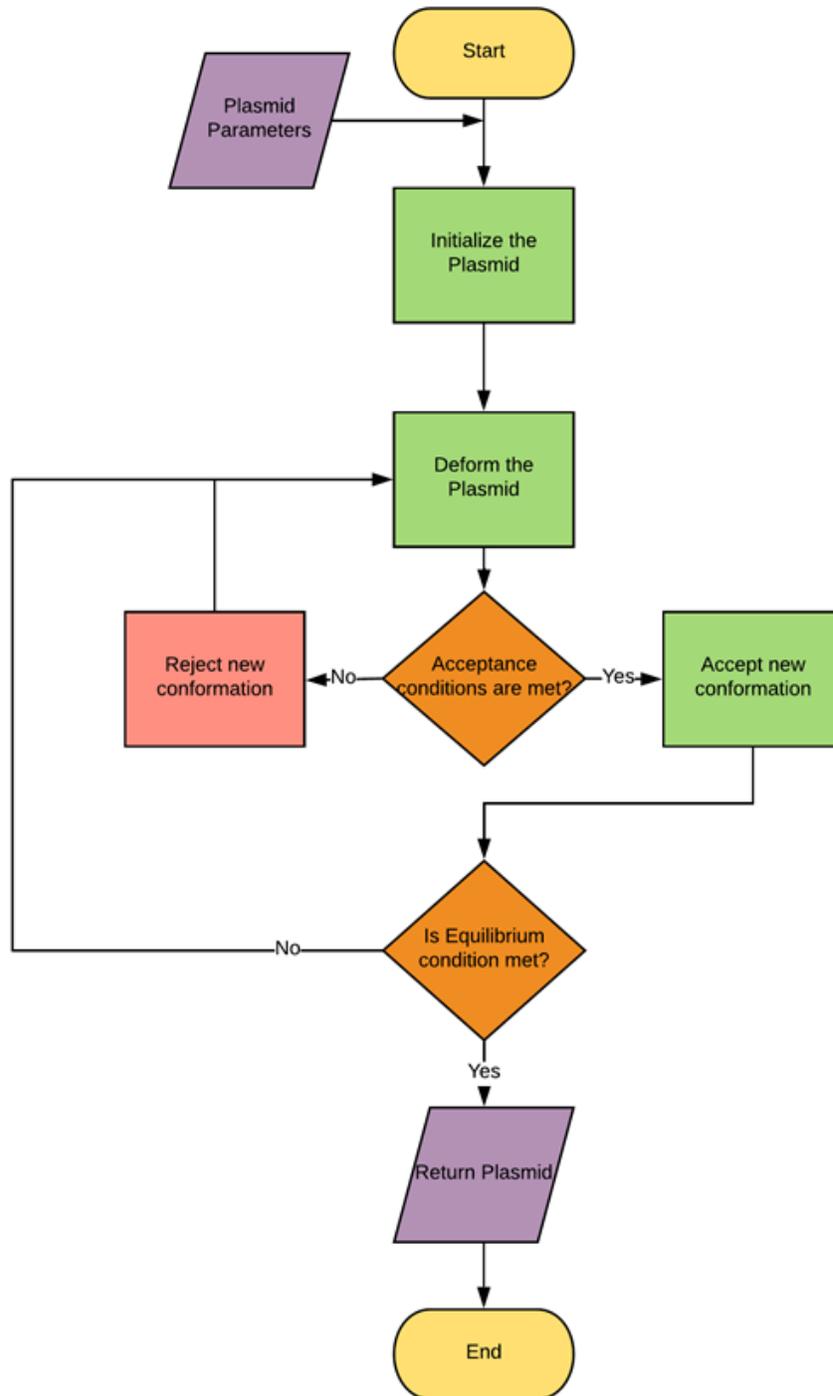


Figura 2.4: Diagrama de flujo del algoritmo de generación de plásmidos.

ratura a la cual se tomaron los datos, métodos de medición, etc.) por lo que la lectura de estos archivos para análisis suele ser llevada a cabo utilizando alguna herramienta que permita leer directamente estos formatos sin recurrir a la programación directa y desarrollo de herramientas propias (lo cual, es posible y no tan tedioso como podría sonar). Geant4-DNA en su ejemplo “pdb4dna” muestra como cargar un plásmido de ADN a partir del conocimiento de las posiciones atómicas para los carbonos, hidrógenos, fósforos, etc, que componen dicho ADN.

2.1.3. Clasificación de Rompimientos de ADN

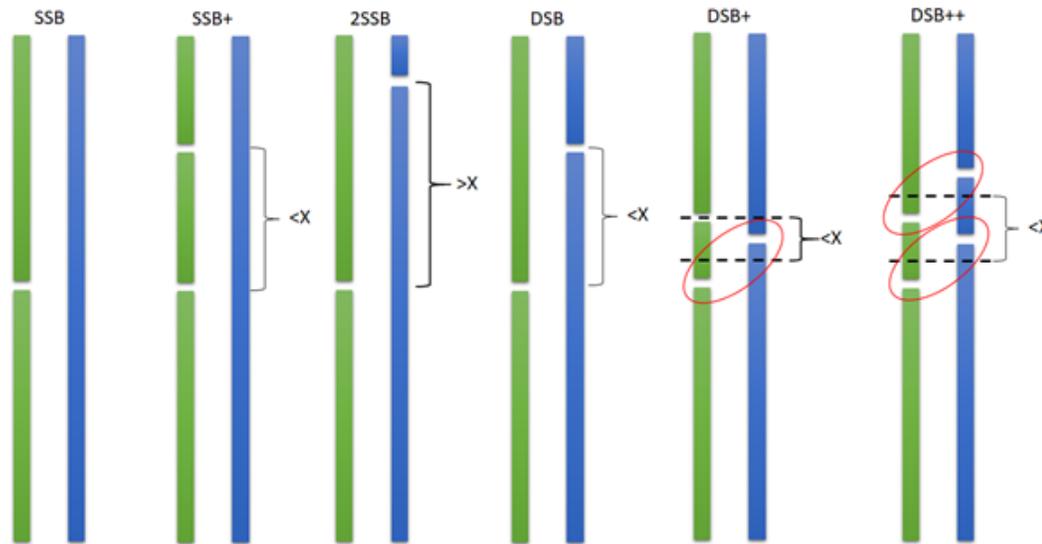


Figura 2.5: Clasificación de rompimientos utilizada en este trabajo, basada en los trabajos de Charlton y Mc Namara, X es una distancia de 10 bps (3.4 nm), los círculos rojos encierran un DSB. Los SSB+ ocurren cuando ocurren dos rompimientos simples a una distancia menor de 10 bps. Los 2SSB ocurren cuando ocurren dos rompimientos en diferentes hebras, pero a una distancia mayor a 10 bps. Los DSB ocurren cuando hay dos rompimientos en diferentes hebras a una distancia menor de 10 bps. DSB+ son tres rompimientos, donde al menos dos de ellos forman un DSB y los DSB++ son dos rompimientos dobles a una distancia menor a 10 bps desde el centro de los DSBs.

Un rompimiento o daño simple de ADN (*single strand break*, SSB) causado por un evento físico ocurre cuando la energía depositada acumulada en la región ocupada por un azúcar o fosfato excede 17.5 eV [41]. Algunos autores también utilizan una distribución lineal de 5 eV a 37.5 eV (Werner Friedland, 2017) para definir un SSB. Por otro lado, un rompimiento o daño causado por una especie química es atribuida a la reacción

de un nucleótido con el radical $\bullet OH$ [56] [20]. La clasificación de los daños que se utilizó en el análisis de datos, fue una clasificación originalmente propuesta por Charlton y mencionada por McNamara [11] [38], debido a que las técnicas experimentales actuales para la determinación del número de rompimientos no tiene la resolución suficiente para distinguir rompimientos complejos (SSB+, 2SSB, DSB+, DSB++), solo se utilizaran los rompimientos SSB y DSB.

2.1.4. Códigos Monte Carlo *Track Structure*

En esta sección se describen con mayor detalle los puntos clave que distinguen a los códigos Monte Carlo *Track Structure*, diseñados para cuantificar el transporte de radiación para micro y nano dosimetría, así como su implementación en Geant4-DNA (subsecuente en TOPAS-nBio). Por lo tanto, los procesos físicos descritos en esta sección son solamente eléctricos en el rango de energía relevante para el estudio del daño directo al ADN (hasta 1 MeV para electrones).

Track de la Partícula

La definición del track (trayectoria, trazado, etc.) de una partícula depende de la aplicación a la cual esta enfocada. En modelación Monte Carlo se puede decir que el track de una partícula esta formado por todos los puntos de ionización y excitación tanto de la partícula primaria, como de sus productos secundarios, que se generan conforme la partícula viaja en el medio. Cuando hablamos del track, nos referimos a partículas con la capacidad de ionizar el medio, es decir radiación ionizante. La radiación ionizante es aquella con energía suficiente para extraer electrones de los átomos del medio en el cual está viajando, por ejemplo: rayos-x, rayos- γ , y partículas cargadas (e^- , *protones*, *alfas*, etc). Los rayos-x (con energías que van de unas decenas de eV a unos cuantos cientos de keV) son capaces de producir electrones al interactuar con el medio con una energía E_e cercana a la de los rayos-x incidentes $E_{e^-} = E_{rx} - \phi$, donde ϕ es la energía de enlace del electrón (generalmente de unas decenas de eV). Estos pueden ser producidos por el frenamiento súbito de electrones en materiales de alto número atómico o naturalmente por la recombinación de electrones orbitales. Los rayos- γ (energías de unos cientos de keV a

unos cuantos MeV) pueden producir foto electrones, pero si la energía aumenta también pueden producir electrones libres mediante dispersión Compton y positrones mediante la creación de pares. Por lo tanto, los rayos- γ pueden producir electrones libres en un amplio espectro de energías. Estos son producidos naturalmente por radioisótopos, por ejemplo, ^{60}Co o por interacciones nucleares de partículas altamente energéticas.

Por otro lado, cuando las partículas cargadas interactúan con la materia producen electrones secundarios en un espectro amplio de energías a lo largo de su trayectoria conforme son frenadas en el medio. Por lo tanto, todas las partículas que conforman la radiación ionizante producen los mismos productos (es decir e^-), por lo que la diferencia en daño biológico de diferente radiación ionizante de diferente calidad debe estar relacionada con la distribución espacial de estas interacciones, que depende de la energía y masa de la partícula incidente.

La caracterización del *track* de una partícula y su relación con el daño al ADN es una investigación en desarrollo [50] [13] [58]. Una de las primeras cantidades utilizadas para caracterizar (en energía y distribución espacial) al *track* es el “potencial de frenado”, en particular el “potencial de frenado restrictivo”, o LET (linear energy transfer). El potencial de frenado se define como la pérdida de energía por unidad de longitud conforme la partícula atraviesa el medio y es desacelerada. El potencial de frenado es una cantidad bastante estudiada debido a que es fácilmente obtenible experimentalmente (por ejemplo, a través de la dosis depositada) y teóricamente calculable con precisiones mayores a 99 % [3].

El potencial de frenado contiene suficiente información para cuantificar la respuesta del medio a la radiación en escenarios macroscópicos (arriba de algunos micrómetros), a través de la energía absorbida localmente. De esta manera, el llamado potencial másico de frenado, es decir, el potencial de frenado multiplicado por la densidad del medio se relaciona con la dosis depositada a través del espectro de fluencia. Este método es utilizado por los códigos Monte Carlo de uso general o de transporte condensado para cuantificar las distribuciones de dosis en radioterapia, por ejemplo, los más utilizados: EGS [44] Geant4 [1] MCNP [61] PENELOPE [51]. Sin embargo, en general los códigos de transporte condensado no proveen de la resolución necesaria para cuantificar detalladamente el transporte de radiación en escenarios micro y nanoscópicos, que es de unos cuantos armstrongs.

Los códigos Monte Carlo *track structure*, calculan el transporte de la radiación de una manera interacción por interacción considerando todos los procesos físicos teóricamente posibles, a diferencia de los algoritmos de transporte condensado (que condensan en una sola interacción muchas interacciones individuales de baja energía para reducir significativamente el tiempo de cómputo), por lo que son capaces de simular daño biológico a nivel molecular [10].

Para poder desarrollar un modelo del *track* usando Monte Carlo *track structure* primero es necesario conocer las secciones eficaces para los procesos físicos relevantes a estas escalas, para las partículas de interés. En este trabajo se utilizó TOPAS-nBio [38] [50] desarrollado en TOPAS [47], que facilita el desarrollo de simulaciones *track structure* aprovechando toda la física provista por Geant4-DNA [25].

La Física de Geant4-DNA

Con todo lo que se ha mencionado anteriormente, se puede entonces comenzar a explicar cómo solventa estos problemas Geant4-DNA. Geant4-DNA calcula las secciones eficaces de ionización y excitación usando la primera aproximación de onda plana de Born (FBA), siendo ampliamente utilizada para calcular procesos inelásticos. Debido a que el medio de interés es tejido biológico, y ya que las células se componen en su mayoría (> 70 %) de agua entonces se utiliza agua como primera aproximación. Por lo tanto, las secciones eficaces que se necesitan son las del agua. Considerando a la molécula del agua, se consideran cinco estados de excitación y cinco de ionización [17]. Se utiliza la función de respuesta del dieléctrico (DRF) de la molécula para calcular las secciones eficaces de las partículas incidentes. Utilizando la FBA, las transferencias de energía y momento se relacionan con la función de pérdida de energía (ELF) a través de $Im[-1/\varepsilon(E, q)]$, donde ε es la DRF compleja que caracteriza a la molécula objetivo. Entonces, el inverso del camino medio de la partícula (IMFP) está dado la doble diferencial de Born en términos de la pérdida de energía E y la transferencia de momento q por:

$$\frac{d^2\Sigma(T, E, q)}{dEdq} = \frac{1}{\pi\alpha_0 Tq} Im \left[\frac{-1}{\varepsilon(E, q)} \right] \theta[q - q_-(E, \tau)]\theta[q_+(E, \tau - q)]\theta[\tau - E] \quad (2.13)$$

Donde α_0 es el radio de Bohr, τ es la energía cinética de la partícula, $T = (m|M)\tau$

es la energía cinética de un electrón viajando con la misma velocidad de la partícula considerada, m es la masa del electrón, M es la masa de la partícula incidente, θ es la función escalonada y $\varepsilon = \varepsilon_1 + i\varepsilon_2$ representa la función compleja del dieléctrico del material objetivo. Los límites de transferencia de momento van como sigue:

$$q_{\pm} = \sqrt{2M}(\sqrt{\tau} \pm \sqrt{\tau - E}) \quad (2.14)$$

Así, la sección eficaz diferencial simple y la IMFP integrada se pueden obtener con las dos expresiones simples siguientes:

$$\frac{\Sigma(T, E)}{dE} = \int \frac{d^2\Sigma(T, E, q)}{dEdq} dq \quad (2.15)$$

$$\Sigma(T) = \int dE \int \frac{d^2\Sigma(T, E, q)}{dEdq} dq \quad (2.16)$$

El punto más importante es el cálculo de la ELF del medio. En Geant4-DNA, el primer paso consiste en tomar datos experimentales del DRF para hacer un ajuste en el límite óptico ($q = 0$) en una serie de ecuaciones de Drude, después se aplica el modelo de dispersión para el rango de transferencia de momento ($q > 0$).

Para bajas energías, cuando la velocidad de la partícula incidente se acerca a la de los electrones orbitando la partícula objetivo (< 1 keV para electrones y < 300 keV para protones), la FBA ya no es válida. En este caso, las secciones eficaces de los protones se calculan usando una combinación de modelos semi-empíricos como la fórmula de Rudd de ionización y las fórmulas de Miller y Green para la excitación. Para la ionización del electrón, la FBA se corrige usando el termino de intercambio propuesto en el reporte número 37 del ICRU [4] y una corrección simple del campo de Coulomb el cual contabiliza la energía potencial ganada por el electrón en el campo de la molécula objetivo. La sección eficaz para ionización a una energía dada T , es entonces calculada para un valor aumentado $T' = T + B_j + U_j$, donde B_j y U_j son la j -ésima energía de enlace de la capa y la energía cinética promedio del electrón en esta misma capa, respectivamente. Para la excitación $T' = T + 2E_j$, donde E_j es la j -ésima energía de excitación. A bajas energías los procesos de interacción para los electrones que dominan son los elásticos, donde el electrón es

deflectado sin perder energía varias veces hasta que ocurre un proceso inelástico. Estas colisiones determinan la densidad de distribución espacial de la nube de ionizaciones y excitaciones y son descritas por dos modelos alternativos en Geant4-DNA: el modelo de apantallamiento de Rutherford y los modelos de Champion [9].

En teoría, los electrones con energía cinética menor a 8 eV no pueden ionizar a la molécula de agua dado que esta energía es menor al del orbital más externo del blanco, por lo que a estos electrones se les llama electrones sub-excitados. Estos electrones pueden producir excitaciones vibracionales y rotacionales además de colisiones elásticas hasta llegar a la termalización completa (a 2.5 meV). Al no disponer de una teoría que describa tal fenómeno, en Geant4-DNA se utilizan las secciones eficaces publicadas por Michaud [43] para moléculas de hielo utilizando un ajuste para el cambio de fase. Para los electrones con energías mayores a 10 keV, se consideran efectos relativistas y se consideran las interacciones longitudinales y las trasversales. La sección eficaz total de ionización para electrones de 10 keV a 1 MeV está dada por:

$$\Sigma_j = \Sigma_j^L + \Sigma_j^T \quad (2.17)$$

Donde

$$\Sigma_j^L = \frac{2}{\pi\alpha_0\beta^2(T)mc^2} \left(\int_{E_{min}}^{E_{max}} dE \int_{k_{min}}^{k_{max}} Im \left[-\frac{1}{\varepsilon(E, k)} \right]_j \frac{dk}{k} \right) \quad (2.18)$$

Y

$$\Sigma_j^T = \frac{1}{\pi\alpha_0\beta^2(T)mc^2} \left(\int_{E_{min}}^{E_{max}} Im \left[-\frac{1}{\varepsilon(E, 0)} \right] dE \right) \times \left[\ln \left(\frac{1}{1 - \beta^2(T)} \right) - \beta^2(T) \right] \quad (2.19)$$

Los límites de transferencia de momento son:

$$k_{max,min} = (ch)^{-1}(\sqrt{T(T + 2mc^2)}) \pm \sqrt{(T - E)(T - E + 2mc^2)} \quad (2.20)$$

$$E_{max} = (T + B_j)/2 \quad (2.21)$$

Para las partículas alfas e iones más pesados como el carbón o el oxígeno, se utiliza un factor de re-escalamiento de velocidad asumiendo que las dos partículas tienen la misma velocidad y la misma sección eficaz multiplicada por un término de carga efectiva. La carga efectiva y la transferencia de electrones para partículas alfa fueron modelados a partir del método de Dingfelder.

$$\frac{d\sigma_{ion}}{dEdq}(v) = Z_{eff}^2(E) \frac{d\sigma_{proton}}{dEdq}(v) \quad (2.22)$$

Donde v es la velocidad del proyectil y Z_{eff}^2 es la carga efectiva al cuadrado del ion incidente.

El uso de la carga efectiva toma en cuenta el efecto de apantallamiento de la carga nuclear causada por los electrones enlazados de la partícula incidente. Depende de la transferencia de energía durante la colisión.

Para iones más pesados que los alfas, es complicado calcular las secciones eficaces para cada estado de carga de la partícula. Por lo que una carga efectiva global, incluyendo el efecto de cambio de carga de Booth y Grant se aplica, dada por la siguiente expresión:

$$\frac{Z_{eff}}{Z} = 1 - \exp(-1,316x + 0,112x^2 - 0,0650x^3) \quad (2.23)$$

Donde

$$x = 100\beta Z^{-2/3} \quad (2.24)$$

A energías de $\sim 8 - 13$ eV también puede ocurrir acoplamiento electrónico, en este caso se utilizan los resultados experimentales de Melton de 1972. Para protones, además de las ionizaciones y excitaciones, los procesos de transferencia de carga ocurren cuando los protones incidentes capturan un electrón y se vuelven átomos neutrales de hidrógeno al ionizar una molécula de agua. El hidrógeno puede producir ionizaciones en el medio y también sobrellevar un proceso de pérdida, perdiendo su electrón orbital y regresando a su estado iónico. En este caso el electrón es eyectado en la dirección del proyectil con la misma velocidad. Los procesos de transferencia y pérdida de energía son estudiados usando formulas semi empíricas, y los átomos de hidrogeno son dispersados usando el modelo de Rudd con un término para contabilizar el efecto de apantallamiento causado

por el electrón enlazado del proyectil.

TOPAS y TOPAS-nBio

El kit de herramientas de código libre de uso general Geant4 [1], desarrollado para física de altas energías, es frecuentemente usado para simulaciones de radioterapia, sin embargo, su flexibilidad de uso y diseño modular basado en objetos lo hace muy complicado para usar. Por esta razón se han desarrollado varias plataformas en Geant4 para facilitar su uso principalmente en radioterapia, como GAMOS (Geant4-based Architecture for Medicine-Oriented Simulations, Arce et al., 2018) y GATE (Geant4 Application for Tomographic Emission, Jan et al., 2011) y TOPAS (TOol for PArticle Simulation [47]). De estos TOPAS sobre sale al ser un software para el modelado radiobiológico, multithreaded y completamente 4D. A pesar de que TOPAS fue originalmente diseñado para las aplicaciones de protón terapia, ha sido extendido a muchas otras aplicaciones de radio terapia. Algunos códigos Monte Carlo *track structure*, son PARTRAC, RITRACKS, KURBUC, y el mismo Geant4-DNA [6] [19]. Sin embargo, estos están depreciados, son comerciales, requieren de una agresiva curva de aprendizaje o están restringidos a los laboratorios o investigadores que los desarrollaron. Esto motivó el desarrollo de TOPAS-nBio [38] [50], como una nueva extensión de TOPAS que envuelve y extiende a Geant4-DNA para aplicaciones de radiobiología, nanodosimetría, microdosimetría y radioquímica.

2.2. Simulación de la Etapa Química Heterogénea de la Radiólisis del Agua

La difusión y reacción de especies químicas, es un proceso independiente al transporte de las partículas físicas cargadas. Mientras que los procesos físicos de transporte de partículas pueden considerarse atemporales, debido a que el tiempo de interacción es instantáneo ($< 10^{-15}$ s) comparado con la etapa química ($10^{-12} - 10^{-6}$ s), en la etapa química el número de especies químicas varían con respecto al tiempo. Así, para la simulación del proceso de radiólisis del agua, varias etapas para cada tipo de proceso dependiendo el tiempo de simulación son de la siguiente forma:

- Etapa Física $< 10^{-15} s$
- Etapa Fisicoquímica: $10^{-15} - 10^{-12} s$
- Etapa Química: $10^{-12} - 10^{-6} s$

En la etapa física se lleva a cabo el depósito de energía por eventos de ionización y excitación y se continúa el transporte de electrones sub-excitados. En la etapa fisicoquímica, las moléculas excitadas e ionizadas de agua, resultado del proceso de interacción de la radiación, se disocian generando las primeras especies químicas, así como los electrones sub-excitados se termalizan y son capturados por un arreglo de moléculas de agua para generar electrones solvados. Subsecuentemente, en la etapa química, las primeras especies de propagan en el medio por difusión reaccionando entre sí, reduciendo el número de especies primarias y creando productos secundarios hasta alcanzar el equilibrio. Más allá de la etapa química está la etapa biológica, en esta etapa se ponen en marcha los mecanismos de reparación del ADN, efectos secundarios en organismos vivos etcétera; dura de $10^{-6} s$ hasta horas-años.

2.2.1. Método Paso a Paso

Geant4 tiene la capacidad de simular la difusión y reacción de especies químicas hasta $1 \mu s$ asumiendo reacciones totalmente controladas por difusión, a través de su extensión Geant4-DNA [29] [28]. El método de transporte es Browniano paso-a-paso, así que se requieren los coeficientes de difusión D (en unidades de $m^2 s^{-1}$) de las especies químicas involucradas. El medio se asume homogéneo y compuesto por agua a temperatura ambiente. La distancia promedio entre cada “paso” después de un cierto tiempo Δt , está dada por $(6D\Delta t)^{1/2}$ en una dirección completamente aleatoria. Después de cada “paso”, se verifica si las especies potencialmente reactivas están a una distancia menor que su radio combinado de reacción a (ecuaciones (2.25) y (2.26)), si esto es satisfecho, entonces se realiza la reacción química correspondiente eliminando a ambas especies químicas del sistema y colocando los productos, se existen, alrededor del punto de reacción [32]. El cálculo del radio de reacción a se lleva a cabo con la ecuación siguiente si las especies químicas involucradas en la reacción son diferentes:

$$k = 4\pi a D' \tag{2.25}$$

o con la ecuación siguiente si las dos especies químicas son iguales:

$$k = 2\pi aD' \quad (2.26)$$

Donde k es la tasa de reacción (en unidades $M^{-1}s^{-1}$, donde $M = mol/dm^3$), y D' es el coeficiente de difusión para la difusión relativa entre las dos especies químicas.

2.2.2. Tiempos de Reacción Independiente

El método paso-a-paso es, sin embargo, muy costoso computacionalmente, por ejemplo, una simulación de 240 historias con tracks cortos ($100 \mu m$) de protones de 300 MeV tarda alrededor de 2 semanas y media en un solo CPU a 2.8 GHz. Una solución alternativa y estocástica es la de tiempos de reacción independiente (IRT por sus siglas en inglés) propuesto hace más de treinta años por Clifford y Green [12] y ampliamente validado, ver por ejemplo [36] [21] [40] [48]. El método IRT no lleva a cabo la difusión de las especies químicas hasta su reacción, sino que calcula el tiempo en el que las especies reaccionarán a partir de la distribución de probabilidad dada por la ecuación (para reacciones totalmente controladas por difusión):

$$W(x, a, t) = \frac{a_{eff}}{x_{eff}} \operatorname{erfc} \left(\frac{x_{eff} - a_{eff}}{\sqrt{4D't}} \right) \quad (2.27)$$

Donde a_{eff} es el radio efectivo y x_{eff} es la distancia de separación inicial efectiva. a_{eff} y x_{eff} tienen correcciones por el efecto de carga de las especies químicas dado por

$$a_{eff} = \frac{-r_c}{1 - \exp(r_c/a)} \quad (2.28)$$

$$x_{eff} = \frac{-r_c}{1 - \exp(r_c/x)} \quad (2.29)$$

o simplemente $a_{eff} = a$ y $x_{eff} = x$ para especies químicas neutras. La cantidad r_c es el radio de Onsager, definido por:

$$r_c = \frac{Z_1 Z_2 e^2}{4\pi\epsilon_0\epsilon_r K_B T} \quad (2.30)$$

donde Z_1e y Z_2e son las cargas iónicas, ϵ_0 es la permitividad del espacio libre, ϵ_r es la

permitividad relativa del medio, K_B la constante de Boltzmann y T la temperatura.

A partir de la ecuación (2.27), el tiempo de reacción se obtiene después de resolver para t , asignando un número aleatorio ξ uniformemente distribuido a la probabilidad W , dado de la siguiente forma:

$$t = \frac{1}{4D} \left[\frac{x_{eff} - a_{eff}}{erfc^{-1}(\xi x_{eff}/a_{eff})} \right]^2, (\xi < a_{eff}/x_{eff}) \quad (2.31)$$

$$t = \infty, (\xi \geq a_{eff}/x_{eff}) \quad (2.32)$$

Donde $erfc^{-1}(x)$ es el inverso de la función de error. Si el tiempo de reacción t es pequeño, entonces la probabilidad de reacción es mayor.

El método IRT comienza con la información de las posiciones y tipo de especies químicas generadas al final de la etapa fisicoquímica. Para todas las combinaciones de pares de especies químicas potencialmente reactivas, y asumiendo que cada par se encuentra en aislamiento total del resto de especies (de aquí el nombre “independiente”), se muestrea el tiempo de reacción t usando la ecuación (2.31). Estos valores de t se vierten en un arreglo que es reordenado de forma temporal ascendente. Con el arreglo reordenado, la realización de las reacciones se realiza comenzando con el par de especies que obtuvieron el tiempo más corto de reacción, quitando el par de especies del arreglo (y los pares en que ambas especies estén involucrados en el resto del arreglo) y colocando los productos correspondientes al tipo de reacción. Si el producto es reactivo, entonces los respectivos tiempos de reacción con el resto de las especies químicas del arreglo son muestreados, los tiempos son incluidos en el arreglo general y el arreglo es reordenado nuevamente. Este proceso iterativo continua hasta que hayan agotado todas las combinaciones posibles o si un tiempo de corte (especificado por el usuario) es alcanzado. Esta técnica está implementada en TOPAS-nBio considerando 72 reacciones clasificadas en totalmente controladas por difusión, parcialmente controladas por difusión y de primer orden. Mayores detalles pueden encontrarse en [52].

2.2.3. Uso de Disolventes Químicos

Los disolventes químicos o *scavengers* son sustancias cuyo propósito es el de eliminar una cierta especie química, al hacerla reaccionar inmediatamente después de su creación. Los químicos, biólogos y farmacéuticos utilizan los disolventes químicos para aislar ciertas reacciones químicas que quieran estudiar al eliminar reactivos no deseados que aporten productos a la solución, esto además facilita el análisis del conteo de especies químicas de interés.

La simulación explícita de las moléculas de los disolventes es impráctica computacionalmente, debido al gran número de especies químicas uniformemente distribuidas en la solución. Para un sistema de M solventes y N especies químicas, la complejidad del algoritmo para encontrar especies químicas vecinas es $O(N^2M)$. Entonces, la solución alternativa es considerar que la solución de disolvente está uniformemente (continuamente) distribuida y que la reacción con las especies químicas no altera significativamente la concentración de disolvente. Por ejemplo, utilizando la siguiente ecuación:

$$W_{dist}(t) = 1 - \exp(-k[B]t) \quad (2.33)$$

donde k es la tasa de reacción, $[B]$ es la concentración del disolvente, generalmente en unidades de $mol\ dm^{-3}$ (también es común verlo como M) y t es el tiempo. En particular la cantidad combinada $k[B]$ es conocida como *scavenging capacity* con unidades s^{-1} . En el método IRT, a las especies químicas reactivas con el disolvente también se muestrean tiempos t de la ecuación (2.33) que se incluyen en el arreglo general.

Los disolventes químicos usados en este trabajo se muestran en el cuadro 2.1. Finalmente, la lista completa de todas las reacciones químicas utilizadas en TOPAS-nBio para la radiólisis del agua y sus respectivas tasas de reacciones y constantes de difusión se pueden encontrar en el Apéndice II de este trabajo, cuya validación fue reportada en [50].

<i>Scavenger</i>	Especie con la que reacciona	Tasas de Reacción ($M^{-1}s^{-1}$)	Concentración (mM)
TRIS	$\bullet OH$	$1,5 \times 10^9$	1
EDTA	$\bullet OH$	$0,1 \times 10^9$	0.1
O^2	e_{aq}^-	$1,74 \times 10^{10}$	1.3
O^2	H^\bullet	$2,1 \times 10^{10}$	1.3

Cuadro 2.1: Datos de los disolventes químicos utilizados en las simulaciones 2 y 3 de este trabajo, los productos de estas reacciones no continúan reaccionando por lo que son removidos por completo de la simulación. M es igual a mol/dm^3 .

2.3. Interacción Forzada de Rayos γ

El camino libre de los rayos gamma de ^{60}Co (~ 1.25 MeV) en agua es 15.82 cm. Ya que para el estudio Monte Carlo del daño en ADN usualmente se utilizan volúmenes con una escala de nanómetros, la probabilidad de interactuar a estas escalas es muy baja. Sin embargo, existen dos métodos para simular la interacción de rayos gamma de ^{60}Co a estas escalas. El primero es obtener el espectro de inicial de energía de los electrones secundarios puestos en movimiento por los rayos γ y utilizarlo para crear una fuente de electrones cuya energía es muestreada de este espectro. El segundo método es utilizar técnicas de reducción de varianza, en particular la llamada interacción forzada [5]. En esta técnica el camino libre de la partícula es re-escalado para incrementar la probabilidad de interacción en regiones de grosor mucho menor al camino libre. Para evitar sesgar el resultado, el peso estadístico de cada partícula es reajustado adecuadamente, para mayor detalle ver la referencia [5]. Este método ha sido recientemente implementado en TOPAS y se utilizó en este trabajo.

Capítulo 3

Metodología

3.1. Plásmidos Super Enrollados y Estructura del Capítulo

El software desarrollado en esta tesis para la simulación de plásmidos genera estructuras súper enrolladas a partir de parámetros simples, pero con determinadas condiciones iniciales. El usuario debe proporcionar los siguientes parámetros en orden:

- Semilla aleatoria: La semilla para el generador de números aleatorios (RANECU).
- Temperatura Final: En Kelvin.
- Temperatura Inicial: En Kelvin.
- Altura de la Hélice: Valor 0 es por defecto una altura de 3.375 nm.
- Densidad de la súper hélice σ : Valor 0 es por defecto $-0,06$.
- Número de Segmentos: Dependerá del plásmido que se genere, 10.2 nm es la longitud de cada uno.
- Tipo de Plásmido: Solo una etiqueta, el archivo de salida final usara esta variable.
- Diámetro Efectivo del ADN: Valor 0 es por defecto 3.5 nm.

Si el usuario no coloca todos los parámetros, el programa marcará un error y le indicará los parámetros en este orden por lo que no es necesario recordarlo. El plásmido simulado fue el pBR322 cuya longitud es de 4362 bases par a temperatura ambiente o 145 segmentos de recta, con una longitud de 10.2 nm para cada segmento recto, esto correspondiendo a 30 bps (teniendo en total 4350 bps como aproximación). La temperatura de los plásmidos usados en la tercera etapa, se mantuvo constante a 293 K, sin embargo, en este trabajo se llevaron a cabo plásmidos con un amplio rango de temperaturas que van desde los 450 K hasta los 10 K. La altura de la hélice fue de 3.375 nm, la densidad de la súper hélice como -0.06 y el diámetro efectivo como 3.5 nm. Por último, el chequeo por equilibrio se realizó cada 4 millones de iteraciones utilizando el radio de giramiento como cantidad a evaluar. Se reportan las métricas evolución de energía y torcimiento en función del número de iteraciones, además del tiempo de simulación.

Se desarrollaron dos tipos de algoritmos para la disminución de la temperatura de los plásmidos, uno lineal y otro exponencial, el cual funcionan de manera similar. Se toma la energía elástica total de la conformación actual y se compara con la de la siguiente iteración exitosa, si esta energía es menor a la de la nueva conformación, entonces se disminuye la temperatura un cierto valor ΔT , se repite esto mientras no se haya alcanzado la temperatura deseada. El valor de ΔT esta determinado por el algoritmo específico a usar, en el algoritmo lineal, esta nueva temperatura toma la forma de $T_{nueva} = T_{anterior} - \Delta T$, donde ΔT es un número muy pequeño (0.00001 K). En el algoritmo exponencial la nueva temperatura toma la forma de $T_{nueva} = (T_{inicial} - T_{final})exp(-\Delta T) + T_{final}$, en este caso $\Delta T = \Delta T + 1 \times 10^{-6}$.

La siguiente parte del trabajo constó de simulaciones de irradiaciones. Esta parte constó de tres etapas, cada una se describe una simulación en una sección diferente que se dividió en geometría, irradiación y detectores. Como software base se utilizó una versión de desarrollo de TOPAS/TOPAS-nBio versión 3, compilado con Geant4 v10.04 parche 2.

La primera etapa se enfocó en la simulación del daño directo a las hebras de ADN, con el fin de demostrar la necesidad de desarrollar modelos geométricos más detallados de ADN. Se realizó una comparación del modelo de ADN de Charlton de tres secciones cilíndricas contra el modelo basado en el trabajo de Meylan [42] de seis volúmenes semiesféricos.

En la segunda etapa del trabajo, se realizó la simulación de la etapa química de la radiólisis del agua en TOPAS-nBio. Se utilizó una técnica de reducción de varianza que permite el uso e interacción de fotones de ^{60}Co en volúmenes cuya dimensión es mucho menor al camino libre de esos fotones. La figura de mérito fue el valor G definido como el número de especies químicas producidas por cada 100 eV de energía depositada. Los siguientes escenarios fueron considerados y comparados con datos experimentales: a) tracks cortos (100 μm) de protones de 300 MeV, b) fotones de ^{60}Co , c) ^{60}Co en presencia de *scavengers*.

Finalmente, en la última etapa se realizó una simulación resultado de una combinación de las etapas anteriores. Se utilizó el modelo geométrico del ADN de seis volúmenes combinado con el modelo atómico de la primera etapa y la simulación de la química heterogénea en TOPAS-nBio de la segunda etapa. Por tanto, se combinaron los tracks físicos, químicos y geométricos para la realización de las reacciones especie-especie y especie-ADN. Se contabilizó el número de rompimientos de ADN a diferentes dosis de radiación utilizando una fuente de rayos γ procedentes del ^{60}Co . Con esto se compararon los resultados obtenidos con el trabajo de Tomita et al. [55] [56] [57], el modelo de ADN se localizó sobre las posiciones calculadas y que corresponden a seis estructuras de plásmidos (pBR322) generados a una temperatura constante de 293K.

3.2. Primera Simulación

La primera simulación corresponde a la comparación de modelos geométricos del ADN bajo las mismas condiciones de irradiación. En esta simulación solamente se consideran daños producidos por los procesos de ionización, es decir, daños directos. El objetivo es cuantificar la diferencia que se obtiene en el número de rompimientos al cambiar los modelos geométricos del ADN y comparar con datos experimentales.

3.2.1. Geometría

La geometría consistió en un anillo de 2000 bases pares de ADN. Como modelo geométrico de ADN se utilizó el de D. E. Charlton (Figura 3.1) o el modelo de elip-

soides (Figura 3.2) basado en el trabajo de Meylan [41]. El modelo de Meylan está hecho en la herramienta DnaFabric en C++ [42], una herramienta que tiene como finalidad hacer estructuras complejas y detalladas de ADN a partir del formato del protein daba bank (PDB) para Geant4/Geant4-DNA.

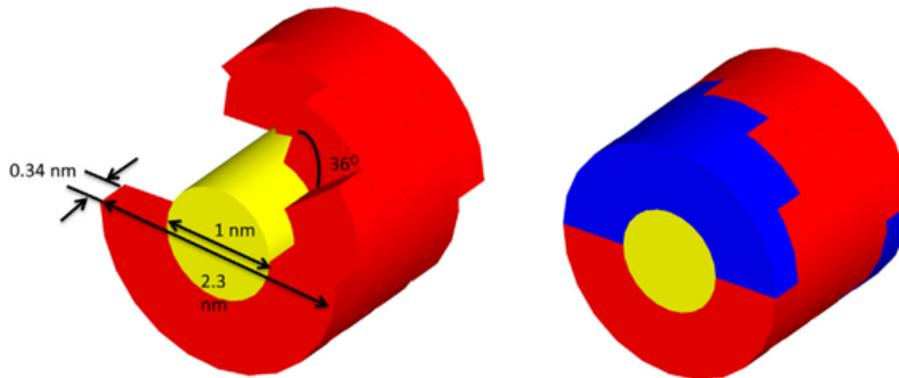


Figura 3.1: Modelo del ADN de Charlton, consiste en tres geometrías por base par, dos secciones de 90° de cilindros y un cilindro en el centro, imagen tomada del artículo de Mc Namara et. al. [38].



Figura 3.2: Modelo del ADN de Meylan. (a) muestra una base par compuesta por seis volúmenes independientes, fosfatos, desoxirribosas y la base par junto con la capa de hidratación. (b) muestra el modelo en una cadena recta, imagen tomada de la referencia [41].

El modelo de Charlton, disponible en la suite geométrica de TOPAS-nBio, se muestra en la Figura 3.1. Para el modelo de elipsoides no se utilizó DnaFabric ya que tal herramienta utiliza como argumento archivos pre-definidos PDB, mientras que los modelos de plásmidos generados en este trabajo son aleatorios (resultado del algoritmo de Metrópolis). Además, DnaFabric utiliza sólidos booleanos que degradan la eficiencia en el cálculo del transporte de radiación. Por lo tanto, el modelo de elipsoides fue desarrollado desde cero basado en

el modelo de Meylan, tomando en cuenta tres factores:

1. Mantener aproximadamente el mismo volumen.
2. Mantener el mismo diámetro del ADN.
3. Evitar sólidos booleanos y la intersección entre volúmenes.

Con estas tres condiciones el modelo propuesto cambia su forma final de la mostrada en la Figura 3.2. El efecto de la forma final se explica con mayor detalle en la sección de resultados.

El anillo se colocó en el centro del mundo (una caja compuesta de aire, densidad = $0,00120479g/cm^3$) de $6 \times 6 \times 6cm^3$. Justo antes del anillo se colocó una caja de Petri de material vinil tolueno (densidad= $1,032g/cm^3$) de 0.8 mm de grosor centrado en $z=-0.71$ mm. Posteriormente se colocó una lámina de vidrio (densidad= $2,4g/cm^3$) de 0.3 mm de grosor centrada en $z=-0.15010$ mm, ver diagrama en la Figura 3.3. El material para todas las componentes de los modelos del ADN es agua líquida a temperatura ambiente.

3.2.2. Irradiación

La fuente de irradiación consistió de un espacio fase generado con TOPAS en modo Monte Carlo de transporte condensado a partir de diez millones de protones. Para esto, se simularon tres escenarios independientes utilizando protones con energías de 10, 20 y 30 MeV. Cada espacio fase fue generado en aire en un plano con dimensiones 60 mm \times 60 mm. La fuente de partículas se posicionó a 3 m SSD (source-to-surface distance) recreando las condiciones experimentales de irradiación para la generación de radiación secundaria. Subsecuentemente, para irradiar el plásmido, se escaló el espacio fase por un factor de 4×10^{-6} en XY debido a que el área de irradiación de el plásmido fue de $120 \mu m \times 120 \mu m$. Por lo tanto, los datos finales fueron corregidos por este factor de escalamiento. Este método de irradiación fue el mismo utilizado en el artículo de McNamara [38]. Cada historia fue considerada independiente de las otras historias.

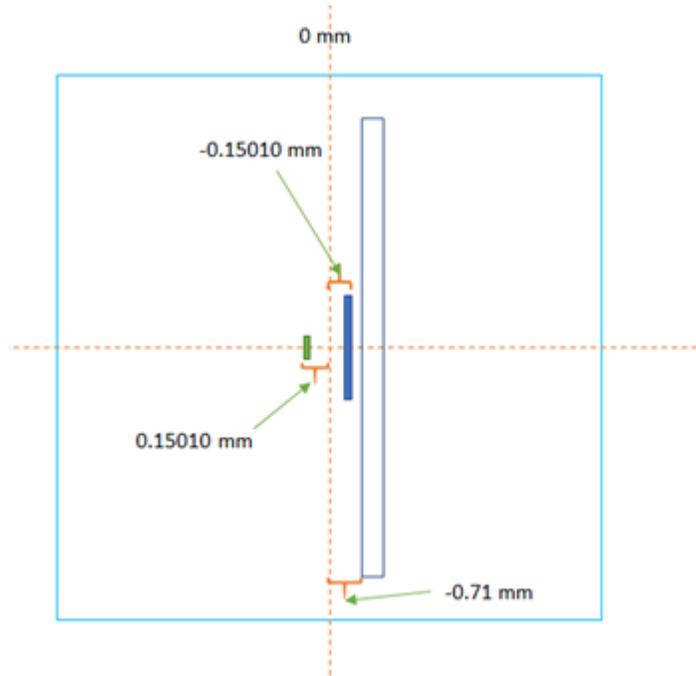


Figura 3.3: Esquema de la geometría de la primera simulación, el rectángulo verde es la posición de plásmido circular (anillo), el rectángulo azul relleno es la posición de la cara inferior de la caja de Petri, y el rectángulo azul sin relleno es la posición de la cara inferior de la lámina de vidrio. Debido a la dirección del haz ($Z+$ a $Z-$) se asume que no hay contribución de las caras superiores tanto de la caja de petri como de la lámina de vidrio.

3.2.3. *Scorers*

En esta simulación se cuantificó el número de rompimientos de ADN causados por el depósito de energía en los volúmenes que componen el modelo de ADN. La salida de datos de TOPAS-nBio consistió de las posiciones y energía depositada de eventos de ionización y excitación que ocurrieron en los volúmenes ubicados en las hebras de ADN. Por lo que, en este caso, se desarrolló un algoritmo externo para procesar los eventos de depósito de energía y cuantificar los rompimientos de ADN. Se consideró que un rompimiento ocurre cuando la energía depositada acumulada excedió 17.5 eV, ver sección 2.1.3.

3.3. Segunda Simulación

En la segunda simulación se calcularon los valores G de las especies químicas producidas por la radiólisis del agua producidos por gammas de ^{60}Co (1.25 MeV). Este tipo de fuente de radiación fue utilizada en la medición experimental de datos de referencia del valor G para el radical $\bullet\text{OH}$ [35], principal causante de daños indirectos al ADN. Las simulaciones de tal escenario utilizan típicamente segmentos cortos de *tracks* con una LET aproximadamente igual al del ^{60}Co ($0,3\text{keV}/\mu\text{m}$) obteniendo resultados muy satisfactorios, por ejemplo, protones de 300 MeV [48] o de electrones de 1 MeV [50]. El uso explícito de fotones de ^{60}Co en la simulación Monte Carlo mejora el detalle de las condiciones experimentales. Más aún, en esta tesis se reporta por primera vez el uso de este tipo de radiación para el cálculo de valores G. Como prueba de consistencia con simulaciones realizadas en trabajos anteriores, se utilizarán también una fuente de *tracks*-cortos de protones de 300 MeV de $100\ \mu\text{m}$ de longitud. Por último, se realizó la simulación utilizando gammas de ^{60}Co en un volumen más pequeño, del mismo tamaño que el utilizado en la simulación de la irradiación de los plásmidos descrita en la sección 3.4, en presencia y ausencia de solventes.

3.3.1. Geometría

La geometría utilizada para calcular los valores G de referencia consistió de una caja de agua (densidad $1\ \text{g}/\text{cm}^3$) semi infinita (de 1 km de lado) para contener a todas las partículas secundarias generadas. Por otro lado, la simulación utilizando el volumen más pequeño consistió de un cubo de $1\ \mu\text{m}$ de lado. Para esta simulación, los disolventes conocidos como TRIS (hidroximetil aminometano), EDTA (ácido etilendiaminotetraacético) fueron incluidos. Las concentraciones de estas sustancias son las siguientes: TRIS $1\ \text{mmol}\ \text{dm}^{-3}$ y EDTA $0.1\ \text{mmol}\ \text{dm}^{-3}$. Además, se simuló una concentración de $5\ \text{mmol}\ \text{dm}^{-3}$ de O_2 (oxígeno) que equivale a la concentración de oxígeno disuelto en agua a 1 atm de presión.

3.3.2. Irradiación

Para la segunda simulación se generaron partículas posicionadas a $(0, 0, -0.5 \mu\text{m})$ con dirección a $+z$, con zero divergencia. Estas partículas fueron *tracks* cortos de $100 \mu\text{m}$ de protones de 300 MeV o gammas de 1.25 MeV. Estas últimas representan la energía más probable del espectro de emisión del ^{60}Co . En el caso de la irradiación con gammas, se utilizó la técnica de reducción de varianza de interacción forzada a $0.5 \mu\text{m}$ de valor más probable de interacción Compton y producción de pares.

3.3.3. Scorers

Se reportó el valor G en función del tiempo (1 ps - 1 μs) utilizando 20,000 historias totales divididas en 4 CPUs asumiendo que cada historia es independiente de las otras historias.

3.4. Tercera Simulación

3.4.1. Geometría

En esta última simulación, se utilizaron seis plásmidos súper enrollados uniformemente distribuidos en un volumen cúbico de $988 \times 988 \times 988 \text{ nm}^3$. Cada plásmido corresponde al plásmido pBR322 cuya longitud es de 4362 bases par ($\sim 1483.08 \text{ nm}$ de longitud)[56]. El modelo geométrico del ADN para estos plásmidos fue el de elipsoides solamente. La composición del cubo de agua y componentes del ADN fue agua a temperatura ambiente y 1 atm de presión. A la par del modelo de ADN, se incluyó en la simulación a las posiciones atómicas de los constituyentes de cada nucleótido (32 átomos para la Adenina, 30 para la citosina, 33 para la guanina y 32 para la timina) con respecto a cada base par. Estas posiciones fueron utilizadas por el algoritmo IRT para incluir la realización de reacciones químicas del radical $\bullet\text{OH}$ con el ADN. El sistema incluyó los solventes químicos TRIS y EDTA del cuadro 2.1, a esas mismas concentraciones.

3.4.2. Irradiación

Se utilizó una fuente de irradiación uniforme con gammas de ^{60}Co , en la cual, cada partícula se generó desde un punto aleatorio sobre la superficie del mundo (cubo de $\sim 1\mu\text{m}^3$) con una dirección aleatoria dictada por un punto uniformemente distribuido dentro del volumen del mundo, ver el diagrama de la Figura 3.4. En esta simulación no se consideró la independencia entre historias, sino la acumulación de varias historias. Por lo tanto, la difusión de especies químicas se realizó cada vez que se obtuvo una dosis integrada de radiación igual a una dosis prescrita. Se usaron los siguientes valores de dosis: 6, 12.5, 25, 50 y 100 Gy. Se usó el método de interacción forzada con una longitud de interacción de $0.57\mu\text{m}$, siendo esta la longitud mínima promedio de una recta que atraviesa aleatoriamente un cubo de $1\mu\text{m}$ por lado.

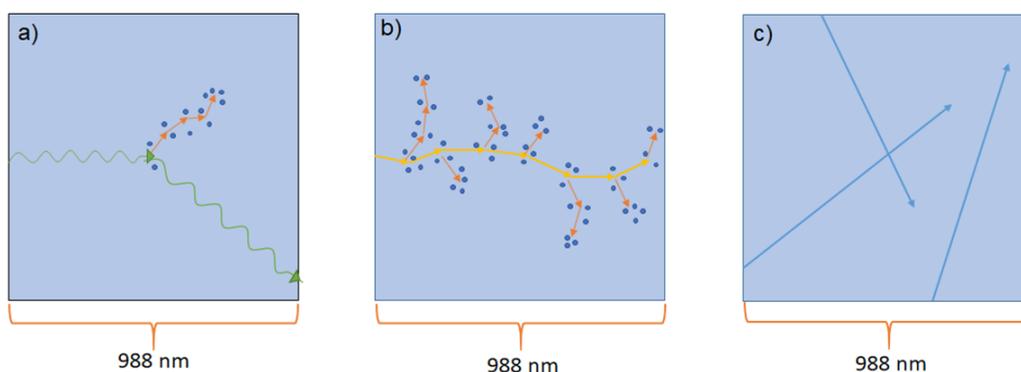


Figura 3.4: El volumen del mundo simulado en esta etapa es agua, con un tamaño de 988 nm por lado. a) el comportamiento de los rayos γ (línea curvada verde) en el volumen simulado, estos interactuarán una sola vez dentro del volumen designado generando un electrón por producción de pares o dispersión Compton, debido a que el rayo γ solo interactúa una vez (deposita energía) su trayectoria no generará especies químicas (puntos azules). b) el comportamiento de los protones, al interactuar un mayor número de veces, se observan depósitos a lo largo de la trayectoria de este y en la trayectoria de los electrones secundarios generados. c) el algoritmo de irradiación genera un punto aleatorio sobre la superficie del mundo y manda una partícula en una dirección aleatoria denotada por un punto dentro del volumen de este (se crean 2 puntos para este método de irradiación).

3.4.3. Scorers

Para la parte química se incluyeron las posiciones atómicas del átomo C4 [56] de los nucleótidos al programa IRT, aunque la elección de otros átomos se dejó como parámetro

Nucleótido	Especie Química	Tasa de Reacción ($mmol\ dm^{-3}\ s^{-1}$)	Radio de Reacción	Probabilidad de Daño
A	$\bullet OH$	$6,1 \times 10^6$	$0,36nm$	50%
A	$H\bullet$	$0,1 \times 10^6$	$0,0018nm$	0%
A	e_{aq}^-	9×10^6	$0,24nm$	0%
T	$\bullet OH$	$6,4 \times 10^6$	$0,38nm$	50%
T	$H\bullet$	$0,56 \times 10^6$	$0,0094nm$	0%
T	e_{aq}^-	18×10^6	$0,48nm$	0%
C	$\bullet OH$	$6,1 \times 10^6$	$0,36nm$	50%
C	$H\bullet$	$0,092 \times 10^6$	$0,0017nm$	0%
C	e_{aq}^-	14×10^6	$0,37nm$	0%
G	$\bullet OH$	$9,2 \times 10^6$	$0,5nm$	50%
G	e_{aq}^-	14×10^6	$0,37nm$	0%

Cuadro 3.1: Reacciones de especies químicas con el ADN, los datos de la tasa de reacción fueron obtenidos de la literatura [41], el radio de interacción fue calculado a partir de la ecuación (2.25) teniendo en cuenta que la difusión de las especies de ADN es cero.

libre en la interfaz final en TOPAS-nBio. Se asumió que los plásmidos son estáticos (no se mueven) durante el curso de la química homogénea (hasta $1\ \mu s$). En el cuadro 3.1 se muestran las reacciones que ocurren con el ADN y de estas las que pueden producir daño. El valor de probabilidad de daño del cuadro 3.1 es un parámetro libre para comparar con datos experimentales u otros trabajos, en este trabajo se utilizó el valor usado por Tomita [56], aunque se han reportado otros valores [16], por ejemplo Friedland [20] utiliza un factor de 65 %.

El electrón solvado e_{aq}^- tiene muy poca posibilidad de causar daño [37], el hidroxilo $\bullet OH$ es el principal contribuyente al daño al ADN. En las referencias mencionados anteriormente [15] [16] [26] [56] ésta especie química es la única que se utiliza para generar daños indirectos. Por otro lado, se tienen registros de que el hidrogeno $H\bullet$ también puede causar daño [49], sin embargo, en este trabajo solo consideraremos al hidroxilo al igual que en las referencias mencionadas anteriormente. El algoritmo IRT utilizó las posiciones del radical C4 para realizar las reacciones químicas con el $\bullet OH$. Una vez que algún rompimiento simple (físico o químico) ocurrió, la base par al número de hebra correspondiente quedó eliminada de futuras reacciones. Se reportó el número absoluto de rompimientos simples y dobles en función de la dosis depositada.

Capítulo 4

Resultados y Discusiones

4.1. Plásmidos de Súper Enrollados

En este trabajo de tesis se generaron plásmidos pBR322 (4362 bps) con 145 segmentos. En la Figura 4.1 se observa una serie de plásmidos creados a partir de diferentes parámetros iniciales para probar diferentes partes del algoritmo desarrollado en este trabajo. El plásmido etiquetado con el índice 1) es un ejemplo de plásmido generado a partir de un bajo número de iteraciones (10,000 iteraciones) a 293 K. El plásmido con el índice 2) fue hecho al número normal de iteraciones (más de 10 millones). En el índice 3) el plásmido se generó bajo las mismas condiciones que el plásmido 1), pero con una semilla aleatoria distinta. El índice 4) es un plásmido creado a temperatura variable (de 400 K a 293 K), el algoritmo de este plásmido no es el final, en este caso se usó el algoritmo lineal para el cálculo de la nueva temperatura, ver sección 3.1. El índice 5) denota un plásmido generado con el criterio final de reducción de temperatura, es decir, se utilizó una reducción exponencial de la temperatura de 400 K a 293 K. Finalmente, el plásmido 6) se generó con el mismo criterio de temperatura variable que el plásmido 5), pero con límites de 450 K a 10 K. El tiempo promedio para producir un plásmido fue de 6 horas. Estos plásmidos se generaron utilizando 1 CPU con procesador 3.1 GHz.

Estos seis plásmidos muestran la susceptibilidad del algoritmo, dado que ligeros cambios en los parámetros iniciales, así como la susceptibilidad a cambios de temperatura

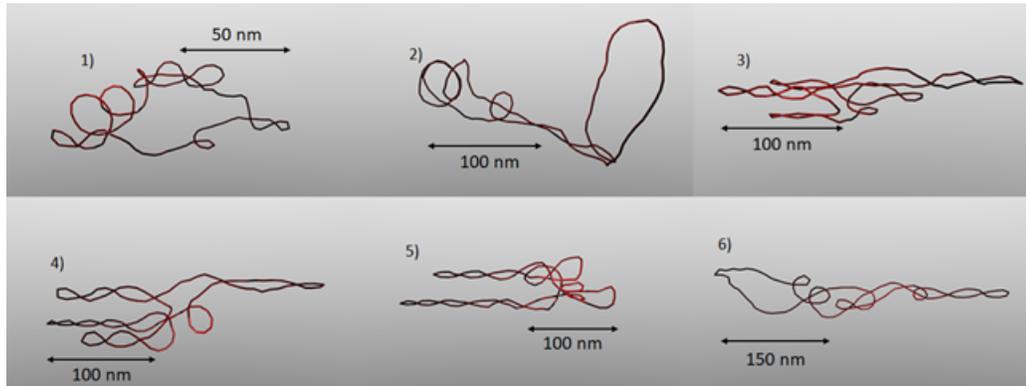


Figura 4.1: En la figura se pueden ver seis plásmidos pBR322 etiquetados como 1), 2), 3), 4), 5) y 6), los cuales fueron creados a partir de diferentes parámetros de Temperatura y número de iteraciones para probar la sensibilidad a los cambios de temperatura del programa.

crear conformaciones completamente diferentes. Esto, de acuerdo a Kummerle (Kummerle, E. A, 2004), facilita que los plásmidos se súper enrollen usando un número menor de iteraciones.

En la Figura 4.2 se muestran las gráficas de evolución de la energía y la temperatura en función del número de iteraciones para los plásmidos 2), 4), 5) y 6). El plásmido 1) se dejó fuera por su bajo número de iteraciones y el plásmido 3), debido a estar hecho a partir de los mismos parámetros que el plásmido 2), también se dejó fuera. Como se muestra en la figura, para el plásmido 2) la energía inicial fue de $2,7 \times 10^{-18}$ J. Para este plásmido se utilizó un chequeo de equilibrio cada 3 millones de iteraciones a partir de 4 millones de iteraciones (i. e. se checo el equilibrio a 7 millones, 10 millones y 13 millones) llegando al equilibrio a los 13 millones con una energía promedio final de $2,65 \times 10^{-19}$ J. Para el plásmido 4), cuya energía inicial y final fue de $2,7 \times 10^{-18}$ J y $5,57 \times 10^{-19}$ J, respectivamente, el chequeo de equilibrio (comparando el radio de giramiento) se realizó una vez que se alcanzó la temperatura final deseada, esto tomó más tiempo de simulación (1.5 veces más) dependiendo del algoritmo de disminución de temperatura, llegando al equilibrio a los 18 millones de iteraciones. Para el plásmido 5) se utilizó un algoritmo exponencial para el cambio de la temperatura (el chequeo de equilibrio igualmente debió esperar a que la temperatura final se alcanzó) a diferencia del algoritmo lineal usado en el plásmido 4) la energía inicial y final son consistentes con los plásmidos anteriores, teniendo una energía final de $5,72 \times 10^{-19}$ J y alcanzando el equilibrio a las 16.17 millones

de iteraciones. Finalmente, para el plásmido 6) la energía final fue $3,56 \times 10^{-19}$ J. La temperatura se redujo de 450 K a 10 K, y se puede ver que, a estos rangos de temperatura, las fluctuaciones en la energía de la configuración varía muy poco.

El comportamiento de la reducción de la energía en función del número de iteraciones es explicado por las ecuaciones (2) y (3). La máxima energía corresponde a un plásmido circular abierto, es decir, a iteración 0. En este caso no existe torcimiento por lo que el número de torcimiento es cero en la ecuación (3). Además, los ángulos Θ_i de la ecuación (2) son constantes y máximos. Por lo tanto, cualquier plásmido deformado tendrá una energía menor. Más aún, el algoritmo de Metrópolis descarta cualquier configuración con energía mayor a la configuración anterior. Como consecuencia, el número de torcimiento se incrementará a mayor super enrollamiento, como se muestra en la Figura 4.3 para todos los plásmidos.

Por otro lado, la elección del algoritmo de disminución de temperatura no altera el comportamiento de la evolución de la energía con el número de iteraciones, si se utilizan bajos cambios de temperatura (de 400 K a 293 K), como se muestra para los plásmidos 4) y 5) de la Figura 4.2. Sin embargo, a grandes cambios (de 400 K a 10 K) (plásmido 6)), si existe cambio significativo, no solo en la evolución de la temperatura sino también en la apariencia de los plásmidos. En este caso, se llega de manera más suave al equilibrio térmico, un comportamiento esperado de acuerdo al trabajo de Kummerle [33]. Todos los plásmidos pBR322 creados usando este cambio de temperatura (400 K a 10 K) adoptan una forma plana.

Debido al comportamiento del número de torcimiento o writhe de las imágenes de la Figura 4.3 se decidió utilizar como algoritmo final de disminución de temperatura el algoritmo exponencial, ya que al usar rangos pequeños de temperatura, asemeja más el comportamiento de la deformación a temperatura constante y a grandes cambios de temperatura contribuye enormemente al equilibrio del plásmido.

La herramienta de plásmidos que se desarrollo es capaz de reproducir una estructura súper enrollada para plásmidos de manera rápida y eficiente, logrando generar plásmidos en un máximo de 8 hrs, esto en parte a las paqueterías del CHLEP las cuales están optimizadas para trabajar con vectores. Los plásmidos generados usando esta herramienta

muestran los comportamientos similares, que los reportados en el trabajo de Kummerle y Vologodskii. Cuyos números de writhe para el plásmido pBR322 son en promedio -25 para plásmidos a temperatura constante de 293 K, -19 para plásmidos con un cambio de temperatura de 400 K a 293 K y -21 para plásmidos con cambio de temperatura de 450 K a 10 K. Las diferencias en los valores de energía y numero de torcimiento se deben a los valores iniciales de los plásmidos y a los valore finales de energía elástica entre los reportados en las referencias y los de este trabajo plásmidos, por lo que se tienen configuraciones de equilibrio equivalentes.

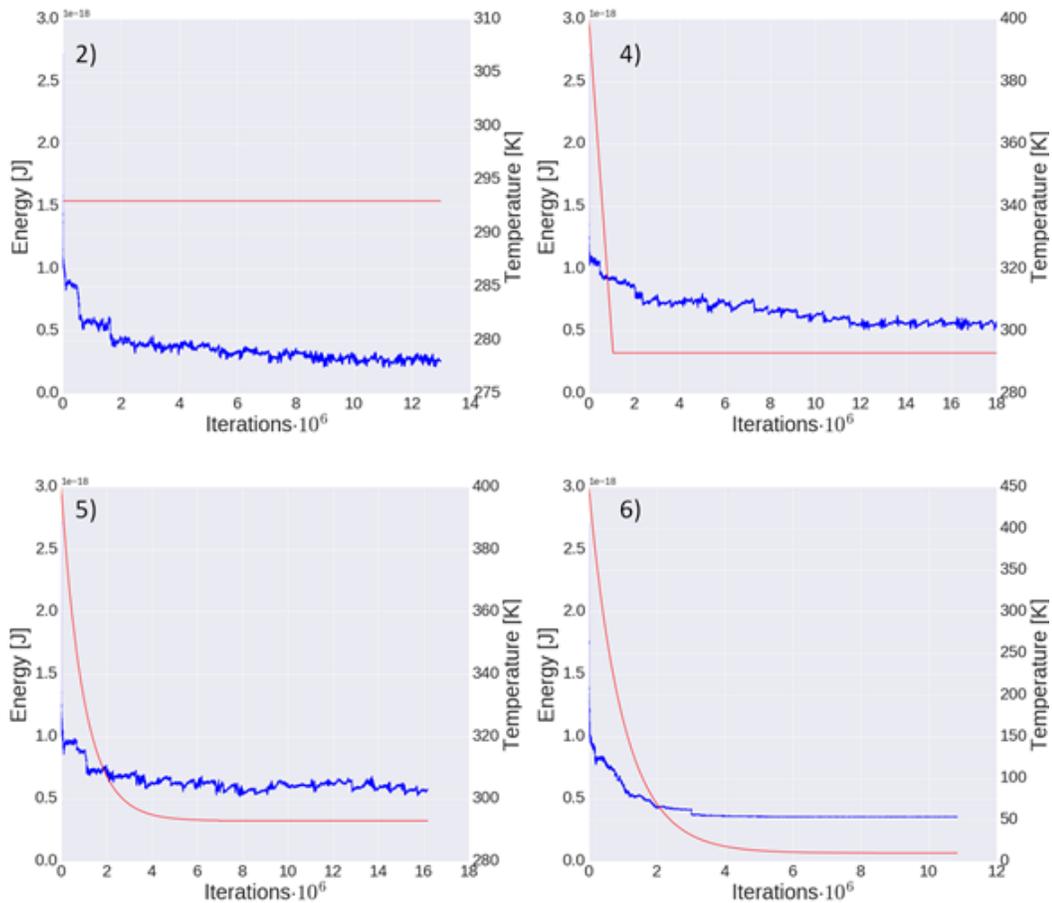


Figura 4.2: Evolución de la energía (mostrado en el eje izquierdo con unidades en Joule línea azul) y temperatura (mostrado en el eje derecho con unidades en Kelvin línea roja) en función del número de iteraciones para los plásmidos 2), 4), 5) y 6) de la Figura 4.1.

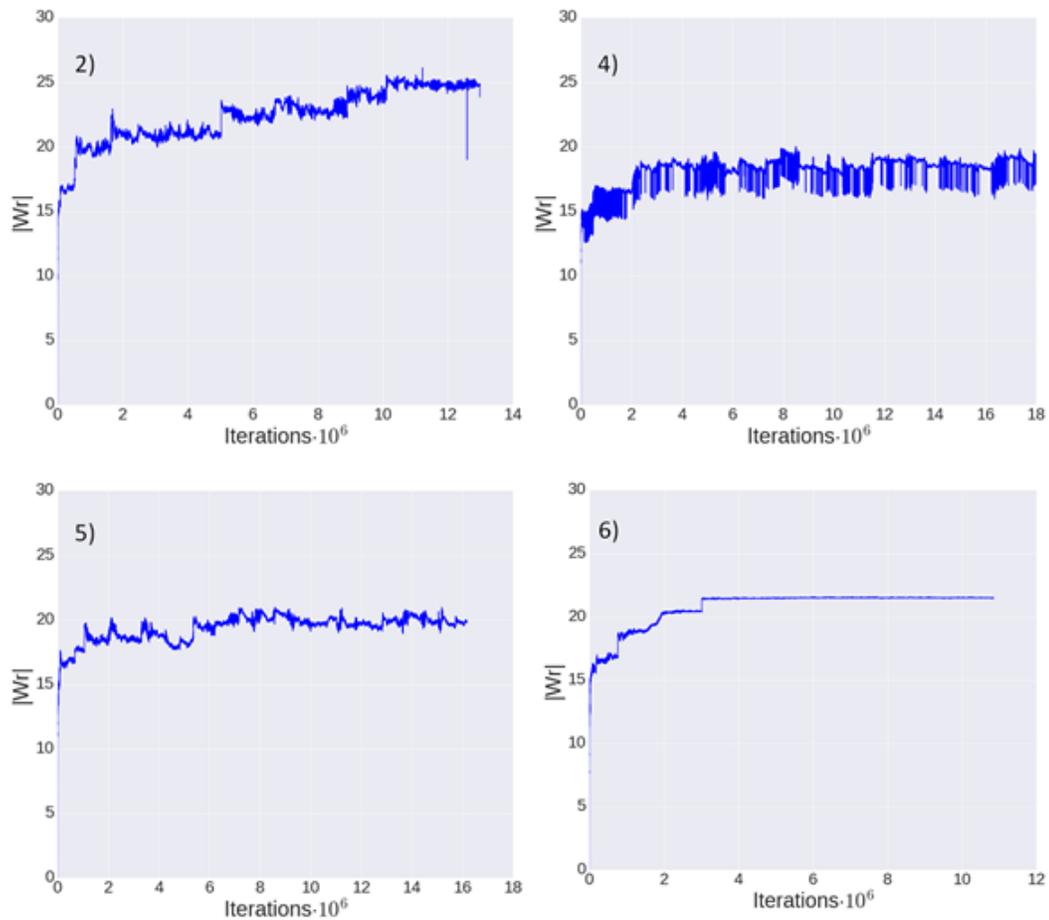


Figura 4.3: : Evolución del *writhe* en función del número de iteraciones para los plásmidos 2), 4), 5) y 6) de la Figura 4.1.

4.2. Modelo de ADN

El modelo geométrico del ADN desarrollado se puede observar en la Figura 4.4. En la figura, para cada base par se utilizaron seis elipsoides. Este modelo fue consistente con el de Meylan, con un diámetro de hélice 2.3 nm y longitud de base 0.34 nm, pero evitando el uso de sólidos booleanos y la intersección entre volúmenes. Esto produce simulación de transporte de radiación más eficientes por la siguiente razón. En los algoritmos de transporte, a cada paso (es decir, después de cada interacción) se calcula la posición de la partícula y la distancia a cada frontera geométrica (cara interna) del volumen en el que la partícula está en ese momento. Entonces, si la simulación consiste, por ejemplo, solo de un cubo, se hace seis veces la evaluación de la misma función para saber la distancia hacia

Modelo	Fosfatos (nm^3)	Desoxirribosas (nm^3)	Base Par (nm^3)
Elipsoides	0.060	0.091	0.093
Charlton		0.9793*	0.1335
Meylan	0.060	0.091	0.093

Cuadro 4.1: Comparación en volúmenes de los modelos de ADN, el modelo de Cilindros de Charlton, el Modelo basado en esferas con sólidos booleanos de Meylan y el modelo de Elipsoides desarrollado y utilizado en este trabajo. *El modelo de Charlton solo cuenta con un volumen que contabiliza tanto los fosfatos como las desoxirribosas.

la siguiente frontera más cercana. Sin embargo, utilizando sólidos booleanos, se invocarán tantas funciones sean necesarias dependiendo de los sólidos involucrados en la creación de tal booleano, aumentando la complejidad y número de las operaciones matemáticas necesarias.

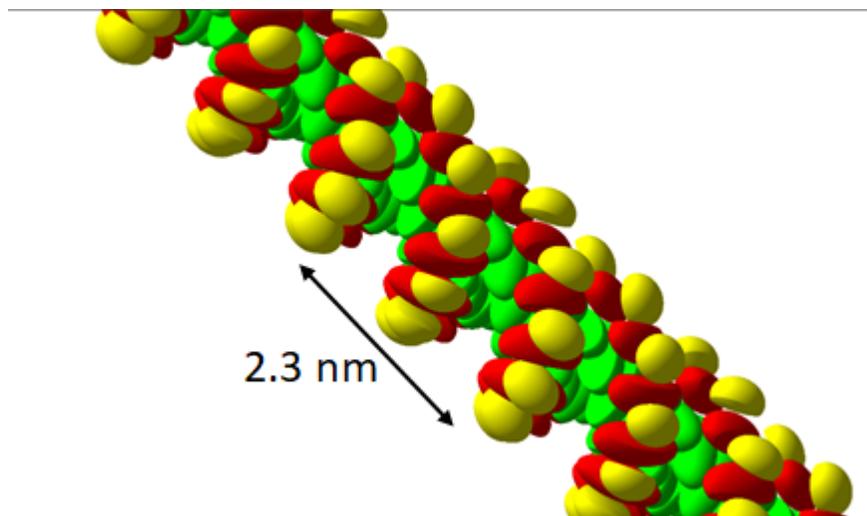


Figura 4.4: Modelo de ADN desarrollado en este trabajo, en la imagen se pueden ver 3 tipos de volúmenes diferentes, con sus respectivos colores (amarillo, rojo y verde), estos tres volúmenes representan los fosfatos, desoxirribosas y las bases pares (respectivamente). Debido a que se repiten 2 veces, cada base par este compuesto por seis volúmenes, dos de cada tipo. El diámetro de la hélice es de 2.3 nm y la altura de una base par es de 0.34 nm

La comparación entre parámetros geométricos fue satisfactoria y se muestra en el cuadro 4.2. Además, en ese cuadro se muestra el modelo geométrico de Charlton.

Aunado a esto, se desarrolló un modelo puramente atómico del ADN para el daño químico, el cual se muestra en la Figura 4.5. Este modelo fue desarrollado a partir de la información espacial de los nucleótidos del Protein Data Bank recopilados en otro trabajo de tesis [46]. A partir de esta información fue posible construir un modelo atómico del ADN,

tomando en cuenta que la posición, cantidad y tipo de átomos varían para cada nucleótido (A, T, C, G), estas posiciones se muestran con una esfera de diferente color y dimensión en la Figura 4.5.

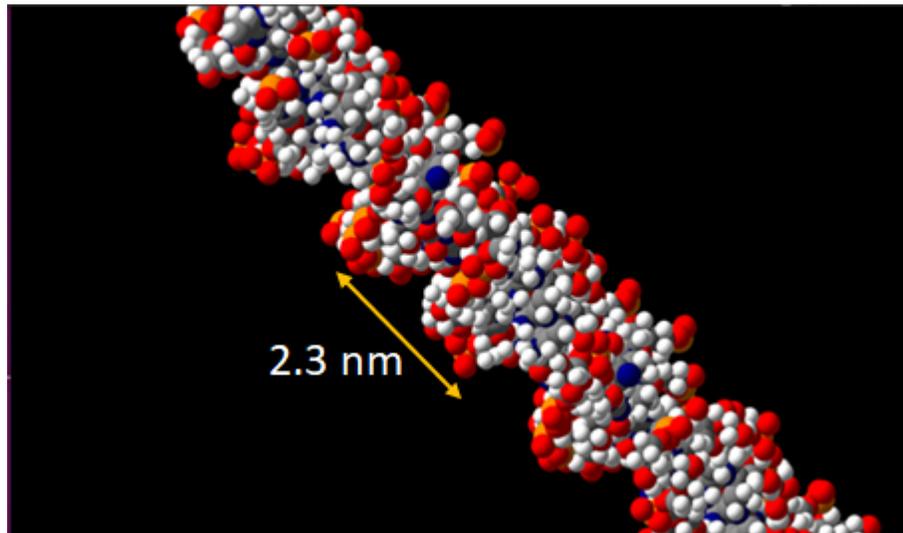


Figura 4.5: Visualización de las posiciones atómicas de las componentes del ADN, se tienen hidrógenos (blancos), carbonos (grises), oxígenos (rojos), nitrógenos (azul oscuro), azufres (amarillos) y fosfatos (naranja).

La última parte de esta sección consistió en combinar ambos modelos alineándolos adecuadamente, como se muestra esquemáticamente para una base par en la Figura 4.6. El primer modelo fue utilizado para contabilizar el daño directo por depósitos de energía en los elipsoides y el segundo modelo para la realización de reacciones con el ADN y contabilizar el daño indirecto.

En los modelos de ADN, Meylan describe la capa de hidratación como un volumen separado al resto del ADN a diferencia del trabajo desarrollado en esta tesis, en el cual la capa de hidratación está implícita en el volumen total, Figura 4.4. Esta capa de hidratación es la única diferencia entre el modelo de Meylan y el utilizado en este trabajo, teniendo en cuenta que los volúmenes y el diámetro de una base par son iguales en ambos modelos. En el modelo de este trabajo no se utiliza la primera capa de hidratación del ADN de forma explícita (como un volumen separado) debido a que esto llevaría al uso necesario de los sólidos booleanos, y, debido a que el material del ADN y de la primera capa de hidratación fue el mismo en todas las simulaciones (agua), esta capa está contemplada dentro de nuestras geometrías elipsoidales.

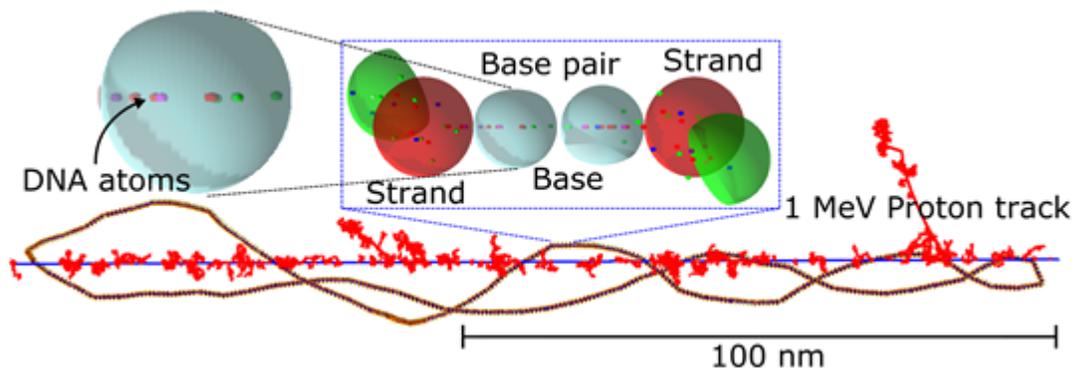


Figura 4.6: Ejemplo de las posiciones atómicas del ADN encerrados dentro de los volúmenes geométricos, en esta imagen se puede ver un plásmido pUC19 con este modelo final del ADN irradiado con protones de 1 MeV.

Este modelo combinado de ADN fue posicionado siguiendo el camino descrito por los lados del plásmido súper enrollado reportado en la sección 4.1. La distancia entre cada base par fue de 0.34 nm y el ángulo de rotación entre cada base par fue de 36 grados alrededor del eje definido por dos vértices subsecuentes del plásmido súper enrollado. En promedio, 29 bases par por lado fueron utilizadas.

4.3. Primera Simulación

En las Figuras 4.7 y 4.8 se muestra el número de SSB y DSB por mega base par por Gy en función de la LET (en $keV/\mu m$), respectivamente, utilizando el modelo de Charlton y el modelo desarrollado en este trabajo de tesis. En la figura también se muestran los datos simulados y experimentales reportados en el trabajo de McNamara et al., [38]. Las condiciones de la simulación se describieron en la sección 3.2. Los valores de energía de las fuentes de protones fue 20-30 MeV. Estos valores corresponden a la región más importante de la curva de dosis terapéutica, el pico de Bragg. En ambos resultados se encontraron diferencias significativas (las incertidumbres estadísticas fueron menores al 1%) entre los modelos de ADN, debido a la diferencia de volúmenes significativa como se ve en el cuadro 4.2. El modelo de Charlton, al ser de mayor volumen, registró más eventos de depósito de energía comparado con el modelo de elipsoides, como se demuestra en las Figuras 4.7 y 4.8 para SSB y DSB, respectivamente. No obstante, el modelo de elipsoides reprodujo de mejor

manera a los datos experimentales de SSB. Por otro lado, para DSB las diferencias con respecto a los datos experimentales fue mucho mayor para ambos modelos. Existen varias causas para esta discrepancia. La técnica experimental de electroforesis utilizada para la medición de datos no cuenta con la resolución necesaria para clasificar DBS complejos de los DBS simples, en la simulación se clasificaron estrictamente DSB simples. Además, el tratamiento enzimático de las muestras después del tratamiento puede afectar el daño de bases par McNamara et al. [38]. Por otro lado, se utilizó un modelo de plásmidos anular ideal de 2000 bases pares, el cual dista de los modelos súper enrollados típicamente utilizados en los experimentos.

4.4. Segunda Simulación

Los resultados de la segunda simulación se muestran en las Figuras 4.9, 4.10 y 4.11, en donde se utilizó una amplia gama de configuraciones para la fuente, ver sección 3.1. Esta simulación es crucial para la última parte de la tesis, puesto que el comportamiento de los valores G durante la simulación influyó el resultado final de la tercera simulación ya que el uso de diferentes fuentes generó un espectro de especies químicas iniciales diferente. Debido a que los datos experimentales con los que se comparó en la última parte de la tesis utilizan al ^{60}Co como fuente de emisión de partículas primarias, se le dio prioridad a esta fuente.

Debido a que el ^{60}Co y los tracks de $100\ \mu\text{m}$ de longitud de protones de $300\ \text{MeV}$ tienen un efecto equivalente (debido a la misma LET) se utilizaron ambas fuentes para compararlas. Además, ya que las especies químicas más utilizadas para contabilizar en daño al ADN son: el hidroxilo ($\bullet\text{OH}$), el electrón solvado (e_{aq}^-) y el hidrogeno ($\text{H}\bullet$), solamente se muestran los valores G, en función del tiempo de estas.

En la Figura 4.9 se muestra el valor G en función del tiempo para el hidroxilo $\bullet\text{OH}$. Se muestra el resultado utilizando *tracks* cortos de protones utilizando los métodos paso a paso (SbS) (línea solida gris) e IRT (línea sólida negra). Los datos de referencia son los de IRT y SbS con una incertidumbre estadística porcentual de 0.28 % y 0.9 %, respectivamente. Como se muestra en la figura, las curvas comienzan con un valor máximo cercano a 5 especies/100 eV al final del estado pre-químico (1 ps) y eventualmente disminuyen confor-

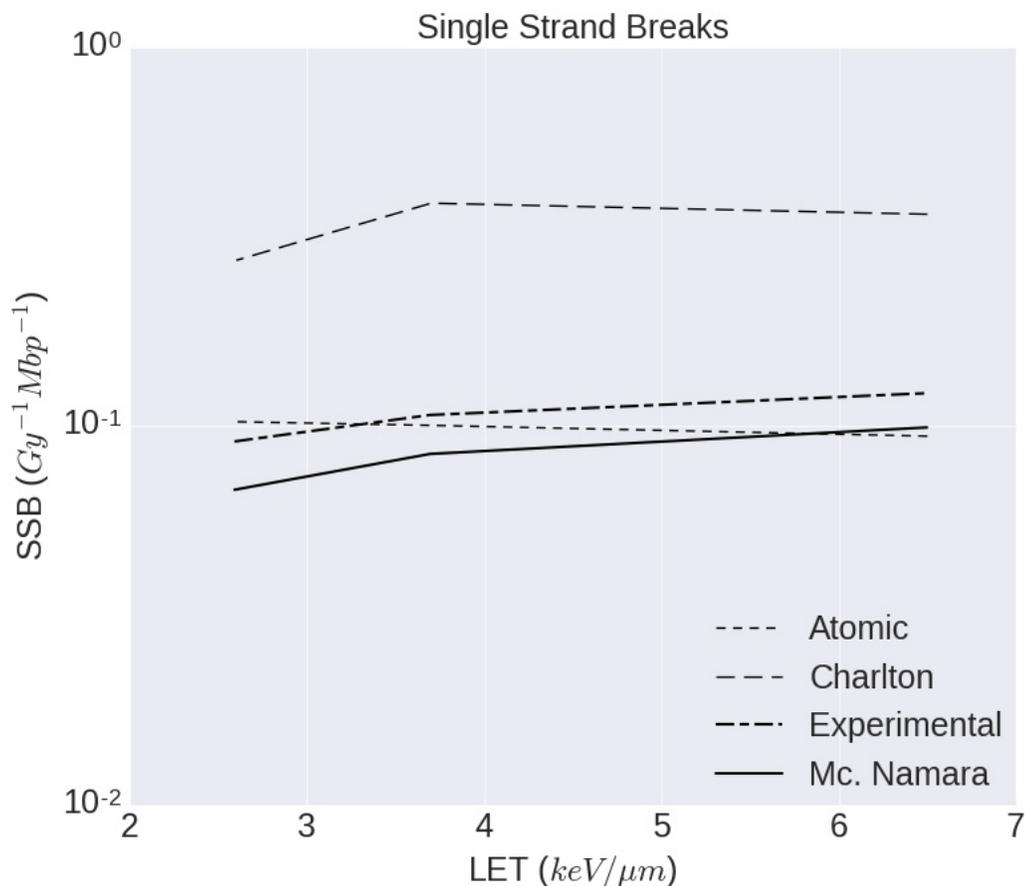


Figura 4.7: Número de rompimientos simples normalizados por dosis y por mega base par, contra la LET de las partículas incidentes. Se tienen datos experimentales de McNamara et al., datos simulados, por la misma autora y resultados de nuestras simulaciones usando el modelo de semi esferas y el modelo de Charlton del ADN, las incertidumbres estadísticas para el modelo atomico es de $\pm 0.16\%$, $\pm 0.74\%$, y $\pm 0.14\%$ para 10, 20 y 30 MeV de protones respectivamente, en el modelo de Charlton tenemos incertidumbres de: $\pm 0.073\%$, $\pm 0.048\%$ y $\pm 0.03\%$.

me se expande la nube de especies reaccionando (y entonces reduciendo el número original de especies) con las demás especies químicas hasta alcanzar el estado estable alrededor de 2.5 especies/100 eV a 1 μs , es decir, las especies químicas están lo suficientemente alejadas como para realizar alguna reacción. IRT reprodujo mejor a los datos experimentales de referencia de LaVerne [35], mostrados con estrellas, quien reportó los valores G de la radiólisis del agua utilizando ^{60}Co . Esta mejor reproducibilidad se debió a que la implementación de IRT en TOPAS-nBio utilizó un esquema de reacciones mucho más completo que el utilizado actualmente por Geant4-DNA, con la ventaja de ser mucho más rápido

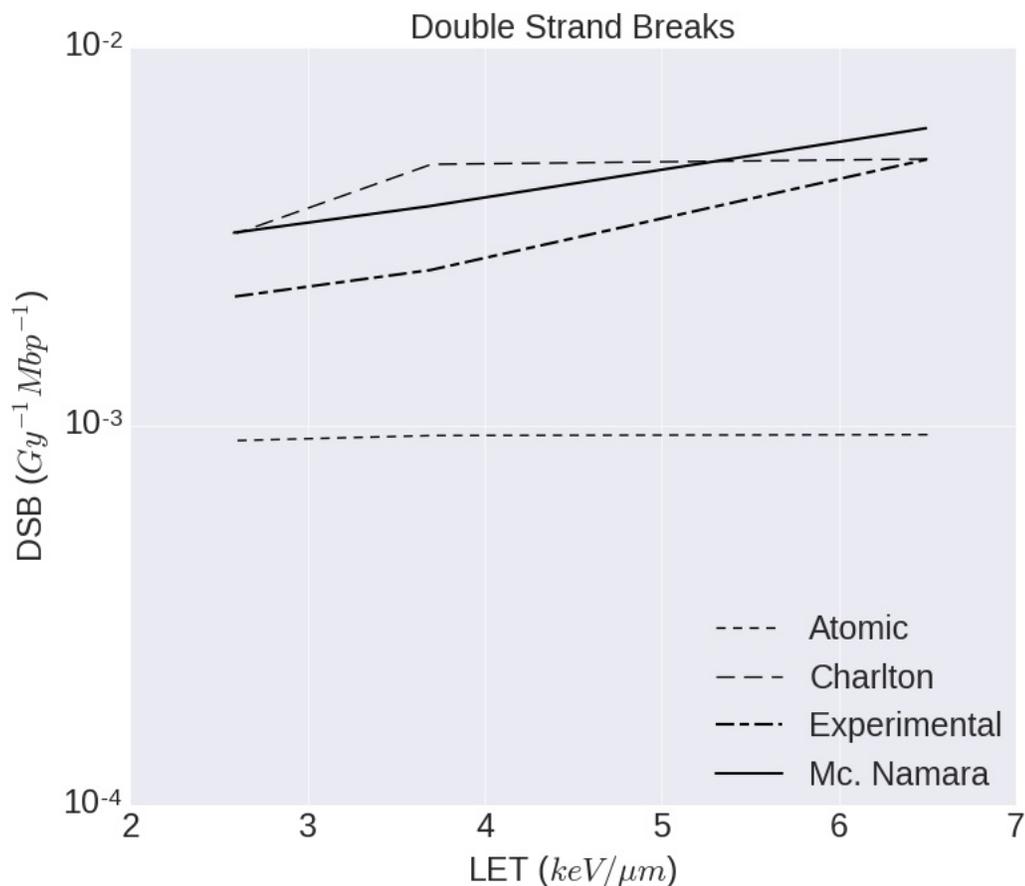


Figura 4.8: Número de rompimientos dobles normalizados por dosis y mega base par, contra la LET de las partículas incidentes, se tienen irradiación para 10, 20 y 30 MeV y la incertidumbre estadística para el modelo de elipsoides (atómico) es de $\pm 3.9\%$, $\pm 9\%$ y $\pm 7\%$ respectivamente, para el modelo de Charlton es de $\pm 2.5\%$, $\pm 1\%$ y $\pm 5.4\%$.

computacionalmente, en relación 150:1. Los resultados utilizando ^{60}Co se muestran con la línea quebrada y gruesa y coinciden, dentro de su incertidumbre estadística de 0.26% con los de IRT. Para estos se demostró, por primera vez, que la técnica de reducción de varianza interacción forzada puede ser aplicable para simulaciones que incluyen el estado químico de la radiólisis. Ya que el camino libre medio de fotones de 1.25 MeV en agua es 15.82 cm, la probabilidad de interaccionar dentro del primer micrómetro de agua es insignificante, por lo que la técnica de reducción de varianza fue esencial para el desarrollo de las simulaciones. La presencia de los solventes TRIS y EDTA en el sistema alteraron la producción de $\bullet\text{OH}$ al reaccionar con este radical, resultados mostrados con la línea gruesa punteada y quebrada, cuya incertidumbre estadística porcentual fue de 0.26%. El

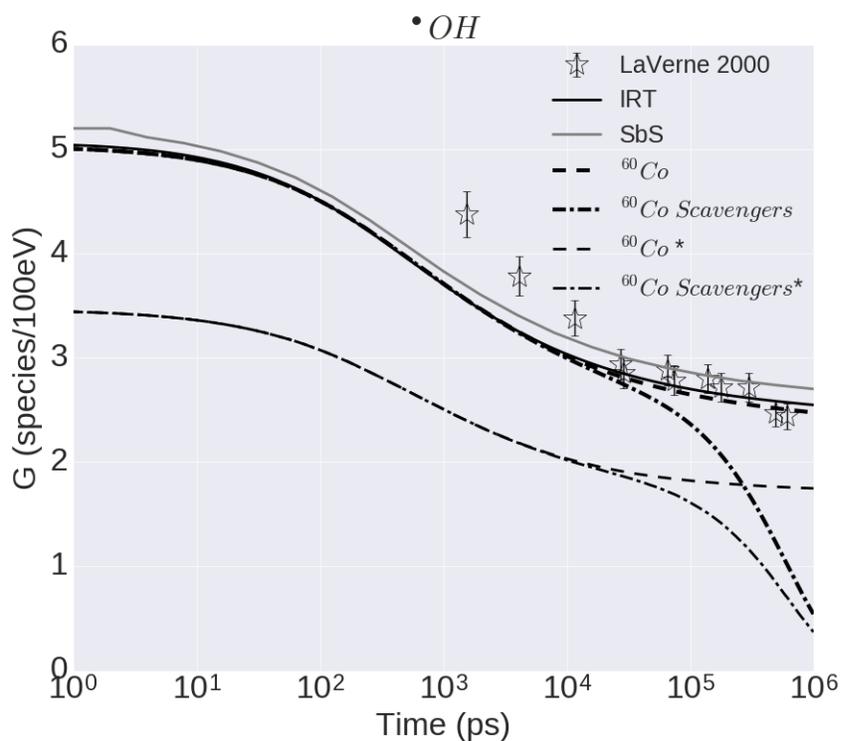


Figura 4.9: Valor G en función del tiempo para el radical $\bullet OH$. Se muestran datos para protones de 300 MeV usando el método IRT y el método step-by-step, datos usando fuentes de cobalto-60 (gammas de 1.25 MeV) con y sin disolventes químicos, y finalmente con Cobalto-60 con y sin disolventes pero usando un volumen reducido de las mismas dimensiones al que se usó en las simulaciones de irradiación de plásmidos finales. Datos experimentales de la literatura se muestran con símbolos, las barras de error representan 1 desviación estándar. El error promedio porcentual para los datos de referencia es de 0.1 %, para el SbS es de 0.25 %, para el $^{60}Co^*$ es de 0.45 %, para el $^{60}Co Scavengers^*$ es de 0.48 %, para el ^{60}Co es de 0.13 %, y para el $^{60}Co Scavengers$ es de 0.26 %

scavenging capacity combinado de este par de solventes es de $1,5 \times 10^6 s^{-1}$ por lo que la reducción significativa de la producción de $\bullet OH$ ocurriría idealmente (es decir, en el sistema aislado $\bullet OH +$ solventes) a $6,6 \times 10^5$ ps, sin embargo, la realización de reacciones es un sistema acoplado; por lo tanto, la contribución de las reacciones con otras especies químicas alteró este valor. Finalmente, al reducir el volumen de agua de la simulación a $1 \mu m^3$, la producción inicial de $\bullet OH$ se redujo significativamente (incertidumbre estadística porcentual de 1.31 %, datos mostrados con la línea quebrada delgada) ya que algunos de los electrones secundarios producidos por los fotones de 1.25 MeV tuvieron suficiente energía para escapar del volumen. Por ejemplo, a esta energía, el efecto dominante es la producción de pares que producirá electrones de 511 keV cuyo rango CSDA es de aproximadamente 2

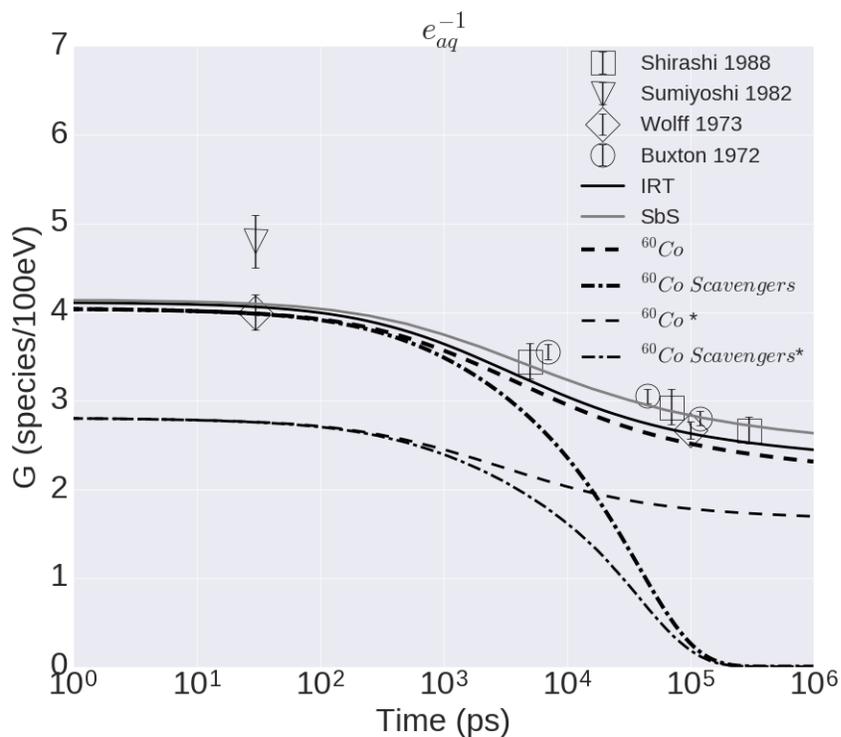


Figura 4.10: Valor G para el electrón solvado, de igual forma se tienen datos para las mismas configuraciones que para el caso de la Figura 4.9, los errores promedios porcentuales en el mismo orden que en la figura anterior son de: 0.08 %, 0.25 %, 0.44 %, 2.13 %, 0.13 %, 0.26 %. Datos experimentales de la literatura se muestran con símbolos, las barras de error representan 1 desviación estándar.

mm. Por último, los resultados incluyendo solventes (incertidumbre estadística porcentual de 1.31 %) se muestran con la línea delgada punteada y quebrada.

Los resultados para el electrón solvado e_{aq}^- se muestran en la Figura 4.10, los datos experimentales hacen referencia a los resultados obtenidos por Shirashi et. al [53], Sumiyoshi et. al. [54], Buxton et. al. [7] y a Wolff et. al. [62]. Los resultados para el hidrógeno H^\bullet se muestran en la Figura 4.11.

4.5. Tercera Simulación

Por último, se muestran los resultados de seis plásmidos pBR322 (~ 4362 bps) a una temperatura constante de 293 K y una densidad de súper hélice de -0.06, en un volumen de

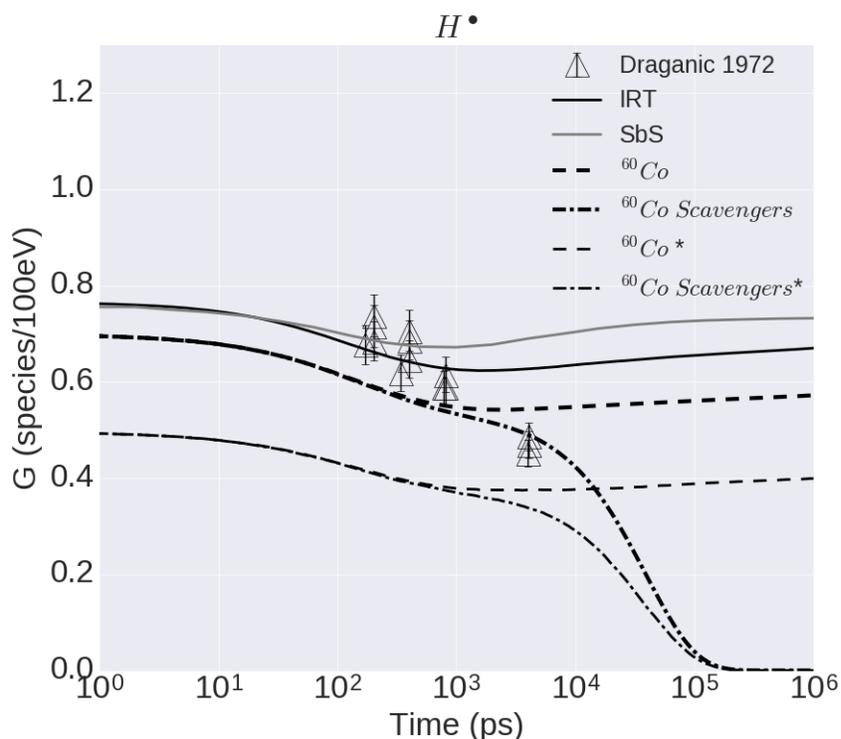


Figura 4.11: Valores G para el hidrogeno. Nuevamente, se tienen las mismas configuraciones que la Figura 4.9, los errores promedios porcentuales en el mismo orden que en la Figura 4.9 son: 0.14 %, 0.44 %, 0.89 %, 3.08 %, 0.16 %, 0.36 %. Datos experimentales de la literatura se muestran con símbolos, las barras de error representan 1 desviación estándar.

$988 \times 988 \times 988 \text{ nm}^3$ para obtener una densidad de $29.75 \mu\text{g}/\text{cm}^3$ de ADN, e irradiados con fotones de ^{60}Co , ver Figura 4.12.

Los resultados de esta simulación para SSB y DSB en función de la dosis prescrita en Gy se muestran en las Figuras 4.13 y 4.14, respectivamente. Los datos se normalizaron por el número de plásmidos (moléculas) de la misma manera que en [56]. En las figuras se muestran los resultados de este trabajo de tesis usando TOPAS-nBio (círculos), datos calculados de la referencia [56] (triángulos) y datos experimentales de la referencia [57]. Estos últimos corresponden a irradiaciones de plásmidos bajo diferentes temperaturas y las concentraciones del disolvente para ver la respuesta del número de rompimientos. Se muestran solo resultados para una concentración específica de disolvente y una temperatura. En este trabajo, para cada valor de dosis la simulación se repitió 200 veces con diferente semilla aleatoria, un mayor número comparado con las 8 repeticiones que se reportan por Tomita; por lo tanto, se redujo la incertidumbre estadística en un factor de

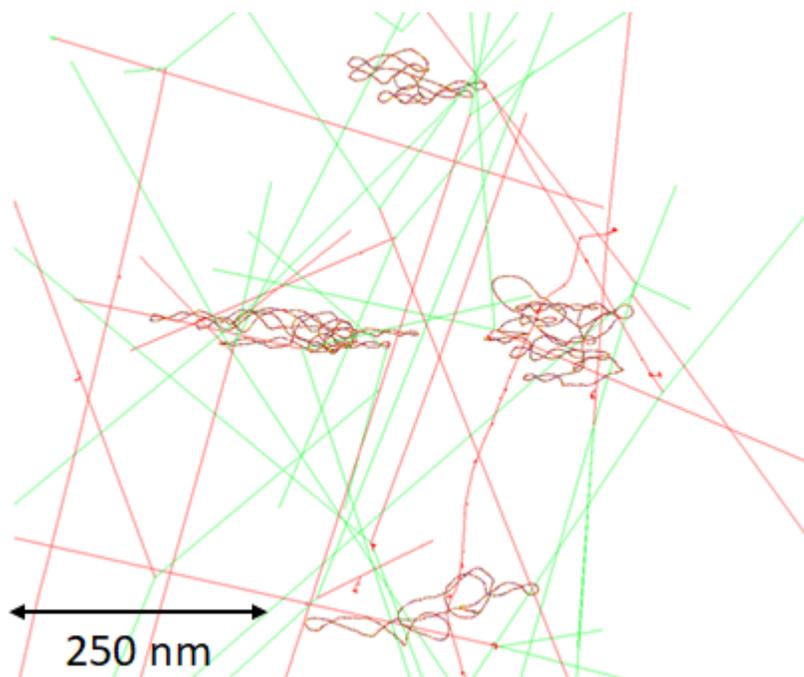


Figura 4.12: Visualización de la tercera simulación. En la imagen se pueden ver cuatro de los seis plásmidos pBR322 utilizados en esta simulación, la fuente de radiación son gammas de 1.25 MeV (líneas verdes) produciendo electrones por efecto Compton o producción de pares (líneas rojas).

5. Las incertidumbres estadísticas para cada dosis se encuentran descritas en la leyenda de cada figura. Para el caso de la dosis, las incertidumbres estadísticas no superaron el 0.01 %.

En la Figura 4.13, se muestra una buena concordancia entre los datos simulados en este trabajo con los datos calculados de [56], dentro de las incertidumbres estadísticas, que fueron obtenidos con el software DBREAK, sin embargo, ambas simulaciones difieren de los datos experimentales de [57]. En el caso de los DSB, Figura 4.14, se muestra una buena concordancia con los datos experimentales a bajas dosis, pero surgen diferencias significativas para a valores mayores a 25 Gy, donde se sobre estiman los valores de DSB. Tomita no reportó el error sistemático en los datos Monte Carlo de DSB, por lo que resulta difícil tener alguna comparativa con sus simulaciones, sin embargo, se infiere que tales errores sean superiores a los de SSB de la Figura 4.13. No obstante, ambas simulaciones reproducen la respuesta no lineal del número de DSB a diferentes valores de dosis. Por otro lado, los errores estadísticos reportados en este trabajo se encuentran por debajo de

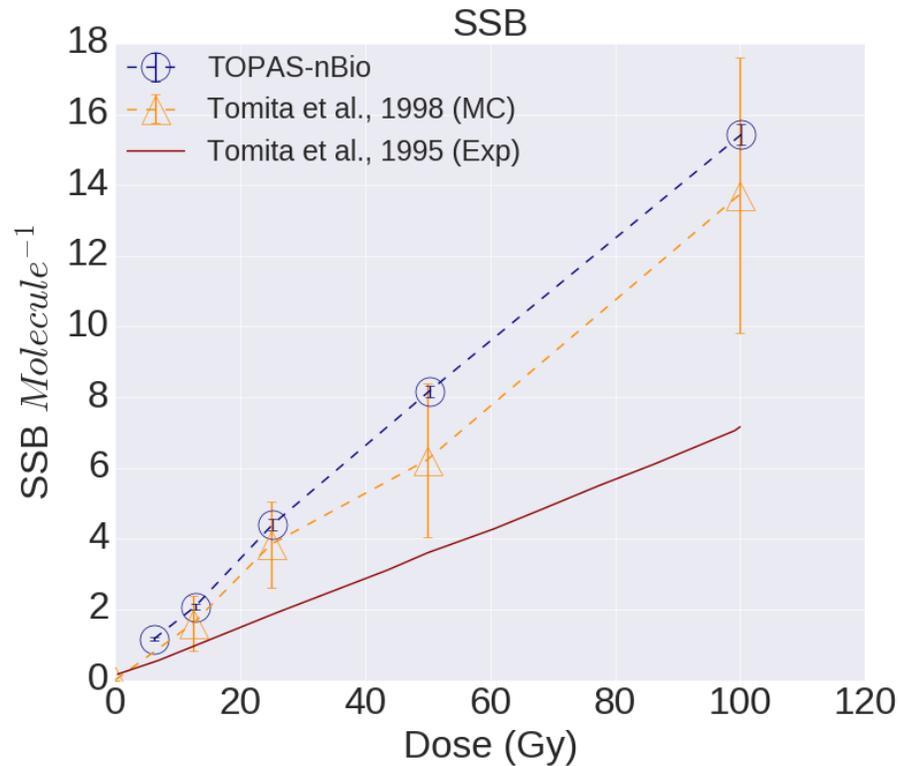


Figura 4.13: Resultados obtenidos en la tercera simulación para los SSB, las marcas circulares son los datos simulados con el método desarrollado, las marcas triangulares son los datos simulados por Tomita, y la línea son los datos experimentales del mismo autor, se tienen configuraciones para 6 Gy, 12.5 Gy, 25 Gy, 50 Gy y 100 Gy, y una incertidumbre estadística de $\pm 0.5\%$, $\pm 0.4\%$, $\pm 0.1\%$, $\pm 0.1\%$, $\pm 0.03\%$ respectivamente para el método utilizado en esta tesis. Para Tomita, sus errores para las dosis de 12.5 Gy, 25 Gy, 50 Gy y 100 Gy es de $\pm 48\%$, $\pm 32\%$, $\pm 35\%$, $\pm 28\%$ respectivamente.

0.5% para los SSB y debajo de 8% para los DSB comparados con los 48% de Tomita para los SSB, por lo tanto, la herramienta Monte Carlo en Topas-nBio permite la simulación de SSB y DSB de manera eficiente. En promedio, cada simulación tomó por repetición: 0.82 hrs (6 Gy), 1.13 hrs (12.5 Gy), 1.69 hr (25 Gy), 3.51 hrs (50 Gy), 11.16 hrs (100Gy). Desde el punto de vista de la simulación Monte Carlo, como se mostró en la sección 4.3, el modelo geométrico del ADN tiene un impacto significativo en el número de SSB y DSB, por lo tanto, más investigación en la configuración geométrica puede ayudar a mejorar la exactitud de las simulaciones. Además, los plásmidos en el arreglo experimental fueron tratados con una concentración de NaCl de 0.01 mol dm^{-3} , que altera el pH del medio e incrementa el número de torcimiento de los plásmidos [59] [56] y el número de SSB

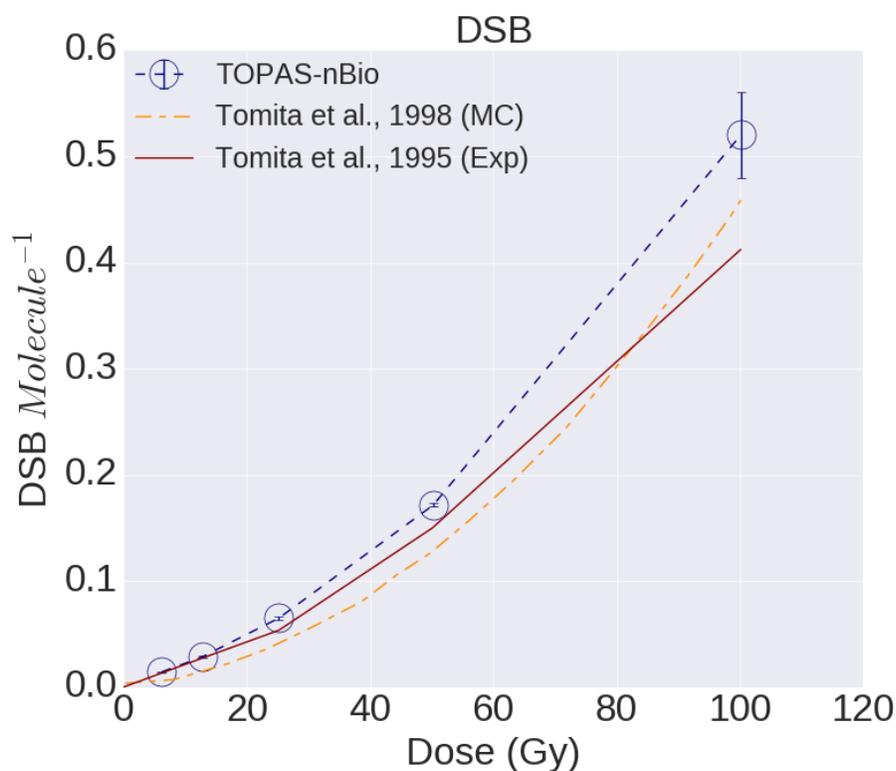


Figura 4.14: Resultados obtenidos en la tercera simulación para los DSB, en este caso se tienen los datos experimentales y simulados de Tomita y los datos obtenidos en este trabajo, las incertidumbres estadísticas de esta parte para nuestros datos son de: $\pm 3.6\%$, $\pm 2.9\%$, $\pm 2.4\%$, $\pm 1\%$, $\pm 7.7\%$, para las dosis de 6 Gy, 12.5 Gy, 25 Gy, 50 Gy y 100 Gy respectivamente. Tomita no reporta errores para esta parte y la gráfica que reporta es posiblemente un ajuste de sus datos.

(ver sección 4.6). Esta característica no fue implementada en la herramienta actual, sin embargo, es considerada como futura y necesaria implementación que puede mejorar la exactitud del modelo geométrico. De la misma manera, ya la probabilidad de que alguna reacción $\bullet OH + ADN(\text{nucleótido})$ genere un SSB es un valor que continúa en investigación [20] [15] y esta se dejó como parámetro libre en la implementación final de la herramienta desarrollada en este trabajo, por lo que un estudio de sensibilidad a este parámetro permitirá entender su influencia en el número de SSB y DSB. Por último, solo se utilizó una lista de física de Geant4-DNA que incluye modelos físicos elásticos e inelásticos para electrones, aunque la investigación y desarrollo de modelos físicos es un tema de interés actualmente [6] [34]. Todos estos parámetros están disponibles en TOPAS-nBio y le permitirán a usuarios experimentados estudiar detalladamente diversas configuraciones (fuera

del objetivo de este trabajo de tesis) que eventualmente mejoren la comparación con datos experimentales a los presentados aquí.

4.6. Prueba de Configuración del ADN

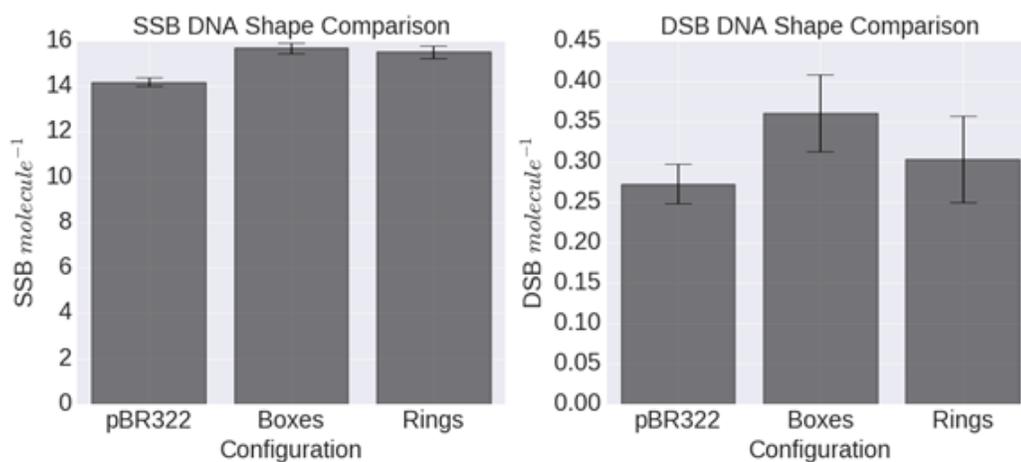


Figura 4.15: Resultados de la diferencia de configuración del ADN, tenemos los rompimientos simples para los plásmidos pBR322, la configuración de cajas rectangulares y de anillos de ADN. Se tienen diferencias significativas para los plásmidos en los rompimientos simples con respecto a los anillos y las cajas, en los rompimientos dobles las diferencias solo son significativas con respecto a las cajas, las desviaciones estándar de cada configuración son de: $\pm 1.3\%$, $\pm 1.5\%$ y $\pm 1.8\%$ para los rompimientos simples en sus configuraciones de pBR322, cajas y anillos respectivamente. Para los rompimientos dobles tenemos: $\pm 9\%$, $\pm 13.1\%$, y $\pm 17.6\%$.

Como se mencionó en la sección anterior, el grado de súper enrollamiento puede afectar el número de SSB. Para mostrar esta suposición, se realizaron tres simulaciones utilizando diferentes configuraciones para contabilizar los SSB y DSB: 1) seis plásmidos súper enrollados (pBR322); 2) seis plásmidos con forma rectangular de 4362 bps, 1×71 segmentos de 10.2 nm (Boxes) y 3) seis plásmidos circulares de 4362 bps (Rings). Los resultados se muestran en la figura 28. En estas simulaciones no se utilizaron disolventes químicos. La dosis prescrita fue de 50 Gy. En la Figura 4.15 se muestran diferencias fuera de la incertidumbre estadística combinada para SSB de pBR322 comparado con Boxes y Rings, mientras que en los DSB dobles no hay una clara tendencia debido a la gran incertidumbre estadística ($>9\%$).

Capítulo 5

Conclusiones

En esta tesis de maestría, se satisficieron los objetivos propuestos. Se concluye lo siguiente:

- Se desarrolló un modelo geométrico de elipsoides y atómico para las bases par de ADN en TOPAS-nBio que evita el uso de solidos booleanos de Geant4.
- Se desarrolló una herramienta Monte Carlo para la generación de plásmidos súper enrollados compatible con TOPAS-nBio y complementada con el modelo de elipsoides y atómico del ADN antes mencionado y los procesos de reacción difusión de especies químicas.
- Se comparó el modelo de elipsoides ADN con el modelo de D. E. Charlton. El modelo de elipsoides produce $29\% \pm 0.1\%$ menos de SSB y $21.25\% \pm 1.5\%$ menos de DSB que el modelo de Charlton.
- Se desarrolló un scorer de rompimientos SSB y DSB usando las extensiones TOPAS compatible con la etapa química de la radiólisis del agua usando el algoritmo IRT.
- Se validó la simulación de radiólisis del agua utilizando ^{60}Co y el método de interacción forzada para gammas. La concordancia con datos experimentales fue satisfactoria.
- Se llevo a cabo una comparación de la simulación de un sistema de seis plásmidos

pBR322 a 273 K en presencia de solventes e irradiada por ^{60}Co con datos de la literatura. El número de SSB totales (directos e indirectos) sobre estimó los datos experimentales de $111.6\% \pm 9.3\%$ hasta en un $117.5\% \pm 4.2\%$, mientras que para DSB fue desde $6\% \pm 1\%$ hasta $26\% \pm 9\%$. La comparación con datos calculados de la literatura resultó en diferencias de hasta $117.5\% \pm 2\%$ para SSB y hasta $6\% \pm 1\%$ para DSB. En este caso, la respuesta lineal de DSB en función de la dosis fue bien reproducida por los datos de este trabajo de tesis. El tiempo de simulación fue de 324 hrs a 346 hrs en un solo procesador a 2.8 GHz.

Como trabajo futuro a esta tesis, se estudiará el efecto en el número de rompimientos al variar los modelos físicos de los procesos de interacción inelástica y elástica, además del efecto de acoplamiento electrónico y vibración. Además, esta se extenderva la herramienta para el estudio del número de torcimiento de los plásmidos contra el número de rompimientos y la contribución del daño complejo con la base. Sin embargo, el impacto de este trabajo es que las herramientas desarrolladas en esta tesis se encontrarán disponibles en TOPAS-nBio, y esto facilitará estudios para radiobiología de los efectos de la radiación en el ADN a etapas tempranas. Entonces, un experto en radiobiología podrá utilizar estas herramientas para interpretar datos experimentales y diseñar experimentos.

Apéndice A

Monte Carlo

A.1. Historia del Método Monte Carlo

El método Monte Carlo tiene sus inicios en 1777 cuando el Conde de Buffon utilizó un método muy similar para resolver el problema de intersección de agujas con un patrón paralelo de líneas, en este problema, las líneas estaban separadas a una cierta distancia D . El Conde lanzó agujas sobre estas líneas, dependiendo del ángulo con el que cayeron dedujo una expresión para calcular la probabilidad de que las agujas cruzaran las líneas o no. Las agujas cruzarían dichas líneas solo si: $y \leq L \sin(\theta)$. Por lo que probabilidad de que una aguja cruzara una línea es de: $P = \frac{2}{L} \pi D$.

El formalismo del método de Monte Carlo fue desarrollado durante la Segunda Guerra Mundial en el Proyecto Manhattan, motivado por la dificultad de resolver las ecuaciones integro-diferenciales del transporte de neutrones. A simples palabras, el Método de Monte Carlo fue utilizado para el cálculo numérico de integrales. Stanislaw Ulam acuñó el nombre Monte Carlo en el artículo seminal (Metrópolis y Ulam, 1949) ser este, el nombre del casino más famoso durante ese periodo.

A palabras de Alex F. Bielajew en el libro *Fundamentals of the Monte Carlo method for neutral and charged particle transport*, el método de Monte Carlo “Representa un intento para modelar la naturaleza a través de una simulación directa de la dinámica esencial

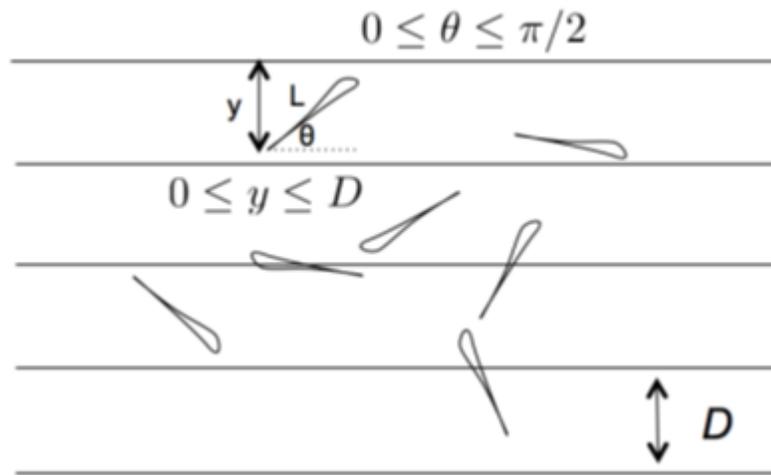


Figura A.1: Imagen ilustrativa del problema de la intersección de las agujas con líneas.

de un sistema en cuestión. En este sentido, el método Monte Carlo es esencialmente una aproximación - una solución a un sistema macroscópico a través de una simulación de sus interacciones microscópicas.”

Para usar este método, se necesita:

1. Conocer la función de densidad de probabilidad de un cierto proceso
2. Conocer las técnicas de muestreo aleatorio adecuadas

A partir de estos dos puntos, se puede simular un proceso físico usando un gran número de repeticiones. Este es el talón de Aquiles de Monte Carlo, debido al múltiple uso de muestreos aleatorios, el tiempo de computo puede ser demasiado grande (horas, días, meses,...). Sin embargo, la flexibilidad que brinda este método para el uso de geometrías complejas y un gran número de procesos físicos simultáneos, hacen que el método Monte Carlo sea de gran utilidad, generalmente cuando el método estocástico clásico falla en reproducir resultados en concordancia con datos experimentales, o simplemente por la falta de técnicas de resolución numéricas para procesar los modelos matemáticos.

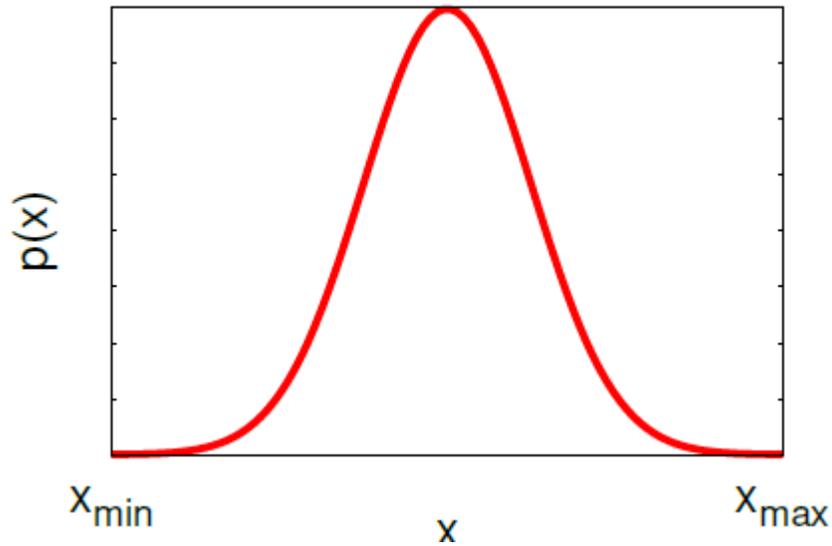


Figura A.2: Ejemplo de una función de densidad de probabilidad.

A.2. Básicos del Método Monte Carlo

En física, una función de densidad de probabilidad (FDP) puede representar la propiedad física más probable de una partícula: posición espacial, energía cinética inicial o después de interactuar, momento, etcétera. Estas se pueden obtener a partir de modelos teóricos o parametrizaciones de modelos y datos experimentales, tales como las secciones eficaces, ángulos de dispersión característicos, etcétera.

Las condiciones que una función tiene que cumplir para que sea FDP son tres:

- Ser positiva, $P(x) \geq 0$.
- Estar normalizadas, tal que $\int P(x)dx = 1$ en todo su dominio.
- Debe de estar acotada.

A.3. Técnicas de Muestreo

Como se ha venido mencionando, las técnicas de muestreo son vitales para el funcionamiento apropiado del método Monte Carlo.

Las técnicas de muestreo dependen de que se conozca la función de distribución de Probabilidad exacta $P(x)$ para un proceso dado, o se tenga una distribución Discreta del proceso con su respectiva probabilidad. Como ejemplo tenemos los espectros de emisión de diferentes fuentes radiactivas. Estos se obtienen tras llevar cabo varias mediciones de la energía de las partículas emitidas por la fuente en un detector (sea una cámara de ionización, estado sólido, etc), haciendo un histograma de estos datos y normalizando para que la integral en todo el dominio sea unitaria.

Sin embargo, existen funciones de distribución que no están acotadas o que son difíciles de trabajar analíticamente, por lo que se usa la función cumulativa de probabilidad (CDF por sus siglas en inglés), esta se define como:

$$c(x) = \int_{x_{min}}^x dx' p(x') \quad (\text{A.1})$$

donde $p(x')$ es una FDP. $c(x)$ tiene las siguientes propiedades:

- $p(x) = \frac{dc(x)}{dx}$
- $c(x_{max}) = 1$
- $c(x_{min}) = 0$

De manera que siempre se puede recurrir a estas funciones. Un ejemplo de FDP es la siguiente ecuación:

$$p(z) = \mu e^{-\mu z} \quad (\text{A.2})$$

esta es la ecuación que determina la distancia que viaja un fotón (z) antes de tener una interacción (*mean free-path*). Dicha función de distribución de probabilidad no está acotada, pero usando su CDF, el problema de muestreo de esta FDP se facilita.

A.4. Generadores de Números Aleatorios

Los números aleatorios (estrictamente pseudo-aleatorios) utilizados por códigos Monte Carlo son secuencias de números no relacionados generados por una ecuación iterativa.

Los buenos generadores de números aleatorios requieren de:

1. Repetibilidad solo tras un gran número de iteraciones.
2. Ser secuencias no uniformemente distribuidas.
3. Deben ser reproducibles a partir de la misma semilla inicial.

Un generador de números aleatorios es una función, que toman como punto de partida una semilla aleatoria y a partir de ella comienza un proceso iterativo, por ejemplo:

$$R_{k+1} = (R_k + C) \text{mod} M \quad (\text{A.3})$$

Esta ecuación funciona si M es un número increíblemente grande. La semilla aleatoria sería el primer valor de R y de ahí se podrían generar los demás. Por supuesto, los generadores de números aleatorios más sofisticados no usan este método como lo es el algoritmo RANECU, sin embargo, en este trabajo no se profundiza en el tema.

En la práctica los generadores de números aleatorios producen números uniformemente distribuidos entre $[0,1]$.

Como se necesita que estos números se apeguen a los procesos físicos de interés, no pueden ser uniformes, así que utilizando técnicas de muestreo es como se obtienen estos comportamientos, por ejemplo, usando:

- Funciones de distribución analíticas (modelos teóricos).
- Distribuciones tabuladas (datos experimentales).

A.5. Muestreo Aleatorio

Como se describió en secciones pasadas, generar números aleatorios de manera uniforme no es adecuado para representar a los procesos físicos que uno desea simular mediante el método Monte Carlo. Por lo que es necesario convertir dichas secuencias de números uniformemente distribuidos en números que sigan funciones de distribución de probabilidades. Los dos métodos de muestreo más utilizados son el método directo y el método

de rechazo, cada uno usado en diferentes circunstancias, más métodos de muestreo se desarrollaron a partir de estos dos.

El primero, el método directo, supone que se conoce la función $c(x)$ y su función analítica inversa $c^{-1}(x)$. Se asigna un número aleatorio uniformemente distribuido r a $c(x)$, tal que $c(x) = r$, finalmente se resuelve para x , siendo esta última el valor de interés a obtener. Claro está, que a veces $c(x)$ no es fácil de invertir o simplemente no se tiene un modelo analítico físico, sino, una colección de datos experimentales de los que se puede recuperar una distribución (por ejemplo, secciones eficaces), en dichos casos se usa el método de rechazo. El algoritmo general de este método es el siguiente:

1. Se genera un número aleatorio uniformemente distribuido r_1 entre $[a, b]$ (siendo a y b el intervalo donde está definida la función de distribución).
2. Se genera otro número aleatorio uniformemente distribuido r_2 entre $[0, 1]$.
3. Se evalúa: $r_2 < p(r_1)/p_{max}$, donde p_{max} es el máximo de la función $p(x)$.
4. Si se cumple la condición 3) se acepta r_1 , en caso contrario se va a 1).

Note que en este método se puede usar una distribución discreta, a diferencia del método directo, por lo que es más flexible pero más costoso computacionalmente, ya que, debido al uso de dos números aleatorios y la posibilidad de rechazar los números, este método toma más tiempo de cómputo que el método directo por lo que es preferible usarlo solo en el caso de que no nos sea posible usar el primer método. El método de rechazo es el método Monte Carlo más conocido, y los demás métodos que se emplean hoy en día son derivados del método directo y del método de rechazo.

A.6. Incertidumbres del Método Monte Carlo

Debido al uso de números aleatorios, y simulaciones de procesos aproximados, existen incertidumbres (también llamados errores estadísticos) en las simulaciones Monte Carlo. Estas incertidumbres pueden ser de dos tipos: estadísticas y sistemáticas.

Las del primer tipo son causadas por la falta de historias simuladas y son erradicadas

incrementando el número de historias, sin embargo, aumentar el número de historias simuladas ocasionara que el tiempo de simulación también aumente. Para evitar esto se usan “técnicas de disminución de varianzas “aproximaciones para el incremento de la eficiencia”, sin embargo, estas últimas sesgan de cierto modo los resultados finales; otro método es usar equipo de cómputo más avanzado o el uso de multithreading (programación paralela).

El segundo tipo de incertidumbres suceden por falta de detalle en las simulaciones o por la divergencia de las aproximaciones para los procesos físicos. Estas no pueden erradicarse, pero pueden mitigarse, por ejemplo, refinando los modelos físicos usados o el detalle geométrico de la simulación (dimensiones, materiales).

Apéndice B

Tablas de Reacciones Químicas

La información para llevar a cabo la radiólisis, difusión y reacción de las especies químicas generadas utilizadas en este trabajo, se encuentra en este apéndice.

Cuando hablamos de radiólisis del agua, se comienza con la etapa pre-química (de 1 fs - 1ps) en la cual los procesos disociativos o de decaimiento de las moléculas de agua ocurren debido al depósito de energía por parte de la radiación mediante ionizaciones y excitaciones. Los procesos disociativos utilizados por defecto en Geant4-DNA y por ende en TOPAS-nBio se encuentran en el cuadro B.1. Estos procesos pueden generar distintas especies químicas dependiendo del proceso físico considerando hasta cinco niveles orbitales para excitación e ionización, para cada caso se tiene una probabilidad de disociación.

Para la simulación de la parte química de la radiólisis (de 1ps - 1 μ s) , el transporte

Proceso			Probabilidad
Estado de Ionización	Decaimiento Disociativo	$H_3O^+ + \bullet OH$	100
A^1B_1	Decaimiento Disociativo	$\bullet OH + H\bullet$	65
	Relajación	$H_2O + \Delta E$	35
B^1A_1	Auto-Ionización	$H_3O^+ + \bullet OH + e_{aq}^-$	55
	Auto-Ionización	$\bullet OH + \bullet OH + H_2$	15
	Relajación	$H_2O + \Delta E$	30
Rydberg, bandas de difusión	Auto-Ionización	$H_3O^+ + \bullet OH + e_{aq}^-$	50
	Relajación	$H_2O + \Delta E$	50

Cuadro B.1: Procesos de disociación de especies químicas utilizados en este trabajo para la radiólisis del agua, donde se muestran los procesos físicos, los productos generados en tales procesos y la probabilidad de ocurrencia.

CAPÍTULO B TABLAS DE REACCIONES QUÍMICAS

Especies Químicas	Nombre en Topas-nBio	Coeficientes de Difusión D ($10^{-9} m^2 s^{-1}$) a 25° C		Radio de Reacción (nm)
		Geant4-DNA Default	TOPAS-nBio	
e_{aq}^-	SolvatedElectron	4.9	4.9	0.5
$\bullet OH$	Hydroxil	2.8	2.2	0.22
$H\bullet$	Hydrogen	7.0	7.0	0.19
H_3O^+	Hydronium	9.0	9.46	0.25
H_2	Dihydrogen	5.0	4.8	0.14
OH^-	Hydroxide	5.0	5.3	0.33
H_2O_2	HydrogenPeroxide	1.4	2.3	0.21
O_2	Oxygen	-	2.4	0.17
O_2^-	SuperoxideAnion	-	1.75	0.22
HO_2	HydroPeroxide	-	2.3	0.21
HO_2^-	Dioxidanide	-	1.4	0.25

Cuadro B.2: Coeficientes de difusión para las especies químicas generadas en la radiólisis del agua.

Reacción	Constante de Velocidad de Reacción k_{obs} ($10^{10} M^{-1} s^{-1}$)	
	Geant4-DNA Default	TOPAS-nBio
$e_{aq}^- + e_{aq}^- \rightarrow H_2 + 2OH^-$	0.5	0.636
$e_{aq}^- + \bullet OH \rightarrow OH^-$	2.95	2.95
$e_{aq}^- + H\bullet \rightarrow H_2 + OH^-$	2.65	2.5
$e_{aq}^- + H_3O^+ \rightarrow H\bullet$	2.11	2.11
$e_{aq}^- + H_2O_2 \rightarrow OH^- + \bullet OH$	1.41	1.10
$\bullet OH + \bullet OH \rightarrow H_2O_2$	0.44	0.550
$\bullet OH + H\bullet \rightarrow H_2O$	1.44	1.55
$H\bullet + H\bullet \rightarrow H_2$	1.2	0.503
$H_3O^+ + OH^- \rightarrow H_2O$	14.3	11.3

Cuadro B.3: Reacciones químicas disponibles en Geant4-DNA y TOPAS-nBio para la radiólisis del agua.

usando movimiento Browniano requiere de las contantes de difusión, los valores usados en este trabajo se muestran en el cuadro B.2.

Finalmente, las reacciones químicas posibles para la radiólisis del agua usadas en este trabajo se muestran en el cuadro B.3. Esta información se encuentra con más detalles y con su respectiva validación en el artículo [50] y no entraremos en mayor discusión respecto a este tema.

Bibliografía

- [1] Agostinelli, S., Allison, J., Amako, K., Apostolakis, J., Araujo, H., Arce, P., ... Zschiesche, D. (2003). GEANT4 - A simulation toolkit. *Nuclear Instruments and Methods in Physics Research, Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, 506(3), 250-303. [https://doi.org/10.1016/S0168-9002\(03\)01368-8](https://doi.org/10.1016/S0168-9002(03)01368-8)
- [2] Alexander, J. W. (1928). Topological Invariants of Knots and Links. *Transactions of the American Mathematica Society*, 30(2), 275-306.
- [3] Attix, F. H. (1986). *Introduction to Radiological Physics and Radiation Dosimetry*. <https://doi.org/10.1002/9783527617135>
- [4] M. J. Berger M. Inokuti H. H. Anderson H. Bichsel J. A. Dennis D. Powers S. M. Seltzer J. E. Turner. Report37. *Journal of the International Commission on Radiation Units and Measurements, Volume os19, Issue 2, 1 December 1984, Pages NP*, <https://doi.org/10.1093/jicru/os19.2.Report37>
- [5] Bielajew, A. F., & Rogers, D. W. O. (1988). Variance-reduction techniques. In T. M. Jenkins, W. R. Nelson, & A. Rindi (Eds.), *Monte Carlo transport of electrons and photons* (p. 407). Ettore Majorana International Science Series.
- [6] Bordage, M. C., Bordes, J., Edel, S., Terrissol, M., Franceries, X., Bardies, M., ... Incerti, S. (2016). Implementation of new physics models for low energy electrons in liquid water in Geant4-DNA. *Physica Medica*, 32(12), 1833-1840. <https://doi.org/10.1016/j.ejmp.2016.10.006>
- [7] Buxton, G. V. (1972). Nanosecond Pulse Radiolysis of Aqueous Solutions Containing Proton and Hydroxyl Radical Scavengers. *Proceedings of the Royal*

-
- Society A: Mathematical, Physical and Engineering Sciences, 328(1572), 9-21.
<https://doi.org/10.1098/rspa.1972.0065>
- [8] Casali, N., & Preston, A. (2003). E.coli Plasmid Vectors. *Methods in molecular biology* (Vol. 253). Totowa, New Jersey: Humana Press. https://doi.org/10.1007/978-1-60761-820-1_12
- [9] Champion, C. (2013). Quantum-mechanical predictions of electron-induced ionization cross sections of DNA components. *Journal of Chemical Physics*, 138(18). <https://doi.org/10.1063/1.4802962>
- [10] Champion, C., Incerti, S., Perrot, Y., Delorme, R., Bordage, M. C., BardiÃ's, M., ... Campos, L. (2014). Dose point kernels in liquid water: An intra-comparison between GEANT4-DNA and a variety of Monte Carlo codes. *Applied Radiation and Isotopes*, 83, 137-141. <https://doi.org/10.1016/j.apradiso.2013.01.037>
- [11] Charlton, D. E., Nikjoo, H., & Humm, J. L. (1989). Calculation of Initial Yields of Single-Strand and Double-Strand Breaks in Cell-Nuclei From Electrons, Protons and Alpha-Particles. *International Journal of Radiation Biology*, 56(1), 1-19. <https://doi.org/10.1080/09553008914551141>
- [12] Clifford, P., Green, N. J. B., Oldfield, M. J., Pilling, M. J., & Pimblott, S. M. (1986). Stochastic Models of Multi-species Kinetics in Radiation-induced Spurs. *J. Chem. Soc., Faraday Trans. 1*, 82, 2673-2689. <https://doi.org/10.1039/F19868202673>
- [13] Conte V, Selva A, Colautti P, Hilgers G and Rabus H 2017 Track structure characterization and its link to radiobiology *Radiat. Meas.* 106 506-11 Online: <http://linkinghub.elsevier.com/retrieve/pii/S1350448717304274>
- [14] Cozzarelli, A. V. V. and N. R. (1994). Conformational And Thermodynamic Properties of Supercoiled DNA. *Annu. Rev. Biophys. Struct.*, (23).
- [15] de la Fuente Rosales, L., Incerti, S., Francis, Z., & Bernal, M. A. (2018). Accounting for radiation-induced indirect damage on DNA with the Geant 4-DNA code. *Physica Medica*, 51(June), 108-116. <https://doi.org/10.1016/j.ejmp.2018.06.006>
-

-
- [16] Dizdaroglu, M., & Jaruga, P. (2012). Mechanisms of free radical-induced damage to DNA. *Free Radical Research*, 46(4), 382-419. <https://doi.org/10.3109/10715762.2011.653969>
- [17] Emfietzoglou, D., Cucinotta, F., & Nikjoo, H. (2005). A complete dielectric response model for liquid water: a solution of the Bethe ridge problem. *Radiation Research*, 164(2), 202-211. <https://doi.org/10.1667/RR3399>
- [18] F. Lory, P. J. (1953). *Principles of polymer chemistry* (1st ed.). New York: Cornell University Press. <https://doi.org/10.1002/pen.760050321>
- [19] Francis, Z., Incerti, S., Karamitros, M., Tran, H. N., & Villagrasa, C. (2011). Stopping power and ranges of electrons, protons and alpha particles in liquid water using the Geant4-DNA package. *Nuclear Instruments and Methods in Physics Research, Section B: Beam Interactions with Materials and Atoms*, 269(20), 2307-2311. <https://doi.org/10.1016/j.nimb.2011.02.031>
- [20] Friedland, W., Schmitt, E., Kunderát, P., Dingfelder, M., Baiocco, G., Barbieri, S., & Ottolenghi, A. (2017). Comprehensive track-structure based evaluation of DNA damage by light ions from radiotherapy-relevant energies down to stopping. *Scientific Reports*, 7(September 2016), 1-15. <https://doi.org/10.1038/srep45161>
- [21] Frongillo, Y., Goulet, T., Fraser, M.-J., Cobut, V., Patau, J. P., & Jay-Gerin, J.-P. (1998). Monte Carlo Simulation of Fast Electron And Proton Tracks In Liquid Water - II. Nonhomogeneous Chemistry. *Radiation Physics and Chemistry*, 51(3), 245-254. [https://doi.org/10.1016/S0969-806X\(97\)00097-2](https://doi.org/10.1016/S0969-806X(97)00097-2)
- [22] Fujimoto, B. S., & Schurr, J. M. (2002). Monte carlo simulations of supercoiled DNAs confined to a plane. *Biophysical Journal*, 82(2), 944-962. [https://doi.org/10.1016/S0006-3495\(02\)75455-6](https://doi.org/10.1016/S0006-3495(02)75455-6)
- [23] Harris, B. A., & Harvey, S. C. (1999). Program for analyzing knots represented by polygonal paths. *Journal of Computational Chemistry*, 20(8), 813-818. [https://doi.org/10.1002/\(SICI\)1096-987X\(199906\)20:8<813::AID-JCC7>3.0.CO;2-I](https://doi.org/10.1002/(SICI)1096-987X(199906)20:8<813::AID-JCC7>3.0.CO;2-I)
- [24] Huerta, E. V. (n.d.). El polinomio de Alexander como invariante de Nudos.
-

-
- [25] Incerti, S., Ivanchenko, A., Karamitros, M., Mantero, A., Moretto, P., Tran, H. N., ... Zacharatou, C. (2010). Comparison of GEANT4 very low energy cross section models with experimental data in water. *Medical Physics*, 37(9), 4692-4708. <https://doi.org/10.1118/1.3476457>
- [26] J. Fullford, H. Nikjoo, D. T. G. & P. O. (2001). Yields of SSB and DSB induced in DNA by Al K ultrasoft X-rays and α -particles comparison of experimental and simulated yields. *International Journal of Radiation Biology*, 77(10), 1053-1066.
- [27] James D. Watson, A. B. (1923). *DNA The Secret Of Life*. New York: Knopf, Alfred A.
- [28] Karamitros, M., Luan, S., Bernal, M. A., Allison, J., Baldacchino, G., Davidkova, M., ... Incerti, S. (2014). Diffusion-controlled reactions modeling in Geant4-DNA. *Journal of Computational Physics*, 274, 841-882. <https://doi.org/10.1016/j.jcp.2014.06.011>
- [29] Karamitros, M., Mantero, A., Incerti, S., Friedland, W., Baldacchino, G., Barberet, P., ... Zacharatou, C. (2011). Modeling Radiation Chemistry in the Geant4 Toolkit. *Progress in Nuclear Science and Technology*, 2(0), 503-508. <https://doi.org/10.15669/pnst.2.503>
- [30] Klenin, K., & Langowski, J. (2000). Computation of writhe in modeling of supercoiled DNA. *Biopolymers*, 54(5), 307-317. [https://doi.org/10.1002/1097-0282\(20001015\)54:5;307::AID-BIP20;3.0.CO;2-Y](https://doi.org/10.1002/1097-0282(20001015)54:5;307::AID-BIP20;3.0.CO;2-Y)
- [31] Konstantin V. Klenin, Alexander V. Vologodskii, Vadim V. Anshelevich, A. M. D. and M. D. F. (1991). Computer Simulation of DNA Supercoiling. *Methods in Enzymology*, 203(C), 403-432. [https://doi.org/10.1016/0076-6879\(91\)03023-A](https://doi.org/10.1016/0076-6879(91)03023-A)
- [32] Kreipl, M. S., Friedland, W., & Paretzke, H. G. (2009). Time- and space-resolved Monte Carlo study of water radiolysis for photon, electron and ion irradiation. *Radiation and Environmental Biophysics*, 48(1), 11-20. <https://doi.org/10.1007/s00411-008-0194-8>
- [33] Kummerle, E. A, E. P. (2004). A computer-generated supercoiled model of the pUC19 plasmid. *Eur Biophys*, 34(13-18), 6.
-

-
- [34] Kyriakou I., Sefl M., Nourry V. and Incerti S. (2016). The impact of new Geant4-DNA cross section models on electron track structure simulations in liquid water J. Appl. Phys. 119 194902 Online: <http://aip.scitation.org/doi/10.1063/1.4950808>
- [35] LaVerne, J. A. (2000). OH Radicals and Oxidizing Products in the Gamma Radiolysis of Water. *Radiation Research*, 153(2), 196-200. [https://doi.org/10.1667/0033-7587\(2000\)153\[0196:ORAOPI\]2.0.CO;2](https://doi.org/10.1667/0033-7587(2000)153[0196:ORAOPI]2.0.CO;2)
- [36] LaVerne, J. A., & Pimblott, S. M. (1993). Yields of Hydroxyl Radical and Hydrated Electron Scavenging Reactions in Aqueous Solutions of Biological Interest. *Radiation Research*, 135(1), 16. <https://doi.org/10.2307/3578391>
- Lemaire, D. G. E., Bothe, E., & Schulte-Frohlinde, D. (1984). Yields of radiation-induced main chain scission of poly u in aqueous solution: Strand break formation via base radicals. *International Journal of Radiation Biology*, 45(4), 351-358. <https://doi.org/10.1080/09553008414550491>
- [37] Lemaire, D. G. E., Bothe, E., & Schulte-Frohlinde, D. (1984). Yields of radiation-induced main chain scission of poly u in aqueous solution: Strand break formation via base radicals. *International Journal of Radiation Biology*, 45(4), 351-358. <https://doi.org/10.1080/09553008414550491>
- [38] McNamara, A., Geng, C., Turner, R., Mendez, J. R., Perl, J., Held, K., ... Schuemann, J. (2017). Validation of the radiobiology toolkit TOPAS-nBio in simple DNA geometries. *Physica Medica*, 33, 207-215. <https://doi.org/10.1016/j.ejmp.2016.12.010>
- [39] Melton, C. E. (1972). Cross Sections and Interpretation of Dissociative Attachment Reactions Producing OH⁻, O⁻, and H⁻ in H₂O. *The Journal of Chemical Physics*, 57(10), 4218-4225. <https://doi.org/10.1063/1.1678051>
- [40] Meesungnoen, J., Benrahmoune, M., Filali-Mouhim, A., Mankhetkorn, S., & Jay-Gerin, J. P. (2001). Monte Carlo calculation of the primary radical and molecular yields of liquid water radiolysis in the linear energy transfer range 0.3-6.5 keV/micrometer: application to ¹³⁷Cs gamma rays. *Radiation Research*, 155(2), 269-278. [https://doi.org/10.1667/0033-7587\(2001\)155\[0269:MCCOTP\]2.0.CO;2](https://doi.org/10.1667/0033-7587(2001)155[0269:MCCOTP]2.0.CO;2)
- [41] Meylan, S., Incerti, S., Karamitros, M., Tang, N., Bueno, M., Clairand, I., & Villagrasa, C. (2017). Simulation of early DNA damage after the irradiation
-

-
- of a fibroblast cell nucleus using Geant4-DNA. *Scientific Reports*, 7(1), 1-15. <https://doi.org/10.1038/s41598-017-11851-4>
- [42] Meylan, S., Vimont, U., Incerti, S., Clairand, I., & Villagrasa, C. (2016). Geant4-DNA simulations using complex DNA geometries generated by the DnaFabric tool. *Computer Physics Communications*, 204, 159-169. <https://doi.org/10.1016/j.cpc.2016.02.019>
- [43] Michaud, M., Wen, A., & Sanche, L. (2003). Cross Sections for Low-Energy (1-100 eV) Electron Elastic and Inelastic Scattering in Amorphous Ice. *Radiation Research*, 159(1), 3-22. [https://doi.org/10.1667/0033-7587\(2003\)159\[0003:CSFLEE\]2.0.CO;2](https://doi.org/10.1667/0033-7587(2003)159[0003:CSFLEE]2.0.CO;2)
- [44] Nelson, W. R., Hirayama, H., & Rogers, D. (1985). The EGS4 Code System. Report SLAC-265, Stanford Linear Accelerator Centre, Stanford, California, 265. [https://doi.org/10.1002/1098-108X\(198707\)6:4<455::AID-EAT2260060402>3.0.CO;2-5](https://doi.org/10.1002/1098-108X(198707)6:4<455::AID-EAT2260060402>3.0.CO;2-5)
- [45] Nikjoo, H., Uehara, S., Willson, W. E., Hoshi, M., & Goodhead, D. T. (1998). Track structure in radiation biology: theory and applications. *International Journal of Radiation Biology*, 73(4), 355-364. Retrieved from <http://informahealthcare.com/doi/abs/10.1080/095530098142176>
- [46] Nunes Raposo, A. (2015). Geometric Modeling , Simulation , and Visualization Methods for Plasmid DNA Molecules.
- [47] Perl J., Shin J., Schuemann J., Faddegon B., Paganetti H., TOPAS: an innovative proton Monte Carlo platform for research and clinical applications. *Journal of Medical Physics*, 39 (11). Retrieved from <https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3493036/>
- [48] Plante, I. (2011). A Monte-Carlo step-by-step simulation code of the non-homogeneous chemistry of the radiolysis of water and aqueous solutions- Part II: Calculation of radiolytic yields under different conditions of LET, pH, and temperature. *Radiation and Environmental Biophysics*, 50(3), 405-415. <https://doi.org/10.1007/s00411-011-0368-7>
- [49] Prise, K. M., Davies, S., & Michael, B. D. (1993). Evidence for Induction of DNA Double-Strand Breaks at Paired Radical Sites., 134(1), 102-106.
-

-
- [50] Ramos-Méndez, J., Perl, J., Schuemann, J., McNamara, A., Paganetti, H., & Faddegon, B. (2018). Monte Carlo simulation of chemistry following radiolysis with TOPAS-nBio. *Physics in Medicine and Biology*, 63(10), 0-12. <https://doi.org/10.1088/1361-6560/aac04c>
- [51] Salvat, F., & Fern, M. (2015). PENELOPE - A Code System for Monte Carlo Simulation of Electron and Photon Transport ORGANISATION FOR ECONOMIC CO-OPERATION AND DEVELOPMENT. PENELOPE, a Code System for Monte Carlo Simulation of Electron and Photon Transport, (July).
- [52] Schuemann J, McNamara A, Ramos-Méndez J, Perl J, Held K, Paganetti H, Incerti S and Faddegon B. TOPAS-nBio - An extension to the TOPAS simulation toolkit for cellular and sub-cellular radiobiology. *Radiat. Res.* Accepted. (2018)
- [53] Shiraishi, H., Katsumura, Y., Hiroishi, D., Ishigure, K., & Washio, M. (1988). Pulse-radiolysis study on the yield of hydrated electron at elevated temperatures. *Journal of Physical Chemistry*, 92(10), 3011-3017. <https://doi.org/10.1021/j100321a061>
- [54] Sumiyoshi, T., & Katayama, M. (1982). the Yield of Hydrated Electrons At 30 Pico-seconds. *Chemistry Letters*, 11(12), 1887-1890. <https://doi.org/10.1246/cl.1982.1887>
- [55] Tomita, H., Kai, M., Kusama, T., & Ito, A. (1997). Monte Carlo simulation of physicochemical processes of liquid water radiolysis. The effects of dissolved oxygen and OH scavenger. *Radiation and Environmental Biophysics*, 36(2), 105-116. <https://doi.org/10.1007/s004110050061>
- [56] Tomita, H., Kai, M., Kusama, T., & Ito, A. (1998). Monte Carlo simulation of DNA strand-break induction in supercoiled plasmid pBR322 DNA from indirect effects. *Radiation and Environmental Biophysics*, 36(4), 235-241. <https://doi.org/10.1007/s004110050077>
- [57] Tomita, H., Kai, Y., & Kusama, A. T. (1995). Strand Break Formation in Plasmid DNA Irradiated in Aqueous Solution: Effect of Medium Temperature and Hydroxyl Radical Scavenger Concentration. *Radiat. Res.*, 36, 46-55.
- [58] Villegas F, Bäckström G, Tilly N and Ahnesjö A. (2016). Energy deposition clustering as a functional radiation quality descriptor for mo-
-

- deling relative biological effectiveness *Med. Phys.* 43 6322-35 Online:
<http://scitation.aip.org/content/aapm/journal/medphys/43/12/10.1118/1.4966033>
- [59] Vologodskii, A. V., Levene, S. D., Klenin, K. V., Frank-Kamenetskii, M., & Cozzarelli, N. R. (1992). Conformational and thermodynamic properties of supercoiled DNA. *Journal of Molecular Biology*, 227(4), 1224-1243. [https://doi.org/10.1016/0022-2836\(92\)90533-P](https://doi.org/10.1016/0022-2836(92)90533-P)
- [60] Watson, J., & Crick, F. C. (1953). MOLECULAR STRUCTURE OF NUCLEIC ACIDS: A Structure for Deoxyribose Nucleic Acid, (4356), 4356.
- [61] C.J. Werner, et al., MCNP6.2 Release Notes, Los Alamos National Laboratory, report LA-UR-18-20808 (2018).
- [62] Wolff, R. K., Bronskill, M. J., Aldrich, J. E., & Hunt, J. W. (1973). Picosecond pulse radiolysis. IV. Yield of the solvated electron at 30 picoseconds. *Journal of Physical Chemistry*, 77(11), 1350-1355. <https://doi.org/10.1021/j100630a008>