UNIVERSITY OF LILLE
Faculty of Sciences and Technologies
Doctoral School of Engineering Sciences

THESIS

To obtain the degree of

Doctor in Sciences for the University of Lille
Discipline: Biophysics and Medical Imaging

Submitted by

PARVATHY A. P. SARMA

COMPUTER SIMULATION AND EXPERIMENTS OF RADIO-INDUCED CELL SENESCENCE

Thesis Advisor 1: Prof. Fabrizio Cleri
Thesis Advisor 2: Prof. Corinne Abbadie

Thesis defended on **28 March 2025** in front of the Committee composed by:

Prof. Elise Dumont	Ecole Normale Supérieure de Lyon	Reviewer
Dr. Pascale Bertrand	Commissariat à l'Énergie Atomique et aux Énergies Alternatives	Reviewer
Prof. Carla Molteni	King's College London	Examiner
Dr. Nicholas Henthorn	University of Manchester	Examiner
Dr. Ralf Blossey	University of Lille	President
Prof. Fabrizio Cleri	University of Lille	Thesis Advisor
Prof. Corinne Abbadie	University of Lille	Thesis Advisor

UNIVERSITÉ DE LILLE
Faculté des Sciences et Technologies
Ecole Doctorale des Sciences de l'Ingénieur

THESIS

Pour obtenir le diplôme de

Docteur en Sciences de l'Université de Lille
Discipline: Biophysique et Imagerie Médicale

Submitted by

PARVATHY A. P. SARMA

SIMULATION PAR ORDINATEUR ET EXPÉRIENCES DE LA SÉNESCENCE CELLULAIRE RADIO-INDUITE

Thesis Directeur 1: Prof. Fabrizio Cleri
Thesis Directrice 2: Prof. Corinne Abbadie

Thèse soutenue le **28 mars 2025** devant le Comité composé de:

Prof. Elise Dumont	Ecole Normale Supérieure de Lyon	Rapportrice
Dr. Pascale Bertrand	Commissariat à l'énergie atomique et aux énergies alternatives	Rapportrice
Prof. Carla Molteni	Kings College London	Examinatrice
Dr. Nicholas Henthorn	University of Manchester	Examinateur
Dr. Ralf Blossey	University of Lille	President
Prof. Fabrizio Cleri	University of Lille	Directeur de thèse
Prof. Corinne Abbadie	University of Lille	Directrice de thèse

ABSTRACT

Cancer continues to pose a significant societal challenge to researchers and medical professionals due to its high mortality rates, intricate pathophysiology, and profound impact on public health and quality of life. Among various treatment modalities, radiation therapy is widely employed to target and destroy cancerous cells by inflicting molecular damage to the DNA. Damage may include single-strand breaks (SSBs), base damage (BD), and double-strand breaks (DSBs), each one activating a cascade of cellular pathways leading to cell death. However, radiation therapy can also affect surrounding normal cells. The out-of-field dose radiation often results in a higher accumulation of SSBs in these cells. Consequently, the normal cells may enter a state of senescence (a dormant state) with the potential to transition into secondary cancerous cells under certain conditions.

Given this interplay between DNA damage, cellular senescence, and cancer progression, this thesis investigates the molecular and mechanistic aspects of the DNA damage response (DDR) pathway, focusing on the role of SSBs and their repair. By integrating molecular dynamics simulations, mathematical modelling, and experimental validation wherever possible, this work enriches our understanding of DNA repair mechanisms and their implications for cellular senescence and cancer therapy.

After a first chapter that gives the state of the art, the main part of this thesis focuses on how poly-(ADP-ribose)-polymerase-1 (PARP-1) protein recognizes SSBs. Molecular-scale simulations are employed to analyze PARP-1's interaction, firstly with a hairpin DNA fragment containing an SSB, offering novel insights into the early steps of the single-strand break repair (SSBR) pathway. Building on this foundation, the study moves on to the smallest unit of nuclear chromatin, the nucleosome, for which a similar study is carried out, by placing SSBs at different DNA sites around the nucleosome. The important role of the interactions with histone tails will be highlighted. Further, the structural and mechanical effects of SSBs and DSBs within a dinucleosome, that is a pair of nucleosomes joined together with varying "linker"-DNA lengths will be explored. These investigations reveal how different types of DNA damage induce conformational changes in chromatin structure, thereby facilitating the recruitment of DDR enzymes to the damage sites.

The final part of the thesis transitions to a more macroscopic perspective, modelling the SSBR pathway to understand the activation of cellular responses to DNA damage. The study analyses cell-fate decisions under varying damage conditions, by using mathematical

modeling with systems of differential equations describing different chemical kinetics (Hill, Michaelis-Menten, mass-action). Thanks to a brief visit (1 month) to the PRECISE group in Manchester University followed by a nice collaboration, I also employed GEANT4-TOPAS Monte Carlo simulations to study the whole chain of radiation-induced DNA damage, quantifying the different type of DNA damages induced during radiation treatment. The models is further integrated into the DaMaRiS (DNA Mechanistic Repair Simulator) framework, to model the SSBR pathway to simulate DNA repair dynamics comprehensively. This multidisciplinary approach, combining computational simulations, mathematical modelling, and experimental insights, advances our understanding of DNA repair mechanisms.

Keywords: Single-strand breaks; cellular senescence; DNA; PARP-1; molecular dynamics simulations; Monte-Carlo simulations.

RESUMÉ

Le cancer continue de poser un défi social important aux chercheurs et aux professionnels de la santé en raison de ses taux de mortalité élevés, de sa physiopathologie complexe et de son impact profond sur la santé publique et la qualité de vie. Parmi les diverses modalités de traitement, la radiothérapie est largement utilisée pour cibler et détruire les cellules cancéreuses en endommageant l'ADN. Ces dommages comprennent les cassures simple-brin (SSB), les dommages aux bases (BD) et les cassures double-brins (DSB), chacun active une cascade de voies cellulaires conduisant à la mort cellulaire. Cependant, la radiothérapie peut affecter également les cellules normales environnantes, car la dose de rayonnement hors champ entraîne souvent une accumulation plus élevée de SSB dans ces cellules. Par conséquent, les cellules normales peuvent entrer dans un état de sénescence (un état dormant) avec le potentiel de se transformer en cellules cancéreuses secondaires dans certaines conditions. Compte tenu de cette interaction entre les dommages à l'ADN, la sénescence cellulaire et la progression du cancer, cette thèse étudie les aspects moléculaires et mécanistiques de la voie de réponse aux dommages à l'ADN (DDR), en se concentrant sur le rôle des SSB et leur réparation. En intégrant des simulations de dynamique moléculaire, une modélisation mathématique et une validation expérimentale chaque fois que possible, ce travail permet d'approfondir la compréhension des mécanismes de réparation de l'ADN et de leurs implications pour la sénescence cellulaire et la thérapie du cancer.

Après un premier chapitre d'état de l'art, la partie principale de cette thèse se concentre sur la façon dont la protéine poly-(ADP-ribose)-polymérase-1 (PARP-1) reconnaît les SSB. Des simulations à l'échelle moléculaire sont utilisées pour analyser l'interaction de PARP-1, d'abord avec un fragment d'ADN en épingle à cheveux (hairpin) contenant un SSB, ainsi offrant de nouvelles perspectives sur les premières étapes de la voie de réparation des SSB (ou SSBR). En s'appuyant sur cette base, l'étude se poursuit avec. l'unité structurelle élémentaire de la chromatine nucléaire, le nucleosome, sur lequel l'étude precedent est répété, en plaçant des SSBs à des différents endroits tout au long de l'ADN. Le rôle important des queues d'histone sera mis en avant. Ensuite, nous allons explorer les effets structurels et mécaniques des SSB et des DSB dans un dinucléosome, soit deux nucleosomes reliés par des différentes longueurs d'ADN (le "linker"). Ces recherches révèlent comment différents types de dommages à l'ADN induisent des changements conformationnels dans la structure de la

chromatine, facilitant le recrutement dee enzymes de DDR sur les sites endommagés.

La partie finale de la thèse passe à une perspective plus macroscopique, en modélisant la voie SSBR pour comprendre les réponses cellulaires aux dommages à l'ADN. L'étude analyse les décisions relatives au destin cellulaire dans différentes conditions de dommages, par le biais de systèmes d'equations différentielles en utilisant différents modèles de cinétique biochimique (Hiil, Michaelis-Menten, mass-action). Grace à une visite de 1 mois au groupe PRECISE de l'Université de Manchester, suivie d'une belle collaboration, des simulations Monte Carlo GEANT4-TOPAS sont utilisées pour étudier toute la chaine d'endommagenent à l'ADN induits par les radiations, en quantifiant les différents types de dommages à l'ADN induits pendant la radiothérapie. Ces modèles sont ensuite intégrés dans le cadre DaMaRiS (DNA Mechanistic Repair Simulator), pour modéliser la voie SSBR afin de simuler de manière exhaustive la dynamique de réparation de l'ADN. Cette approche multidisciplinaire, combinant simulations informatiques, modélisation mathématique et connaissances expérimentales, fait progresser notre compréhension des mécanismes de réparation de l'ADN.

Mots-clés : cassures simple-brin, sénescence cellulaire; ADN; PARP-1; simulations de dynamique moléculaire; simulations Monte-Carlo.

PUBLICATIONS

Part of the text and figures have appeared in the following publications:

1. **Sarma, P. A.**, Abbadie, C. and Cleri, F. (2024). Cooperative dynamics of PARP-1 zinc-finger domains in the detection of DNA single-strand breaks. *Scientific Reports*, **14**(1), 23257.

2. **Sarma, P. A.**, Abbadie, C., de Launoit, Y. and Cleri, F. (2024). Cell senescence and the DNA single-strand break damage repair pathway. *DNA*, **4**(4), 530-552.

3. Ghediri, S.,**Sarma, P. A.**, Saravanan, V., Abbadie, C., Blossey, R., Cleri, F. (2025).Mechanisms of DNA Damage Recognition by UDG and PARP1 in the Nucleosome. Biomolecules, 15(5), 649.

ACKNOWLEDGEMENTS

I would like to express my heartfelt gratitude to my supervisors, Prof. Fabrizio Cleri and Prof. Corinne Abbadie, for their invaluable guidance, support, and encouragement throughout my research journey. Their expertise, patience, and constructive feedback have been instrumental in shaping this work. I would also like to thank Dr. Nicholas Henthorn, from the PRECISE group at the University of Manchester, for his mentorship and support during our collaborative work. I extend my sincere thanks to all the jury members for accepting the invitation to evaluate my thesis and for their valuable time and insights. This work was made possible through the funding provided by the University of Lille, under the PEARL program ("Programme for Early-stage Researchers in Lille") as part of the SENESIMEX project ("Simulations and Experiments of Radiation-induced Cellular Senescence") led by my supervisors. I also gratefully acknowledge the supporting contributions from the ANR DYPROSOME grant (ANR-21-CE45-0032) and the internal resources of the CNRS-CANTHER laboratory.

I am truly thankful to my colleagues Bilal, Baker, Milan, Tymoffii, and Safwen, for the enriching environment you helped to create—where scientific discussions were always welcome, and everyday lunches turned into joyful moments of laughter and warmth. Your presence made the workdays lighter and much more enjoyable. To my friends from India—Shubhi, VJ, Shivangi, and Rubina—thank you for always being just a call away. Through every anxious night, emotional spiral, or random burst of joy, you stood by me with unwavering support, despite the distance. You have been a constant source of encouragement and companionship, and I cherish every bit of it.

To my family-my parents, sister, and brother -you are the four pillars of my life, the ones who stood firm even when I faltered. Your unconditional love, patience, and belief in me have carried me through the most difficult days and celebrated every small milestone I achieved. Thank you for enduring this journey with me, for holding space for my breakdowns, and for lifting me every single time I thought I couldn't go on. Without your strength behind me, this journey would have remained incomplete.

I would like to acknowledge Parth for introducing me to Lille and for being a part of this journey for a significant time. To Aditi - thank you for being the most consistent and supportive friend I could have ever asked for. You have carried me through storms without a second thought, with kindness, strength, and grace that are nothing short of extraordinary. And finally, to my Kanha-for being the quiet, ever-present source of strength and grace in my life. You have walked

beside me through heartbreaks, chaos, and doubts, giving me more than I ever had the courage to ask for. I continue to seek your guidance and blessings in all that lies ahead.

This thesis is the result of many minds, hearts, and the unwavering support of those around me. Just an acknowledgement page will never suffice to show my gratitude. I am constantly indebted to every one of you, and this accomplishment is as much yours as it is mine.

CONTENTS

1 INTRODUCTION 1
1.1 Cellular senescence . . . 1
1.1.1 Morphological and functional characteristics of senescent cells . . . 3
1.1.2 Cellular senescence and cancer . . . 4
1.2 Untangling the DNA double helix . . . 7
1.2.1 Components of DNA . . . 9
1.2.2 The DNA assembly with proteins: from nucleosomes to chromatin . . . 10
1.3 DNA damage and cancer radiotherapy . . . 14
1.4 DNA damage repair pathways . . . 17
1.4.1 Base-Excision Repair pathway . . . 19
1.4.2 Single-Strand Break Repair pathway . . . 20
1.5 Single-strand breaks and cellular senescence . . . 22
1.6 Layout of the thesis . . . 25
2 INTERACTION OF PARP-1 WITH SINGLE-STRAND BREAK IN A MODEL DNA-HAIRPIN 27
2.1 Introduction . . . 27
2.2 Stability of the isolated DNA hairpin including an SSB 30
2.3 Binding of PARP-1 monomer to DNA hairpin . . . 32
2.3.1 Retrieving the experimental structure . . . 32
2.3.2 Starting from the straight hairpin . . . 34
2.3.3 Manually fixing the right DNA-PARP-1 contacts 35
2.4 Comparison between monomeric and dimeric PARP-1 37
2.4.1 Dimeric PARP-1: the "experimental" Zn2 orientation . . . 38
2.4.2 Dimeric PARP-1: the 180°-flipped Zn2 orientation 40
2.5 Summary and outlook . . . 43
3 INTERACTION OF PARP1 WITH SINGLE-STRAND BREAKS IN THE NUCLEOSOME 47
3.1 Introduction . . . 47
3.2 Stability of SSB at different positions in the nucleosome 49
3.3 Analyzing the Zn-finger domain contacts with SSB . . 51
3.4 Nucleosomal reorganization by PARP-1 . . . 56
3.5 Mechanical deformation at SSB sites due to histone tails 59
3.6 Summary and outlook . . . 60
4 DAMAGE IN THE CHROMATIN: LOCALIZATION TO THE LINKER DNA 63
4.1 Introduction . . . 63
4.2 Comparison of DNA linker deformation with vs. without a SSB . . . 68
4.2.1 Simulation setup and force application . . . 68

4.2.2 Structural changes induced by SSB on the whole nucleosome . . . 70
4.2.3 Structural and mechanical changes in the linker-DNA . . . 72
4.3 Comparison of DNA linker deformation with vs. without a DSB . . . 76
4.3.1 Mechanical deformation due to DSB and its significance . . . 78
4.4 Summary and outlook . . . 80
5 MATHEMATICAL MODELING AND MONTE CARLO SIMULATION OF THE DNA DAMAGE REPAIR PATHWAYS 83
5.1 Introduction . . . 83
5.2 Mathematical modeling of the SSBR pathway triggered by out-of-field radiation. . . . 85
5.2.1 COPASI simulation of the SSBR pathway . . . 86
5.3 Monte-Carlo simulation of the SSB repair pathway . . . 90
5.3.1 Introduction to **TO**ol for **PA**rticle **S**imulation (TOPAS) 90
5.3.2 DAMARIS for SSB repair . . . 92
5.3.3 SSB repair pathway model: Framework, Assumptions and Limitations . . . 94
5.3.4 Results . . . 96
5.4 Summary and outlook . . . 106
6 CONCLUSIONS AND FUTURE PERSPECTIVES 109

I APPENDIX 115
A SUPPLEMENTARY FIGURES 117
B PROTOCOLS 133
B.1 Molecular simulation protocols . . . 133
B.1.1 Molecular construction of the DNA hairpin . . . 133
B.1.2 Assembly of the Zn finger domains of PARP-1 . . . 133
B.1.3 Molecular structure of the mononucleosome . . . 134
B.1.4 Molecular structure of the dinucleosome . . . 135
B.1.5 Docking protocols . . . 135
B.2 AMBER14 forcefield parameters . . . 136
B.2.1 Molecular dynamics simulations and analysis . . . 137
B.3 GROMACS Pull code . . . 139
B.4 COPASI: COmplex PAthway SImulator . . . 141
B.4.1 Kinetic Models in COPASI . . . 142
B.4.2 Differential equations for the entire out-of-field dose-triggered SSBR cascade pathway . . . 143
B.5 SDD format for TOPAS . . . 145
B.6 Finalized time constant for SSBR pathway . . . 146

II BIBLIOGRAPHY 149
BIBLIOGRAPHY 151

LIST OF FIGURES

Figure 1 Morphological Changes in Senescent Keratinocytes 4
Figure 2 Senescence Induction Flowchart 6
Figure 3 X-ray diffraction pattern of DNA 8
Figure 4 Molecular components of the DNA 10
Figure 5 Nucleosome core particle 11
Figure 6 Hierarchical Organization of DNA into Chromatin and Chromosomes 13
Figure 7 DNA damages due to Ionizing radiation . . . 15
Figure 8 Single-Strand Break Formation via Direct Ionization 16
Figure 9 Repair kinetics of different types of DNA lesions 19
Figure 10 Base-excision repair pathway 21
Figure 11 Reduced SSB repair pathway by the keratinocytes at the margin 24
Figure 12 Molecular structures of the PARP-1 protein and hairpin DNA 28
Figure 13 Analysis on straight hairpin DNA 31
Figure 14 Retrieving the experimental configuration . . . 33
Figure 15 Analysis on manually-adjusted hairpin DNA-PARP-1 structure 36
Figure 16 Interaction of Zn2 with the pre-bent DNA hairpin 38
Figure 17 Analysis on dimeric PARP-1 with experimental "Zn2"orientation 39
Figure 18 Dynamics of Zn1 Interaction with the hairpin DNA in the 180° flipped Zn2 orientation . . . 42
Figure 19 Details on 3 SSB in the nucleosome 49
Figure 20 Analysis on the damage sites without PARP-1 50
Figure 21 Interaction of Zn-finger domains at the three SSB sites 53
Figure 22 Gap distance analysis for the 3 sites docked with PARP-1 54
Figure 23 Mechanical deformation calculated at damage sites 57
Figure 24 Dihedral angle evolution at the damage site with and without PARP-1 59
Figure 25 H-bond analysis on site 1 and site 3 60
Figure 26 The 6 structural variations of the dinucleosome, categorized by defect type and linker length . 67
Figure 27 Before and after pushing the nucleosome together with a SSB. 69

Figure 28 Distance and angle evolution for SSB variation 71
Figure 29 Linker DNA deformation with & without SSB 73
Figure 30 Structural helical parameters of DNA 74
Figure 31 Analysis of Short & Long Linker DNA with & without SSB 75
Figure 32 RMSD and COM analysis on linker-DNA with DSB 78
Figure 33 Structural helical parameters analysis on DSB variation 79
Figure 34 SSBR pathway utilized for COPASI modeling . 87
Figure 35 COPASI time-course simulation results 89
Figure 36 DaMaRiS algorithm 93
Figure 37 SSBR pathway utilized for MC simulations . . 95
Figure 38 Chromatin structure utilized in DaMaRiS model. The colored wired center represents the chromatin material. The blue mesh is the nuclear membrane and the black mesh represent cytoplasm. While the square boxes represent the box size of the entire system under consideration. 98
Figure 39 SSB quantification 101
Figure 40 Trial and error method for fixing time constant for PARP-1 kinetics 103
Figure 41 Protein recruitment kinetics 104
Figure 42 Overall SSB repair kinetics 105
Figure A1 RMSD and RMSF value of isolated DNA hairpin configuration 117
Figure A2 Contact maps of the PARP1-DNA hairpin . . . 118
Figure A3 Zn2 and Zn1 interaction with the straight hairpin 119
Figure A4 Analysis on manually adjusted Zn1-Zn2 starting from the straight hairpin 120
Figure A5 Analysis of PARP1-DNA with and without linker between the Zn-finger domains 121
Figure A6 Interactions between Zn1-Zn2 and the hairpin in the orientation **(2)** 122
Figure A7 Docking with Pydock 123
Figure A8 RMSD of end nucleotides in the break site 3 . 123
Figure A9 Super helical position in nucleosome 124
Figure A10 Position and numbering of the SSB introduced in the DNA wrapped around the nucleosome for H-bond analysis. 124
Figure A11 Helical parameter for short-linker variation . . 125
Figure A12 Helical parameter for long-linker variation . . 126
Figure A13 Different helical parameters for short linker with a DSB. 127
Figure A14 Different helical parameters for long linker with a DSB. 128

Figure A15 Mechanical instabilities in an axially compressed rod 129
Figure A16 Evolution of distance between the two nucleosomes 130
Figure A17 Gap distance of the second strand break in DSB 130

Table 1 DNA damage (N/(cell·Gy)) induced by ionizing radiation vs. endogenous damage due to normal cell activity.[84] 17
Table 2 H-bond formed at the SSB site 55
Table 3 The fitted parameters of eqn (2), for predicting the yield of each damage category as a function of proton dose and LET. 99
Table 4 Sensitive fraction vs Damages 100
Table 5 The different literature used to validate DaMaRiS Model and their R^2 value considered. 102
Table 6 LET values and corresponding damage metrics 131
Table 7 Atomic properties of different elements from Ref.[245] 136
Table 8 DGP parameters in dna.hdb file 137
Table 9 DGP parameters in the dna.rtp file 138
Table 10 Field by Field Summary of the Header Fields and their Type 145
Table 11 Value by Value Definition of the Data Fields to Score DNA Damages 146

1 INTRODUCTION

1.1 CELLULAR SENESCENCE

The early 1900s saw a surge of interest in scientific research and speculation around the concept of immortality of biological tissues, a notion not necessarily mainstream or widely accepted within the scientific community. From 1910 to 1940s, Alexis Carrel, a Nobel-prize-winning French surgeon, observed that fibroblasts isolated from chicken hearts could grow in culture for over 30 years [1]. In his paper titled "On the Permanent Life of Tissues outside of the Organism," Carrel proposed that with ideal culturing techniques, the lifespan of tissues grown *in vitro* could be extended indefinitely, thus suggesting that tissues inherently possess the ability to sustain perpetual life *in vitro*. This finding stirred the scientific community, especially since chickens as a whole have a maximum lifespan of only ~10 years. This work led a few biologists to believe that cells were essentially immortal *in vitro* and would only succumb to aging in an *in vivo* setting [1, 2].

Nevertheless, all of Carrel's notions were questioned when Leonard Hayflick performed a series of experiments demonstrating a finite lifespan for human cells grown *in vitro*. In 1961, Hayflick needed to maintain a huge number of cells alive and healthy for some experiments, so he resorted to a technique called "serial passaging". With this method he was able to maintain the cell culture for months. However, at some point the cells began to degrade and were unable to divide further while remaining alive. Assuming he made some experimental errors, he collaborated with Paul Moorhead to investigate this issue. The duo eventually determined that healthy cells can undergo only a finite set of divisions, about ~40-60 passages. Hence, their study suggested that healthy cells are incapable of dividing infinitely (unlike Carrel's claims), but after a few number of divisions they cease to divide, and enter a state of irreversible growth arrest termed as **replicative senescence** [3]. This phenomenon was later observed in a plethora of species, and eventually, the limit of cell division was termed as '**Hayflick limit**'. As a possible explanation, Hayflick assumed that the cell has a core memory of previous cell divisions, which refers to a cellular mechanism that tracks the number of divisions and directs it to enter replicative senescence once exhausted [3]. Over the century, scientists concluded that this cellular machinery is working on the genetic material of the cell, rather than on the organelles and protein content [4].

Between 1970 and '90s, Elizabeth Blackburn and Carol Greider discovered that the telomeres play a major role in this process. Telomeres are a set of repetitive DNA sequence at the end of the eukaryotic chromosomes, which help to prevent incomplete replication and instability [4, 5]. They observed that after each division there is a gradual loss of a few nucleotides, which results in **telomere shortening**. As the telomeres reach a certain length, they become unable to bind to certain telomere-capping proteins and expose the DNA ends. This triggers the DNA repair pathways, and eventually leads to cell senescence and cell death [4, 6].

Senescence, once thought to be just a flaw in tissue culture techniques, has since been recognized as a critical process in both physiological and pathological contexts. It is triggered by various signals or stresses, which include telomere shortening, mitogenic signals, radiation, oxidative and genotoxic stress etc [7, 8]. These stress induce cell senescence by ceasing the cells to proliferate further. The entire process of senescence is one of the main reasons for cellular aging, as they secrete inflammatory factors which leads to tissue deterioration which leads to aging and age related diseases [9].

Cellular senescence plays both beneficial and harmful roles. It has key physiological functions in normal development, tissue remodeling, wound healing, and limits tumor progression by preventing damaged cells from proliferating. However, senescence also contributes to aging and age-related diseases. The accumulation of senescent cells can impair tissue repair and regeneration, deplete stem and progenitor cell reserves, and trigger the release of the senescence-associated secretory phenotype (SASP)[8–10]. Also, senescent cells can evolve into neoplastic cells at much later times [11–13] and be at the origin of secondary cancers. The link between senescence and **unrepaired DNA damage** is one key issue that motivates the main subject of this thesis.

Rather generally, one can distinguish cell senescence into two main categories [14, 15]:

1. **Replicative** senescence, and

2. **Stress-induced** premature senescence (SIPS)

Replicative senescence is a result of spontaneous telomere shortening after several dozens of cellular division [4]. On the other hand SIPS - as the name suggests - is a quick response of the cell to exogenous stress induced by an external factors [16, 17]; this is independent of telomere shortening and is relevant to both normal cells and cancer cells; moreover, it is a major contributor to **oncogene-induced** cell senescence (OIS). In summary, cellular senescence acts as a double-edged sword, playing a crucial role in regulating the cell life cycle.

1.1.1 *Morphological and functional characteristics of senescent cells*

Senescence is a cellular state marked by an irreversible cell-cycle arrest, increased activity of senescence-associated β-galactosidase (SA-β-gal), and resistance to apoptotic stimuli [18]. The irreversible cell-cycle arrest persists even when there is an active mitogenic stimulus; this characteristic distinguishes the senescent cell from the quiescent, where the cells can resume proliferation in response to appropriate signals. Typically, senescence associated cell-cycle arrest is mediated by the p21/p16 proteins [19], which inhibit the activity of cyclin-dependent kinases (CDK), but without affecting the growth-promoting pathways such as mTOR (mechanistic target of Rapamycin) and MAPK (mitogen-activated protein kinase) pathways. Activity of mTOR during cell-cycle arrest leads to cellular hypertrophy and tissue-specific hyperfunction [18, 20–22]. Hence, they become morphologically flattened, and display enlarged nucleus or multiple nuclei, cytoplasmic granularities, multiple vacuoles in the cytoplasm, and increase in cytoplasmic microfilaments (see Figure 1) [23, 24]. These events are the result of protein accumulation inside the cell, due to lowered proteasome activity [25]. Another hallmark of the senescent cell is over-expression of SA-β-gal, a lysosomal enzyme found in large amounts and often used as a senescent cell marker. This can be used *in vitro* or *in vivo*, and lightens the cell under a microscope with a bright-blue color because of the increased β-gal activity, which helps to visualize and quantify the senescent cells inside a population.

Senescent cells also develop a senescence-associated secretory phenotype (SASP). The proteins secreted in this process form a diverse group, including inflammatory cytokines, chemokines, matrix metalloproteinases, and growth factors, which have biological consequences that vary depending on the context. SASP can perpetuate senescence through autocrine and paracrine signalling, and it also triggers immune responses that may result in the elimination of senescent cells. The inflammatory component of SASP is primarily driven by the transcription factors NF-κB and C/EBPβ, which are activated in response to DNA damage, along with other interconnected signalling pathways, such as GATA4, mTOR, and the p38MAP kinase. [20]. Senescent cells accumulate enlarged and dysfunctional mitochondria. This could be due to the decrease in mitochondrial membrane potential, increased proton leak and reduced rates of fusion and fission. Altogether, these disturb the ability to generate ATP and increase the production of reactive oxygen species within the cell and outside the cellular environment. The increase in oxidative stress results in misfolded proteins and protein aggregation. To accommodate these mishaps within the cell, the cell tries to enlarge its organelles and trigger various cascade pathways to rectify it [20].

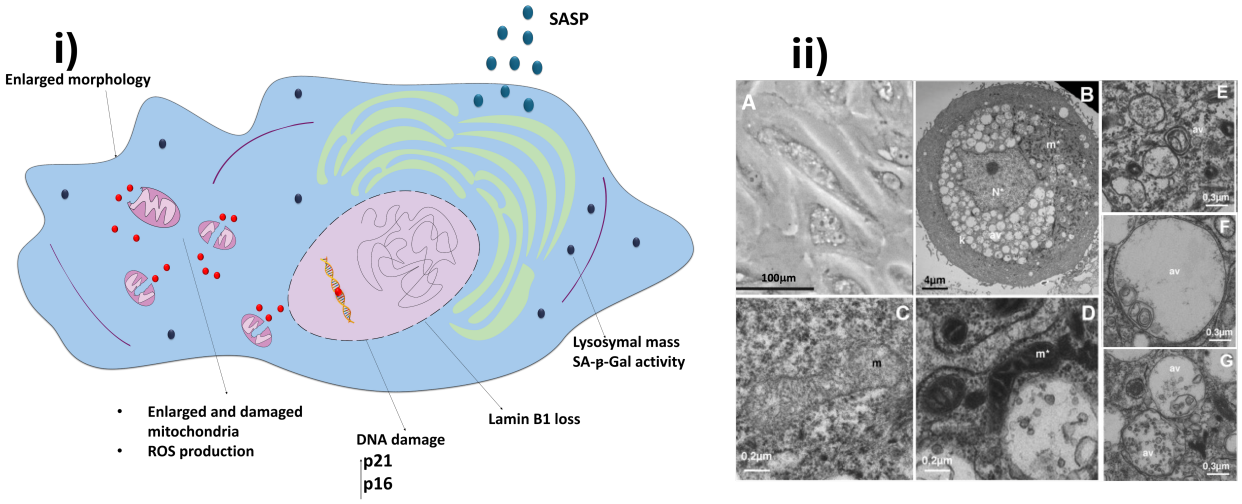


Figure 1: **(i)** Schematic representation of the main morphological changes that happen within a senescent cell. **(ii) (A)** Senescent keratinocytes under phase-contrast microscopy; note the granular and vesicular-like material around the nucleus and the polynucleation (from [26]). **(B)** Keratinocytes at the senescence plateau; N* is the deformed nucleus of a senescent cell; k is the cytokeratin network encircling the nucleus and the autophagic vacuoles; and m* indicates mitochondria aggregated close to the nucleus. **(C,D)** Details of mitochondria morphology in the growth phase and at the senescence plateau, respectively. **(E–G)** Details of autophagic vesicules (av) in keratinocytes at the senescence plateau (from [27]).

1.1.2 *Cellular senescence and cancer*

Cellular senescence has long been known to have a link with cancer, since most cancer cells acquire the capability to proliferate uncontrollably. Notably, senescence with its distinctive characteristic of cell-cycle arrest may act as a limiting criterion for cancer cell progression [28]. Oncogene-induced cell senescence (OIS), is a phenomenon due to aberrant activation of oncogenes, especially members of BRAF and RAS families [29]. Such mitogenic changes overwhelm cellular division mechanisms by triggering cascade pathways and unscheduled DNA replication. In response, the cells tries to induce proliferation-arrest, taking advantage of the benefiting role of senescence. Thereby, this almost switches-off the malignant nature of cancerous cells.

It is worth noting that cells can go into a senescent state not just due to cancer, but also due to cancer treatments. Treatments such as chemotherapy, radiotherapy and targeted drug deliveries have been known to induce cell senescence, in a form termed **therapy-induced cell senescence** (TIS). Exposure to chemotherapy drugs or radiation induces senescent markers in cells. Several types of cancer treatments can induce senescence including exposure to alkylating agents [30] like cisplatin, cyclophosphamide, and temozolomide; topoisomerase inhibitors [30] like doxorubicin, etoposide, and camptothecin; γ-radiation [29]; microtubule inhibitors like paclitaxel [31]; and, to a lesser extent, vinca alkaloids like vincristine [32].

Accumulation of unrepaired DNA damage is a main inducer of cell senescence.

Irrespective of the detailed treatment mechanisms, almost all cancer therapies lead to **DNA damage**. Accumulation of damage is recognized as the primary mechanism inducer of senescence in both normal and cancer cells. Irreparable DNA damage, including even just a single unrepaired DSB, has been shown to be sufficient to induce cellular senescence [33]. In this respect, it is important to note that the level of DNA damage required to induce senescence may differ between malignant and non-malignant cells, similar to the varying thresholds for apoptosis. For instance, patients with breast cancer were found to have increased number of T-cells expressing $p16^{INK4a}$ when administered with different chemotherapeutic agents; hence indicating immuno-senescence as a bystander effect[29, 34]. Similarly, radiation therapy also led to the upregulation of $p16^{INK4a}$ in non-malignant cells [13, 35]. [1]

Induction of cellular senescence with respect to TIS/OIS occurs through distinct pathways, described in Figure 2. TIS, which is the main concern of this work, results in the triggering DNA damage repair (DDR) pathways, leading to activation of the key proteins p53 and $p21^{WAF1/CIP1}$; alternatively, OIS results in the activation of either or both of the p53/$p21^{WAF1/CIP1}$ and $p16^{INK4A}$/Rb tumor-suppressor pathways, and the Ras-Raf-MEK-ERK, PI3K/AKT/mTOR and p38/MAPK signalling pathways [36, 37]. These proteins are essential components, as they inhibit CDK2-cyclin-E and CDK4/6-cyclin-D complexes and prevent the phosphorylation of RB (retinoblastoma) protein. Such events block the entry in S-phase and induce senescence [36, 37]. While the tumor-suppressor proteins p53/$p21^{WAF1/CIP1}$ and $p16^{INK4A}$ are primarily involved in TIS and OIS, senescence can also occur independently of these pathways. Additionally, $p21^{WAF1/CIP1}$ can be activated by pathways other than p53 [38, 39]. Also, prolonged over-expression of any of these four factors – p53, RB, $p16^{INK4A}$, or $p21^{WAF1/CIP1}$ – is sufficient to induce senescence [36]. Disregulated expression, mutations, secondary alterations or complete loss of tumor-suppressor proteins may result in inadequate senescence induction or an escape from senescence [13, 36]. The cell may enter apoptosis or proliferation stage depending on whether the DNA is correctly and thoroughly repaired or not. Regardless of the cancer type, the level of senescence (whether it's OIS or TIS) appears to significantly influence the outcome for cancer patients.

The original motivation for the **SENESIMEX** project at the origin of this thesis, generously funded by the University of Lille through the

[1] Proteins are sometimes indicated with an exponent-like label to specify a particular isoform, family, or regulatory association. The notation helps distinguish between proteins with similar names or multiple functions. For example, $p16^{INK4a}$ refers specifically to the INK4a isoform of the p16 tumor suppressor protein, which is encoded by the CDKN2A gene. Similarly, $p16^{INK4a}$/Rb indicates that $p16^{INK4a}$ is acting on the Rb (retinoblastoma protein) pathway, through the inhibition of cyclin-dependent kinases (CDK4/6) to regulate cell cycle progression.

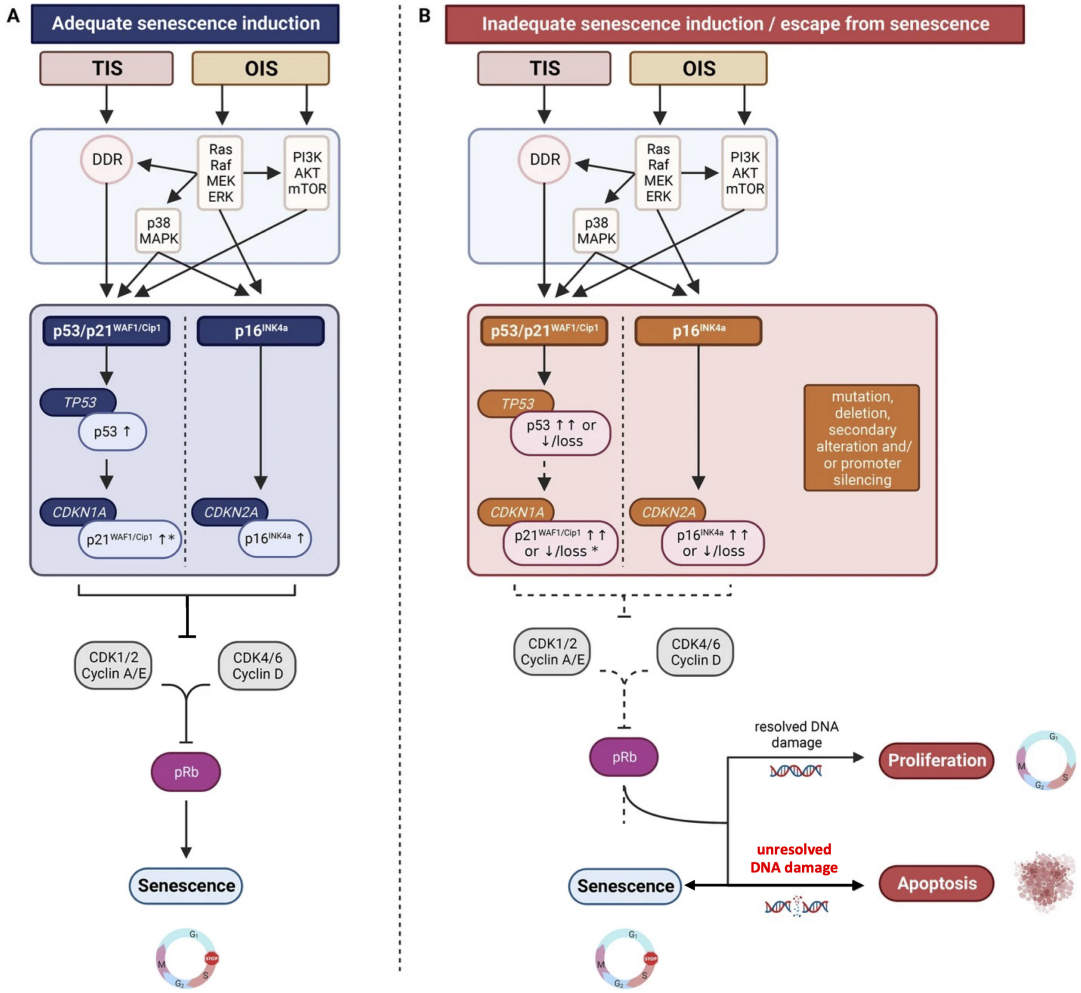


Figure 2: **(a)** Flowchart of "adequate" senescence induction, vs. **(b)** Inadequate cell senescence pathway. Note the bifurcation between senescence and apoptosis for unresolved DNA damage in the latter. This will constitute the subject of mathematical modeling in Chapter 5. (Adapted from [36], w/ permission.)

program PEARL ("Programme for Early-Stage Researchers in Lille"), stems from the interest of our research groups in the 'triangular' connection between cellular senescence, cancer development and therapies, and (unrepaired) DNA damage. Given the very wide impact and implications of such extensive research program, which is correlated to ongoing sister projects and activities in the CANTHER and IEMN Laboratories (belonging to CNRS, INSERM and University of Lille), my thesis work focussed on the more molecular aspects of the early steps leading to senescence. As it will be seen in Chapters 2 to 4, one key focus is the molecular correlations of cell senescence with DNA damage and repair, especially as far as the ubiquitous single-strand breaks to the DNA phosphate backbone are concerned (the subject has been developed and detailed in our paper *Cell senescence and the DNA single-strand break damage repair pathway* [40]). For this

purpose, the rest of this introductory Chapter provides some basic notions about DNA molecular structure and organization inside the cell nucleus; molecular mechanisms of DNA damage and repair; and their direct and indirect relationship with cancer and cellular senescence.

1.2 UNTANGLING THE DNA DOUBLE HELIX

The first steps for understanding the role of genetic material started around the late 1860s when Gregor Mendel explained how traits (physical characteristics of a living organism, like height, the color of the eyes, etc.) are transferred from one generation to another [41]. Three years later, in 1869, a Swiss physiological chemist, Friedrich Miescher, first discovered a substance called 'nuclein' inside the nucleus of human white blood cells, which had a high phosphorus content and strong resistance to proteolysis [42]. Despite these early discoveries, it took more than 50 years for the Mendel's theories and Meischer's study to be acknowledged by the scientific community.

The early 20th century brought further advances, when Phoebus Levene proposed that nuclein should be constituted by a long filament of *nucleotides* (a nitrogenous base linked to a sugar and a phosphate) connected via their phosphate groups, and called the chemical structure 'deoxyribonucleic acid' (which later W. Astbury shortened to DNA). The work of scientists like Oswald Avery in the 1940s provided compelling evidence that DNA, and not proteins, was the substance responsible for heredity. The famous experiments by Avery and his colleagues showed that DNA could transfer genetic information from one organism to another, a process they termed "transformation" [43]. Seven years after Avery's discovery, A. D. Hershey and Martha Chase proved the transforming principle. With the help of radioactive-marked phosphorous in DNA and radioactive-marked sulfur in proteins, they demonstrated that in bacteriophages, indeed, DNA is being transferred from one generation to another, and not proteins [44]. This pivotal discovery opened the door to further research, enabling a few years later Erwin Chargaff to decipher the composition of DNA [45].

At this time, many scientists were racing to determine the precise structure of DNA. One of the notable attempts was made by Linus Pauling, a two-time Nobel laureate. In 1953, Pauling proposed a model in which DNA was composed of three intertwined strands with phosphate groups at the core and the bases pointing outward [46]. However, this model turned out to be incorrect due to a fundamental flaw: it failed to account for the negative charge of the phosphate groups, which would repel each other in such an arrangement. This error left the door open for Watson and Crick who, with the help of Rosalind Franklin's critical X-ray diffraction images (among which

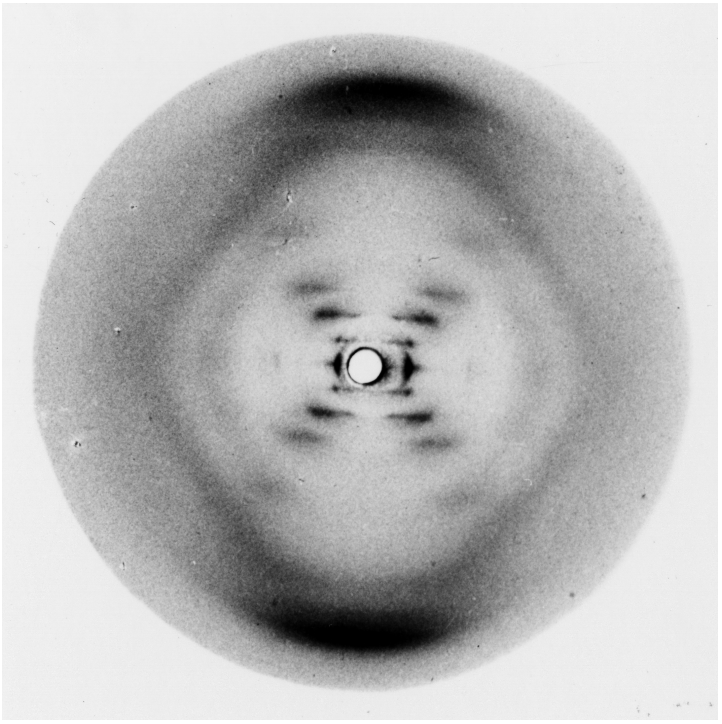


Figure 3: The "Photo 51", showing X-ray diffraction pattern of DNA, taken by Raymond Gosling, a post-graduate student working under the supervision of Maurice Wilkins and Rosalind Franklin at King's College London, while working in Sir John Randall's group. The image was tagged "Photo 51" because it was the 51st diffraction photograph that Franklin had taken. It was critical evidence in identifying the structure of DNA, since the X-shaped spots visible in the reciprocal space of the diffracted image unambiguously point to a helical arrangement of the scattering centers.

the famous "Photo-51", Figure 3, developed the correct structure of DNA.

Already the early chemical model by Levene clearly comprised nucleotides formed by one of four variant nitrogenous bases (adenine, thymine, guanine, cytosine), a sugar molecule which he correctly identified in the deoxyribose, and a negatively-charged phosphate group [47–49]. Structural studies of DNA started in the early 1950s when DNA was first examined by X-ray diffraction analysis. The early X-ray diffraction results indicated that DNA comprised of two strands of the polymer wound into a helix. The observation that DNA was double-stranded was crucial and provided one of the significant clues that led to the Watson-Crick deciphering of the structure of DNA. Their structure proposed that DNA is a double helix with two antiparallel strands running in opposite directions [50]. The sugar-phosphate backbone forms the outer structure, while the nitrogenous bases (adenine, thymine, guanine, and cytosine) face inward, pairing through hydrogen bonds (adenine pairs with thymine, and guanine pairs with cytosine) [50]. This complementary base-pairing explained Chargaff's rules and also provided a natural mechanism for DNA replication [45]. DNA's potential for replication and information en-

coding could became apparent only when this model was proposed and verified.

1.2.1 *Components of DNA*

A DNA molecule is composed of two long chains of nucleotides, which are the basic building blocks of DNA. Each of these chains is referred to as a DNA strand. The two strands are held together by hydrogen bonds between the nucleotide bases. Nucleotides comprise a five-carbon sugar, one or more phosphate groups, and a nitrogenous base. In DNA, the sugar is deoxyribose, attached to a single phosphate group, and the nitrogenous base can be adenine (A), cytosine (C), guanine (G), or thymine (T) (Figure 4). These nucleotides are connected by covalent bonds through the sugar and phosphate groups, creating a "backbone" that alternates between sugar and phosphate along the DNA strand. Additionally, the nitrogenous bases are attached to the sugar moiety via N-glycosidic bonds, which links the base to the nucleotide structure. Since the base is the only varying component among the four types of nucleotides, each DNA strand can be compared to a necklace (the backbone) with four different types of beads (the bases A, C, G, and T). These symbols (A, C, G, and T) are also used to represent the four nucleotides, which include the base, sugar, and phosphate groups [51].

The nucleotide sequence in a DNA strand has a specific chemical polarity. Each sugar molecule can be seen as a block with a phosphate protruding at the $5'$ carbon atom, on one side, and a hydroxyl at the $3'$ carbon, on the other (see again Figure 4). These blocks connect sequentially by forming a phosphodiester $5' - 3'$ bond (which eliminates a water molecule) to form a single DNA strand; the two ends of the strand are distinct, with one end terminated by a $3'$ hydroxyl group, and the other by a $5'$ phosphate group. This polarity is indicated by referring to the ends as the $3'$ terminus and the $5'$ terminus [51].

The three-dimensional structure of DNA, known as the "double helix", results from the chemical and structural characteristics of its two nucleotide strands (hence, the term dsDNA to indicate double-stranded DNA). These strands are held together by hydrogen bonds between bases on opposing sides, positioning all the bases inside the double helix while the sugar-phosphate backbones remain on the outside. Each base pair comprises a purine (a two-ring structure) or a pyrimidine (a single-ring structure). According to the chemical rule assessed by E. Chargaff, in normal DNA adenine pairs with thymine, and guanine pairs with cytosine. Such a complementary base-pairing allows the bases to align in the most energetically favorable way within the double helix, maintaining a uniform distance between the sugar-phosphate backbones. The complementarity imposes a mirror-

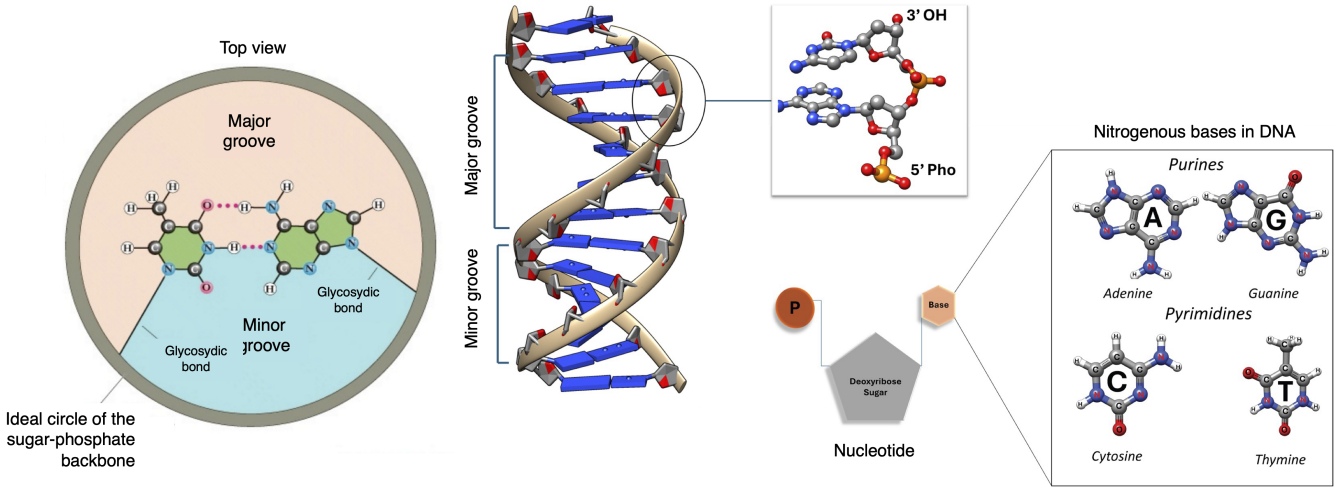


Figure 4: The main molecular components of the DNA structure. On the left, the double-helix in which the base-pairs are arranged in a spiral staircase, and the top-view scheme of the glycosidic bond arrangement. The pitch of the helix is not regular because the insertion points of the two phosphates of each base pair are not symmetric with respect to the central axis, thereby giving rise to a major and a minor groove. On the right, the two purine and the two pyrimidine nitrogen bases are depicted in ball-and-stick style (red spheres = oxygen, white=hydrogen, grey=carbon, blue=nitrogen).

inversion symmetry of the two strands in the double helix: the two sequences are identical, however one runs in the 5′-3′ direction, while the other runs in the opposite 3′-5′ direction. This has drastic consequences on the transcription and duplication of DNA, since all the enzymes involved in the process can only add nucleotides to polymerize a nucleic acid in the $5' \rightarrow 3'$ direction.

The packing of base-pairs is optimized as the two sugar-phosphate backbones twist around each other, creating a double-helix structure with one complete turn of 360 degrees about the DNA main axis occurring about every ten base pairs. Interestingly, the backbone geometry leads to the formation of a "major" and a "minor" groove, i.e. a helix with uneven pitch. The angle between the glycosidic bonds (sugar moiety and DNA) is 120 degrees for the narrow angle, and 240 degrees for the wide angle; as a result, the two phosphates of each base-pair, inserted in an imaginary circle perpendicular to the DNA axis, are not symmetric (see Figure 4, left-side). It turns out that when the base-pairs stack upon each other, the narrow angle between the sugars on one edge generates a minor groove, while the large angle generates a major groove, along the helical structure.

1.2.2 *The DNA assembly with proteins: from nucleosomes to chromatin*

The total length of the about 6 billion base-pairs of DNA inside the human cell nucleus would be about 2 m, if they were stretched end to end. This humongous stretch of DNA must be compactly packed in-

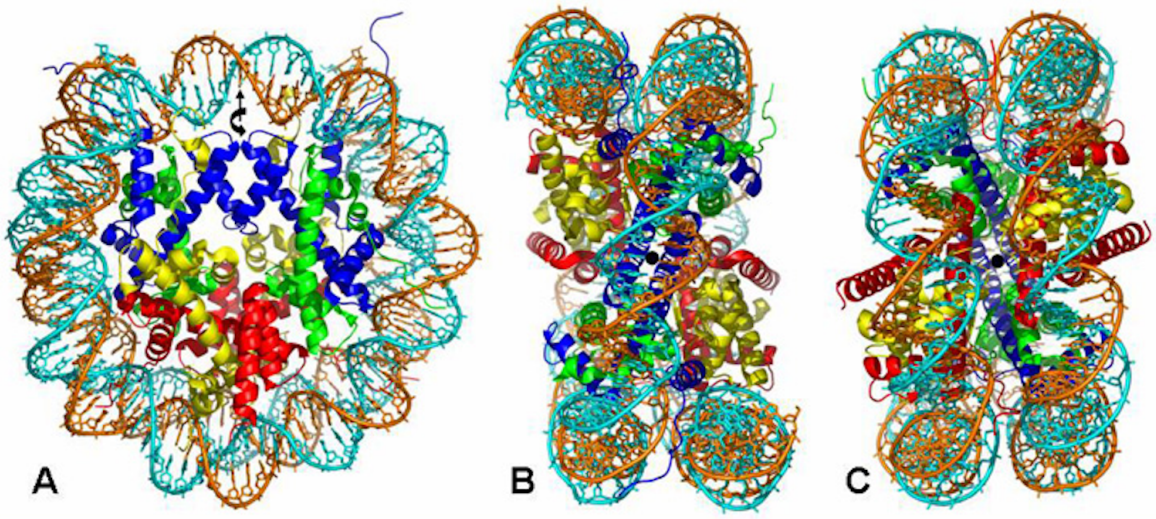


Figure 5: Nucleosome core particle: **(a)** top view along the superhelical (SHL) pseudo-symmetry axis, running perpendicular to the page through the geometrical centre of the DNA loops; **(b)** back-side view showing the dyad symmetry point; **(c)** front-side view showing the double loop. The 147-bp DNA are depicted in ribbon and stick representation for the two antiparallel strands (gold and cyan); the eight histone protein dimers are represented in ribbons (H2A:Green, H2B:Magenta, H3:Red, H4:Yellow). The molecular structure is taken from the RCSB PDB data base, entry 1KX5 [56].

side the cell nucleus, whose diameter is less than ~10 μm. To achieve such an extreme degree of compaction, the DNA is wrapped around clusters of specific proteins, the histones, forming a structure known as the "nucleosome."

The nucleosome, which serves as the chromatin basic structural and functional unit, comprises eight histone proteins and 147 base pairs of DNA [52, 53]. The discovery of nucleosomes was initially suggested by electron microscopy observations, where chromatin appeared as a "beads-on-a-string" structure [54, 55]. Further evidence supporting the existence of nucleosomes came from experiments involving the chemical cross-linking of histones in chromatin, which revealed that the histones dimers H2A, H2B, H3, and H4 form a distinct octameric structure, consistent with a repeating histone unit within the chromatin fiber as shown in Figure 5 [56].

A typical nucleosome core mainly consists of DNA tightly wrapped around eight proteins. Each nucleosome encompasses two copies of the histones H2A, H2B, H3, and H4, which assemble into a histone octamer. This octamer binds and wraps 1.7 turns of DNA, corresponding to about 147 base pairs, in the "nucleosome core particle", or NCP. Since the helix repeat of DNA is about 10.4 bp, the DNA makes contact with the histone octamer surface at 15 spots, including the two entry/exit sites. The axis going through the center of the nucleosome, perpendicular to the plane containing the DNA turns, is the "superhelical" axis (a pseudo-symmetry axis), the DNA-histone contact cor-

responding to super-helical locations usually numbered from -7 to +7, with the 0 located at the center of planar mirror symmetry, called "dyad" (see figure A9 in Appendix A for a description of super-helical location/SHL positioning).

When the 'linker" histone H1 is added, it further wraps an additional 20 base pairs of DNA, resulting in two complete turns around the octamer, and forming what is known as a chromatosome [54, 57]. Because of the small amount of DNA associated with each nucleosome, a chromosome that contains on average over 100 million DNA base pairs, consists of hundreds of thousands of aligned nucleosomes. Pairs of nucleosomes in the sequence are connected by stretches of "linker" DNA (the nucleosome is properly defined as the NCP + linker DNA), which are typically about 20-80 base pairs in length (but can occasionally be longer, up to 2-300 bp) thus creating a "string of beads" appearance when viewed under an electron microscope [54, 57].

The length of DNA per nucleosome was determined by treating chromatin with DNAses, such as micrococcal nuclease (MNase), which preferentially cuts the linker DNA between nucleosomes before affecting the DNA wrapped around the histone octamers. By controlling the extent of MNase digestion, researchers could stop the reaction before all linker DNA was cleaved, resulting in a mixture of mononucleosomes, dinucleosomes, trinucleosomes, and higher-order structures connected by linker DNA [58]. When DNA from MNase-treated chromatin was analyzed by gel electrophoresis, it produced bands of varying lengths, each corresponding to a multiple of one mononucleosomal DNA [59]. This observation, combined with electron microscopy and histone cross-linking data, led to the acceptance of the "subunit theory" of chromatin [60, 61]. These subunits, later named nucleosomes [62], were eventually crystallized, providing a detailed model of their structure [56, 63].

The organization of DNA into nucleosomes results in a reduction of fiber length by approximately seven-fold. For instance, a DNA strand that is originally 1 meter long is condensed into a "string-of-beads" chromatin fiber that is only 14 centimeters long [51]. However, even this shortened length of chromatin is still far too long to fit within the nucleus, which typically measures only 3 to 5 micrometers in diameter. Consequently, chromatin undergoes further coiling into a more compact and thicker fiber, which has long been known as the "30-nanometer fiber," named for its approximate diameter of 30 nanometers [64]. (See Figure 6.)

The precise mechanism by which nucleosomes are folded into the 30-nanometer fiber, or other competing super-structures, has been and still remains a subject of extensive research and debate. One of the challenges in studying this process is that electron microscopy, although a powerful tool for visualizing chromatin packaging, makes

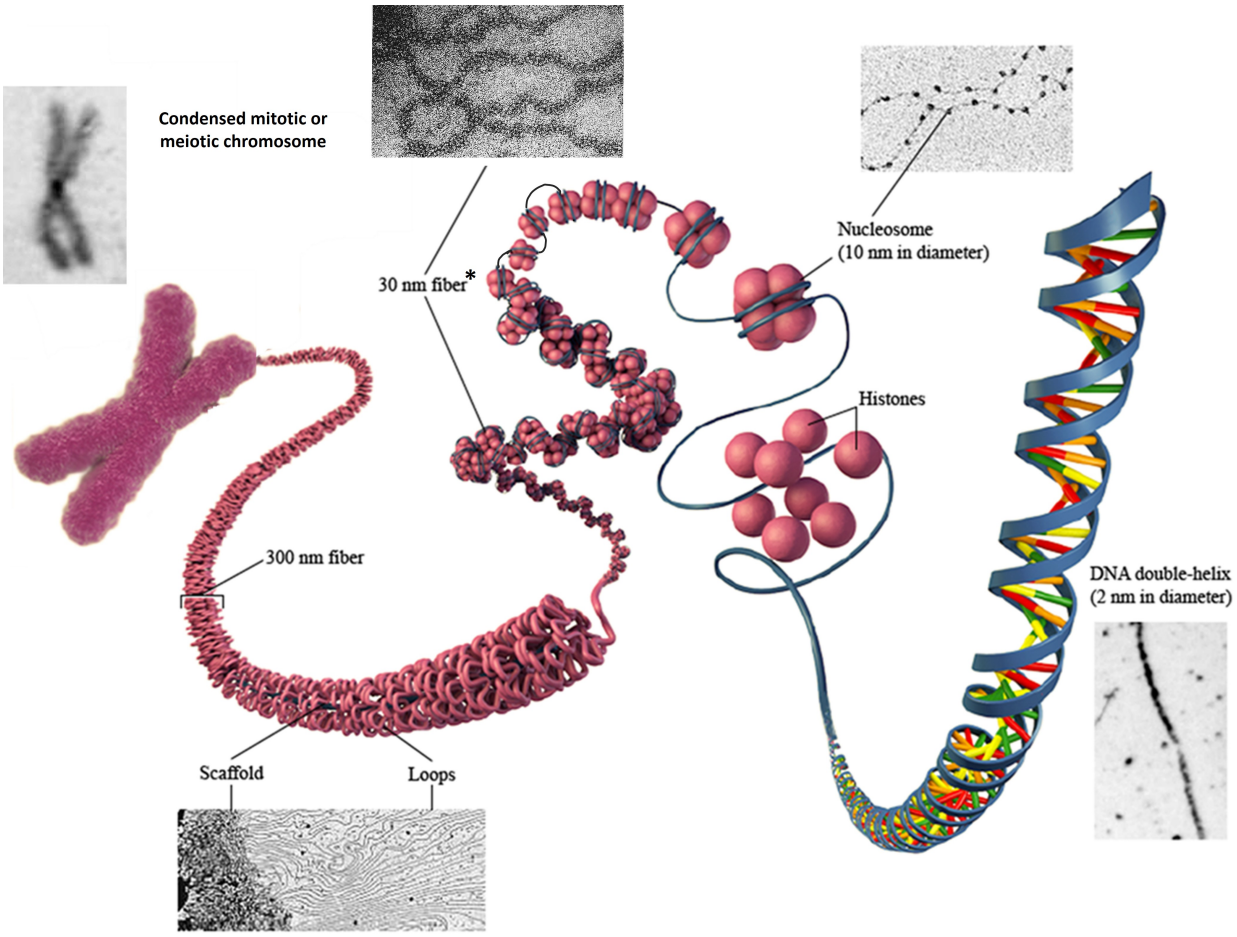


Figure 6: Schematic representation of how DNA is wrapped around the histone octamer, then nucleosomes condensing into chromatin fibers, and eventually self-structuring into a chromosome. Electron-microscope images at different resolutions accompany the subsequent steps in the hierarchic assembly of the idealized chain. *Still not established and still remains a subject of extensive research

it difficult to distinguish individual nucleosomes, once the dense chromatin fibers have formed. Additionally, the appearance of these superstructures can vary depending on whether observations are made on isolated chromatin, or within intact nuclei. As a result, it is increasingly believed that the 30-nanometer fiber may not be as regular or uniform, as it was often depicted in schematic diagrams [65, 66]. Notably, histone H1 plays a crucial role in stabilizing these higher-order chromatin structures, with locally regular stretches of 30-nanometer fiber forming most efficiently in its presence [67].

During cellular nuclear processes such as transcription and replication, the two strands of DNA must temporarily separate, allowing polymerases to access the DNA template. However, the presence of nucleosomes and the subsequent folding of chromatin into structured fibers present significant obstacles to the enzymes responsible for unwinding and copying DNA. To overcome these barriers, cells utilize

chemical mechanisms that either open up chromatin fibers, or temporarily remove histones, thereby enabling transcription and replication to proceed. Broadly, there are two primary mechanisms for increasing chromatin accessibility:

- Histones can be enzymatically modified by the addition of acetyl, methyl, or phosphate groups [68].
- Chromatin remodeling complexes can displace histones, thereby exposing the underlying DNA sequences to polymerases and other enzymes [69].

These processes are reversible (centromeres are indeed difficult to replicate, but finally they are [70]), allowing the chromatin to return to its compact state the once transcription or replication is complete.

1.3 DNA DAMAGE AND CANCER RADIOTHERAPY

Detailed knowledge of the DNA structure and its enzymatic interactions within the nucleus is essential to cancer treatments. Cancer remains a significant concern among researchers and medical professionals due to its high mortality rates, complex pathophysiology, and profound impact on public health and quality of life [71]. Numerous techniques are employed to treat cancer, including surgery, chemotherapy, radiotherapy, targeted drug delivery, and immunotherapy. Among these, the most widely used treatments are radiotherapy and chemotherapy, which can be combined or used independently, depending on the type and severity of the cancer [72]. When cancer has not metastasized and is in its early stages, radiotherapy is often the preferred treatment method. Even though radiotherapy has proven to be an effective mode of treatment, it is always associated with other health-related risks, including the development of secondary malignancies that are rare but aggressive. This alone provoked the scientific community to understand the mechanisms by which radiotherapy induces cancerous cell damage and death [73].

Radiation energy transfer to DNA induces ionization

During radiotherapy, ionizing radiation (IR) is targeted at cancerous cells, which can kill them by inducing DNA damage [74]. IR inflicts damage on molecules such as DNA, proteins, membranes, by imparting amounts of energy capable of causing ionization. The electron loss in ionization can be viewed as an oxidation process. Most of the energy transferred during radiation exposure is mediated by energetic electrons generated from the interaction of photons or other ionizing particles with molecular components of the cell. Importantly, these ionization events are spatially confined along the tracks of the primary ionizing particles or the secondary electrons. BHigh-energy charged particles and electrons near the end of their tracks create dense ionization clusters within small volumes [75, 76]. These clusters can generate complex DNA damage (Figure 7).

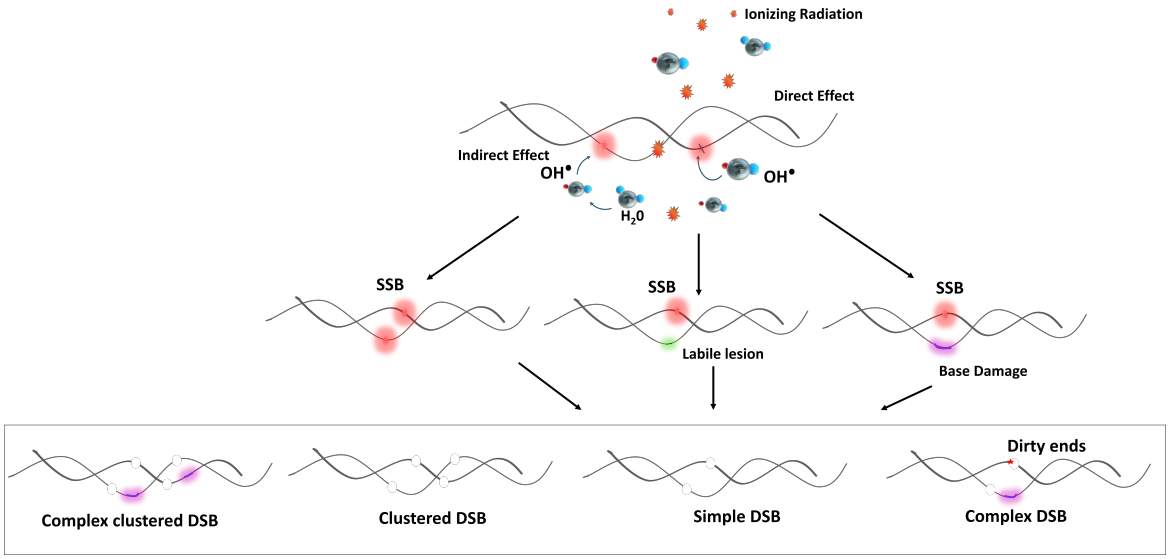


Figure 7: Ionizing radiation (IR) can directly attack the DNA, or it can generate reactive oxygen species (ROS) by interacting with the water molecule (direct vs. indirect effect). These can in turn cause various strand breaks (single-, SSB, or double-strand, DSB), base damage, or labile sugar lesions (these include different form of sugar damage, abasic sites and base damages that affect the stability of sugar). At the same time, IR can also deposit a significant portion of its energy in the form of ionization clusters, which lead to clustered DNA damage (which include combination of SSB+Base damages). New simple or clustered DSBs lesions (with dirty ends (check figure 8), and base damage) can form, starting from these damage clusters, either immediately (after chemical modifications of sugar lesions that initially do not disrupt the sugar-phosphate backbone), or following the enzymatic repair of base damage.

Damage to DNA can occur mainly through two mechanisms: (i) *direct effect* or (ii) *indirect effect* [77]. The *direct effect* is a one-electron oxidation reaction, e.g., direct absorption of a photon that results in a "secondary" electron release by an atom, which interacts within the DNA creating a radical anion or cation in the molecule. In contrast, the *indirect effect* results from the formation of free radicals due to the radiolysis of water (mainly $OH^{\bullet}$, $H^{\bullet}$, e^{-}_{aq}), sugar-based radicals (produced by H-abstraction), and other cellular components within the cell; such free radicals, including the secondary electron, then diffuse and can attack chemically-exposed sites in the DNA. For UV, x-ray or gamma-ray radiation the indirect effect accounts for 60-70% of the DNA lesions, while the direct effect contributes the remaining 30-40% [78], the proportions being reversed for proton or high-energy particle radiation, where the direct effect becomes largely dominant. Both these effects can create various forms of DNA damage, necessitating different DNA repair pathways.

Free-radical attack can oxidize the nitrogenous bases within the DNA double helix, producing such typical species as oxidized guanine (known as 8-oxoG), or deaminated cytosine (which turns to uracil, U). Such defects are highly dangerous, since after replication

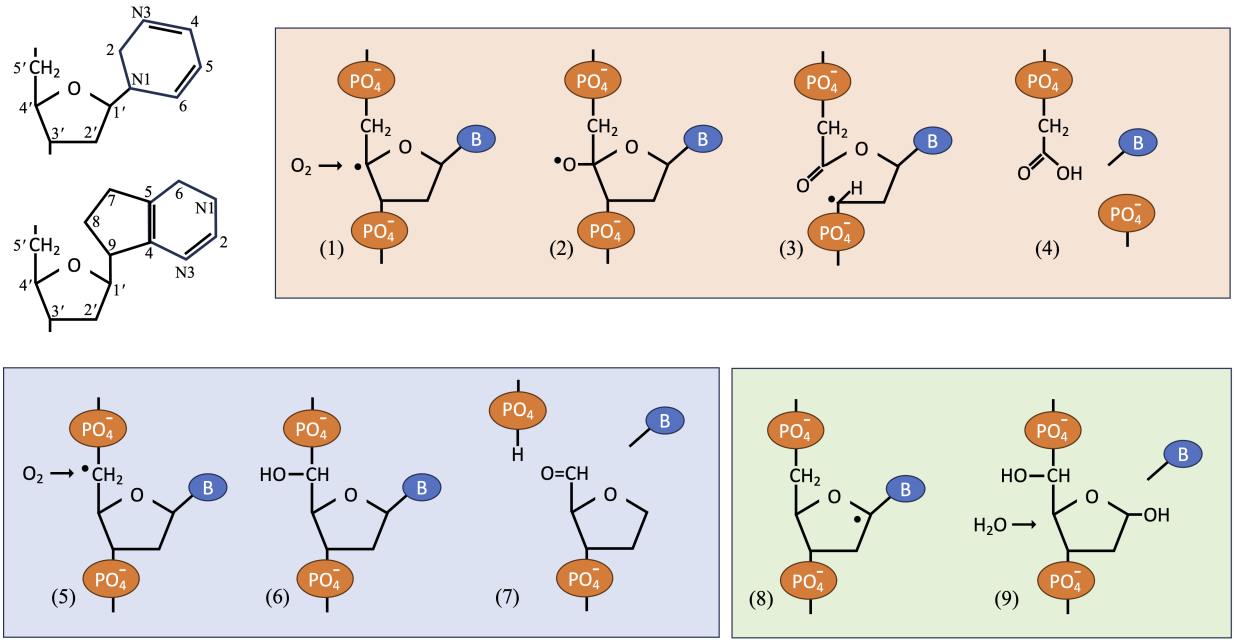


Figure 8: Top-left: numbering convention of the C and N atoms in the nucleotides. (1-4) Direct ionization at C4′ in presence of O_2 leads to SSB with base release, and a phosphoglycolate (CH_2COOH) at the 3′ terminal. (5-7) Direct ionization at C5′ in presence of O_2 leads to SSB with base release, and a aldehyde (HC=O) at the 5′ terminal. (8-9) Direct ionization at C1′ followed by hydrolysis leads to base release and a AP-site.

they can originate inheritable mutations in the DNA; notably, both the 8-oxoG·C and the U·C pairs will produce a C·G→A·T mutation [79]. Instead, nucleophilic radical attack to the sugar-moiety of DNA, mainly leads to the formation of abasic sites and strand breaks (cuts of the phosphate backbone). Abasic sites occur when the N-glycosidic linkage between the sugar and the nitrogenous base is disrupted, resulting in the lack of base while the backbone remains intact [77]. On the other hand, strand breaks can occur due to the transfer of free radicals along the DNA molecule, either in the presence or absence of oxygen, where they attack typically the C5′ and C4′ carbon atom sites [77]. The strand break can occur on one, or both sides of the double helix (**single- or double-strand break**, respectively **SSB** or **DSB**). In the presence of oxygen, the result is the formation of "clean" breaks, terminated with 3′–phosphate and 5′–OH, and less likely with 3′–OH and 5′–phosphate. However, it has also been observed that attack on the C5′ or C4′ carbon can give rise, respectively, to 5′–aldehydes, or 3′–phosphoglycolate (PPG), or other more chemically complex ("dirty") ends [80, 81]. It is worth noting that dirty ends are virtually unprocessable, since DNA-ligases cannot reseal a SSB or a DSB unless the ends are clean [82]. For example, tyrosyl-DNA phosphodiesterase requires several hours to remove PPG from 3′ DNA termini [77, 83] (see e.g. Figure 8).

It has often been observed that low linear-energy transfer (LET) radiation (e.g., x-rays and γ-rays) can lead to two or more ionizations within 1-4 nm apart, termed "ionization clusters", which result in multiple DNA lesions. Such multiple lesions can include more than two SSBs, AP-sites, or DSBs formed within the same, or the opposite strand from the same energy-deposition event (see again Figure 7). These "clustered" lesions are rarely formed by endogenous stress, rather they are primarily contributed by the toxicity of ionizing radiation (see Table 1). [75, 84].

Type of damage	**Radioinduced**	**Endogenous**
Single strand breaks	1000	> 10000
Base damage	2000	3200
Abasic sites	250	12600
Double strand breaks	40	40–50
DNA-protein cross-links	150	?
Non-DSB clustered lesions	122	?
Complex DSB	?	?

Table 1: DNA damage (N/(cell·Gy)) induced by ionizing radiation vs. endogenous damage due to normal cell activity.[84]

1.4 DNA DAMAGE REPAIR PATHWAYS

Irrespective of their origin, DNA damage will trigger specific repair pathways known as DNA-damage-repair (DDR), to defend the cell against genomic instability and accumulation of mutations. According to the type and amount of damage, cells have developed multiple and complex DDR pathways, which enable the cell to identify the damage, and repair it generally with a very high success rate. Even though these pathways are complex, they all follow a common workflow performed by certain specific molecular components, which can be summarized in the following steps [85]:

1. Recognizing the DNA damage - This is performed by DNA damage sensors (like poly-(ADP)-ribose-1 protein; Ku 70/80 Protein etc.[86]))

2. Recruiting repair proteins and excising the damaged segment - This is performed by transducers (like γH2AX, ATM, ATR etc [87]), which include a repair protein bound to a scaffold protein forming a complex at the site of lesion. This complex thereby removes the damaged segment.

3. Re-synthesizing missing parts - This is done by effector proteins, which add new nucleotide and ligate the broken ends.

Among the several distinct pathways dedicated to repairing different types of DNA damage, **Non-Homologous End Joining (NHEJ)** and **Homologous Recombination (HR)** are critical for the repair of double-strand breaks (DSBs), while **Base-Excision Repair (BER)** and **Single-Strand Break Repair (SSBR)** address damage to individual bases, and single-strand breaks, respectively.

It is estimated that a human cell repairs 10,000–20,000 DNA lesions each day. Among these, SSBs are are one of the most frequent, spontaneously-occurring lesions [88]. It is widely assumed that SSBs should be far less lethal than DSBs, which is more complex for both its identification and repair. Indeed, SSBs are repaired relatively quickly (Figure 9) [76, 89]. However, it has been observed that **unrepaired SSBs** (see the saturation at long times in the figure, for both SSB and DSB) can interfere with gene transcription and can evolve into more toxic DSBs during DNA replication; eventually, these combined events can lead to programmed cell death, or "apoptosis" [90, 91]. The origins of unrepaired damage, and notably SSBs, is one very active area of biomedical research [92–94], and while several possible causes have been indicated, there is not a general consensus on the persistence of DNA damage across cell replication, and its effects. It is also important to realize that the amount of SSBs generated during the cell exposure to endogenous/exogenous insults plays a major role. For instance, during an IR-exposure, around 10 more SSBs are generated for each DSB [95]. It is estimated that a IR-dose of 3 Gy leads to the formation of ≈600 SSBs and AP-sites per hour, for each DSB, in normal human cells [96]. Monte-Carlo simulation studies of DNA damage due to proton attack, emphasized that at higher proton energy about twice the SSB are observed, compared to the fraction of DSB [97]. Notably, even a small fraction of non-repaired SSBs can interfere with gene transcription and replication [90, 96]; most likely, non-repaired SSBs in proliferating cells are found to lead to the blockage or collapse of DNA replication forks during the S-phase; such defects can possibly evolve into special one-ended DSBs, whose repair is harder and leads to genomic rearrangement [90, 98]; eventually, these combined events can lead to apoptosis, or to cell senescence [90, 91].

Altogether, any SSBs which go unrepaired, or are "tolerated" during the entire repair pathway, have a potential to change into a DSB. Considering the natural prevalence of spontaneously-formed SSBs [99, 100], it is therefore understandable that cells should have developed highly effective mechanisms to reduce their effects. To ensure that SSBs are entirely repaired before the DNA replication, cells need to detect the SSBs, along with inducing the DDR pathway and delay the cell cycle progression, but until now there is no evidence supporting the existence of such a signaling pathway [101]. Hence, under-

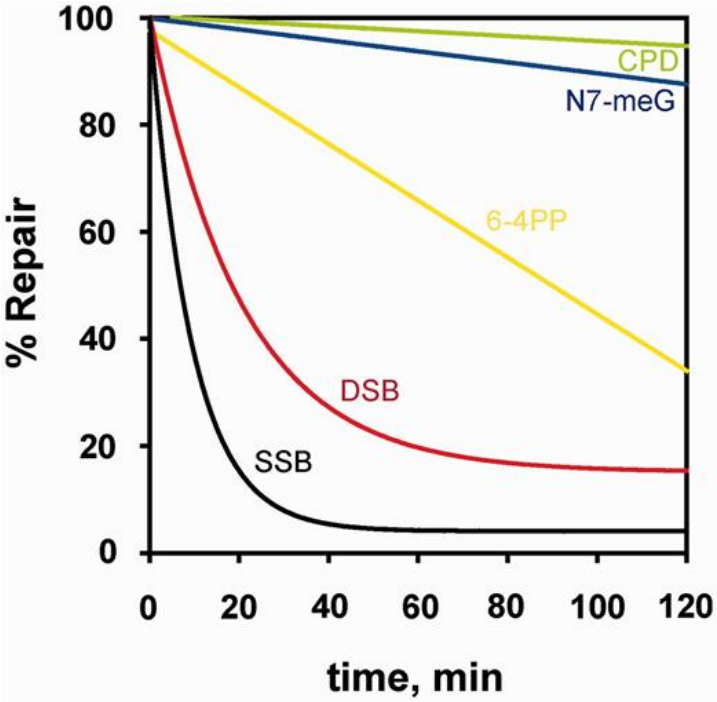


Figure 9: Repair kinetics of different types of DNA lesions. The plots show the kinetics of damage removal from CHO-AA8 cells of, respectively: SSBs, DSBs, 6–4 photoproducts (6–4PP), cyclobutane pyrimidine dimers (CPD); and for human lymphocytes, of N7-methylguanine (N7-meG). (Adapted from. [76])

standing how cells navigate this intricate repair process is not only crucial for grasping a fundamental cellular biology issue, but also for uncovering potential therapeutic strategies in disease management. Among others, a thorough understanding of the base-excision repair (BER) and single-strand break repair (SSBR) is essential, as these processes are integral to preventing genomic instability and ensuring cellular homeostasis. The upcoming Section will delve into the molecular intricacies of BER and SSBR to highlight their pivotal roles in safeguarding DNA integrity.

1.4.1 *Base-Excision Repair pathway*

BER pathway rectifies small base lesions that do not significantly contribute to the distortion of the DNA helix. ROS species mainly induce these small lesions, resulting in oxidized base lesions [102]. BER requires mainly 4 or 5 enzymes to carry out the DNA repair process (see Figure 10). Depending upon the type of lesion, it is initiated by removing the damaged base by one of the 11 glycosylases (out of which a few, called 'bifunctional', include both glycosylase and AP-lyase activity, therefore do not require a further AP-endonuclease for processing the break). This process forms apurinic or apyrimidinic sites (AP) sites. The hanging phosphate is then cleaved by an

AP-endonuclease enzyme (for mono-functional glycosylases), which generates a 3′-OH and 5′-deoxyribose phosphate (dRP) group at the break site. The subsequent step of filling the single nucleotide gap can follow a long or short patch repair. Short-patch repair is considered more dominant; long-patch repair is observed in post-replicative BER initiated by UNG2 or NEIL1 glycosylases, expressed during the S-phase. Long-patch repair is also observed when Polβ cannot remove the 5′-dRP terminus (e.g., if the dRP is oxidized or reduced).

In long-patch BER, the damaged nucleotide, along with several adjacent nucleotides, is replaced through the action of polymerase-δ (Polδ) and polymerase-ϵ (Polϵ) in conjunction with proliferating cell nuclear antigen (PCNA). The displaced strand is excised by Flap-endonuclease 1 (FEN1), and the DNA backbone is subsequently sealed by DNA ligase-I (Lig-I). Differently from the third step of short-patch repair, which does not involve the addition of several nucleotides, DNA polymerase-β (Polβ) adds a single nucleotide to the damaged site, creating a nick between two adjacent nucleotides. Finally, the nick is ligated by the DNA ligase-III, along with the help of X-ray repair cross-complementing protein 1 (XRCC1), which serves as a scaffold for ligase III.

The decision-making process between the choice of short-patch repair avsnd long-patch repair pathway is still under research [103]. Long-patch repair pathway is largely observed in proliferating cells and use of replicating proteins like Pol-ϵ/δ, FEN1, PCNA and Lig-I [104]. One of the hypothesis proposed is that the choice depends on the ATP concentration near the defective AP-site, which is controlled by the Lig-III and XRCC1. The study by Petermann et al. [105] shows that the long-patch repair is preferred in presence of higher ATP concentrations. A few other studies purported that the choice depends on the initiating glycosylase available, and the type of damage. For instance, if the 5′–dRP intermediate produced by APE1 is successfully removed by the Pol-β, then short-patch BER is preferred, and otherwise it follows the long-patch repair pathway [106, 107].

1.4.2 *Single-Strand Break Repair pathway*

SSBs can arise directly via the hydrogen abstraction and sugar fragmentation of the oxidized deoxyribose by the free radicals. This could either result in phosphoglycolate or phosphate termini [80]. During endogenous stress, SSBs are produced indirectly as an intermediate step of the BER pathway (either due to the action of APE1 at the abasic site or by the combined activity of bifunctional DNA glycosylase) [99]. Notably, most SSBs are accompanied by the loss of a single nucleotide at the strand-break site.

Irrespective of their origin, single-strand breaks are sensed by PARP-1.

Irrespective of their origin, all the SSBs are sensed by **poly-(ADP-ribose) polymerase-1 (PARP-1)** enzymes, which are catalytically ac-

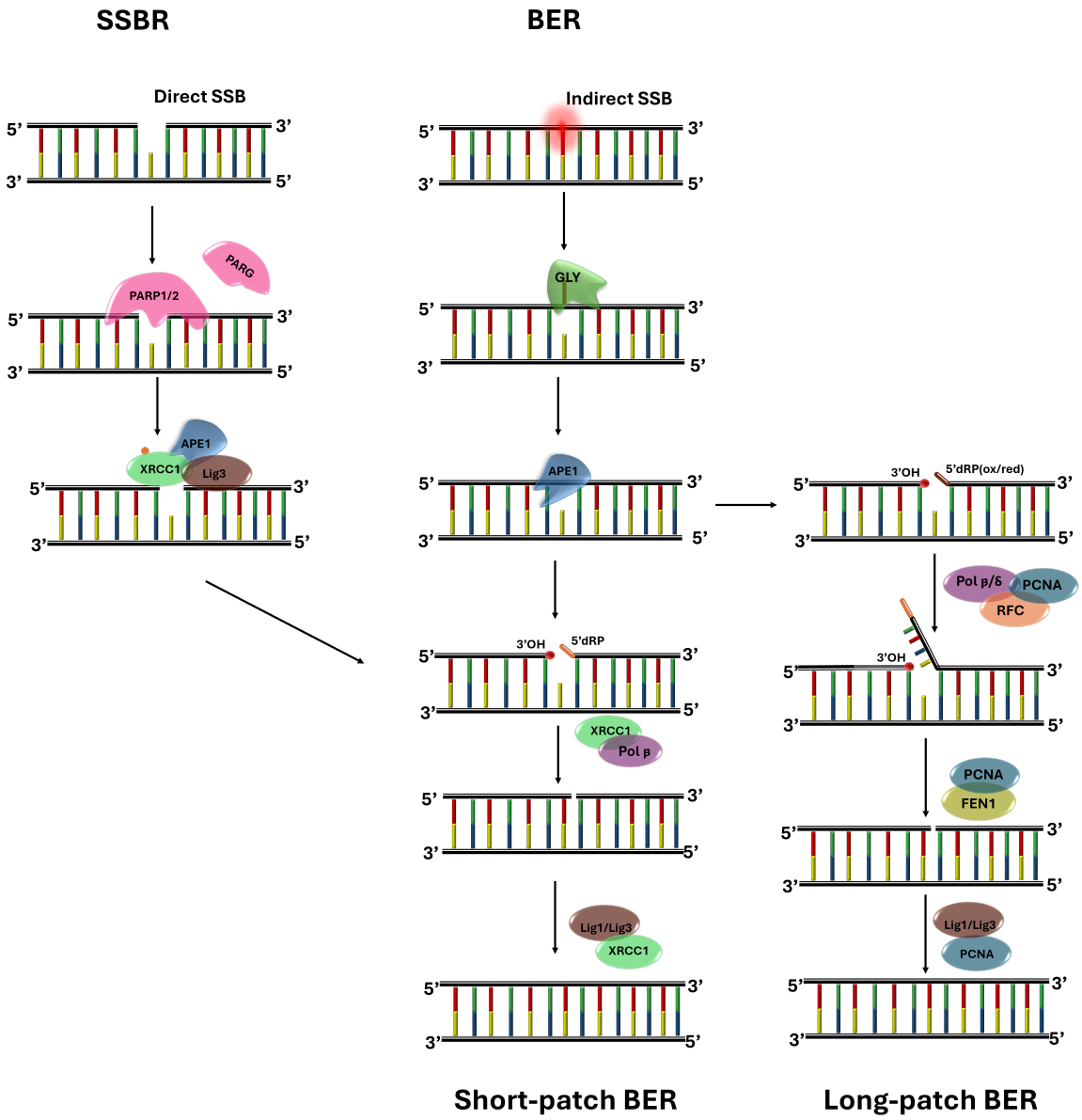


Figure 10: Base-excision repair (BER) is initiated by a glycosylase (GLY) that identifies and removes the wrong base, thus creating an abasic site; most abasic sites are processed by APE1, which cleaves the site and recruits Polβ. Polβ inserts the new nucleotide and repairs the 5′-deoxyribose phosphate (dRP) left by APE1: this creates a clean single-strand break (SSB). SSB-repair (SSBR) is initiated by PARP-1, which (after a sequence of enzymatic reactions) serves as platform to recruit APE1, XRCC1 and other proteins; at this point the two pathways join, and can proceed to completion through either short- or long-patch BER. The strand break (nick) is sealed by XRCC1-Lig3α. In cases where Polβ cannot remove the 5′-dRP (such as when it is oxidized or reduced), the repair is stalled. At these stalled sites, Replication Factor-C(RFC), Pol-δ/ϵ, PCNA are recruited to extend the gap by several nucleotides, and FEN1 cleaves the resulting flap. Ligation is then completed by Lig1. Direct single-strand breaks (SSBs) from sugar damage, and some SSBs arising during BER, are recognized by PARP-1 or PARP-2. PARP then recruits phosphorylated-XRCC1 and Lig3α to form a scaffold for repair. APE1 or PNK processes the damaged termini into 5′-phosphate and 3′-hydroxyl groups, aided by XRCC1. Polβ fills the gap with a nucleotide, followed by ligation by Lig3α.

tivated. Once activated, PARP-1 transfers ADP ribose units to itself or to other proteins, producing a long chain of poly-(ADP-ribose) (PAR) units (see Figure 10). The accumulated PAR chains, in turn, favor the recruitment of XRCC1. XRCC1 is then phosphorylated, and recruits the repair enzymes. It is also worth noting that these PAR chains are rapidly degraded by poly-(ADP-ribose)-glycohydrolase (PARG), and PARP-1 is recycled for subsequent detection of SSB. As mentioned, once XRCC1 is phosphorylated, it acts as a scaffold protein for recruiting enzymes required to repair SSBs. End-processing enzymes like polynucleotide kinase phosphatase (PNKP) and aprataxin (APTX) interact with phosphorylated XRCC1, DNA, polβ and DNA ligase III. PNKP resolves special terminals like 3′-Pho or 3′-PPG formed during irradiation. At the same time, Pol-β colocalizes with XRCC1 to help with polymerase activity. Eventually, the entire process is terminated by nick-sealing activity of ligase-IIIα [108]. Hence, the final few steps of the SSBR pathway are part of the short-patch BER pathway. Hence, these two meticulously designed pathways are crucial for repairing the "spontaneous" SSBs and oxidative damages in the DNA.

1.5 SINGLE-STRAND BREAKS AND CELLULAR SENESCENCE

Irradiation of tumoral volumes can impact various cellular entities through both direct and indirect mechanisms. As explained in Section 1.3, direct damage occurs when radiation interacts with the DNA itself, while indirect damage is mediated by the radiolysis of water, leading to the formation of reactive-oxygen species (ROS). These ROS can inflict damage on proteins, lipids, and nucleic acids, including DNA, which can result in both single-strand breaks (SSBs), double-strand breaks (DSBs) and base oxidation [109]. Traditionally, it was believed that DNA damage would be observed only in the targeted tissue volume, due to the direct interaction of radiation. But this model was questioned once chromosomal damage, translocations and mutations were observed also in cells at nearby/distant tissues [110, 111]. Hence, oxidative stress is not confined to the targeted irradiated area, but can also affect neighboring, non-irradiated cells through intercellular communication, a phenomenon known as the "bystander" effect [112]. This could also increase the carcinogenic risk of the distant tissues, resulting in so-called **second primary cancers (SPC)** [113].

These radiotherapy-induced second cancers are not to be intended as the recurrence of initial cancers. Instead, they originate at newly transformed cancerous cells, from the normal cells which were affected by the radiations. These are seen to develop preferentially near the margins of the irradiated area, rather than in the planning target volume (PTV, where 100% of the irradiation dose is theoretically administered). These neighboring areas are exposed to scattered radiation from various sources, such as leakage from the machine's col-

limators (which shape the beam), the flattening filter (which ensures uniformity of the radiation dose), and scattering that occurs inside the patient's body [114]. The internal scatter was identified as the primary factor influencing the dose deposited in the closest margins around the target area [114]. This marginal radiation (also known as "out-of-field dose") has three key traits: (i) the dose decreases exponentially as the distance from the target increases, (ii) the dose is roughly proportional to the size of the PTV, and (iii) the energy spectrum of the radiation is shifted toward lower energies compared to those within the PTV [114].

In addition to their spatial relationship with the PTV, another remarkable feature of SPCs is their extended latency period, which can range from a few, up to tens of years following the initial cancer treatment [113]. This indicates that normal cells affected by scattered radiation may remain dormant in the body while still retaining the potential for neoplastic transformation. However, the exact biological nature of this dormant state remains unknown [11–13, 114, 115].

Unrepaired DNA damage can accumulate at the margins of irradiated regions.

This uncertainty has been a primary motivation for our research, seeking to explore the chemical and biological changes, particularly the **accumulation of unrepaired DNA damage** that may occur at the margins of the irradiated area. In our research groups, we aim to understand how these changes might induce a senescent state in cells at the PTV margins, potentially leading to the development of secondary sarcomas. The unrepaired SSBs can activate specific cellular pathways, as discussed below.

The ROS generated during irradiation not only cause SSBs but also down-regulate the expression of poly-(ADP-ribose) polymerase-1 (PARP-1), a key enzyme in the DNA damage response (explained in section 1.4.2). Normally, the accumulation of SSBs triggers the SSB repair (SSBR) pathway, which shares its final steps with the Base Excision Repair (BER) pathway, see Figure 11 [116, 117]. However, in keratinocytes undergoing spontaneous senescence, oxidative stress (e.g., induced by radiation) is found to down-regulate PARP-1 expression, reducing the synthesis of PAR, both in concentration and length of the PAR-chains. This reduction impedes the recruitment of CK2 to the damage site, arresting the repair process at the step of XRCC1 recruitment [11]. Similarly, in fibroblasts exposed to radiation doses outside the field, PAR synthesis is also diminished, though the molecular mechanisms remain unclear [13]. Recent studies suggest that out-of-field IR may directly influence the expression of enzymes such as nicotinamide phosphoribosyl transferase (NAMPT) and nicotinamide mononucleotide adenylyltransferase 1 (NMNAT1), which are responsible for synthesizing NAD+, a substrate required by PARP-1 for PAR synthesis during DNA damage. Additionally, out-of-field IR could increase the production of poly(ADP-ribose) glycohydrolase (PARG),

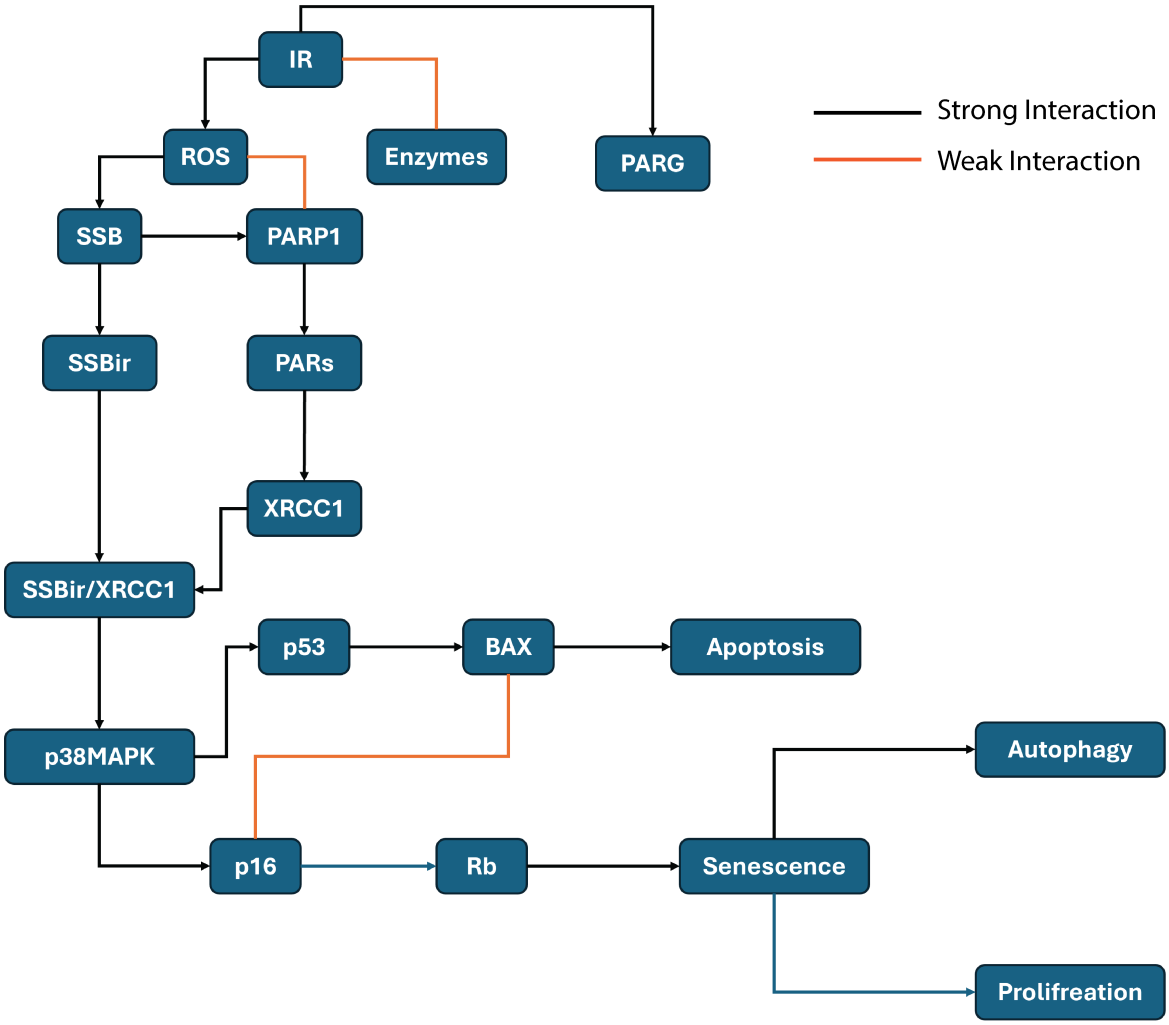


Figure 11: Reduced SSBR pathway followed by the keratinocytes in the margin of irradiated cell constructed using the experimental data (SSBir - Unrepaired SSB).[11, 13] This synthetic flowchart will be used in Chapter 5 to set up the COPASI mathematical modeling of the SSB repair.

which degrades PAR. As a result, PAR formation declines significantly.

Lack of PARP-1 and smaller PAR chains promote unphosphorylated XRCC1 and compromise SSB repair efficiency.

As opposed to normal repair conditions where phosphorylated XRCC1 follows the correct SSB repair chain, research has shown that low levels of PAR, or PAR-chains of shortened length, are sufficient to hamper phosphorylation of XRCC1. This unphosphorylated XRCC1 accumulates at SSB sites and blocks the repair process (see the SSBir/XRCC1 block in Figure 11, probably because unphosphorylated XRCC1 cannot dissociate from DNA [118]); unrepaired SSB triggers the p38 mitogen-activated protein kinase (p38MAPK) pathway, leading to the up-regulation of p16, a key factor in cell cycle arrest [12]. In fibroblasts exposed to out-of-field IR, it is also observed an up-regulation of p21, which contributes to cell cycle arrest. However, the exact mechanism driving this up-regulation remains unknown. Once

the p38MAPK pathway is activated, it can promote either p16 or p53 pathways. While p53 typically leads to apoptosis, p16 drives cells towards senescence. Experimental data indicate that cells with accumulated unrepaired-SSBs preferentially enter senescence over apoptosis, possibly due to the inhibition of p16 on BAX, a pro-apoptotic factor [119]. This preference for senescence over apoptosis is particularly evident in conditions where SSBs are prominent. Senescent cells are typically cleared by autophagy, a major lysosomal degradation pathway. However, in certain cases, such as irradiated fibroblasts located near the radiated volume, cells may enter a state of premature senescence and later escape into a proliferative state which is also termed as inadequate senescence/escape from senescence.

1.6 LAYOUT OF THE THESIS

Given the above overview of cellular senescence, its link with DNA damage, and its relation to cancer, our first step was to zoom in on the intricate details behind identifying the DNA strands breaks and their evolution. For this purpose, this thesis is subdivided into five Chapters. The first Chapter (that you are presently reading) gives a summary of the state-of-the-art research conducted during my PhD, providing a comprehensive overview of how DNA damage can induce cellular senescence, and its eventual link to secondary cancers. This introductory study is now published in our review paper: *Cell senescence and the DNA single-strand break damage repair pathway* [40]. We then started by looking at the very initial steps of the pathway i.e. how nuclear repair proteins identify the DNA damage, and notably strand breaks.

Hence, Chapter 2 focuses on the molecular mechanisms by which PARP-1 protein identifies single-strand breaks (SSBs), studied by a substantial amount of molecular-scale computer simulations. This objective was selected because it represents the initial step in the single-strand break repair (SSBR) pathway, and the mechanistic details of PARP-1's interaction with SSBs remain largely unexplored in the literature. This Chapter aims to offer a detailed analysis of the complex dynamics and molecular interactions involved when PARP-1 engages with a hairpin-fragment of DNA containing an SSB, which has been repeatedly used in experiments as a model system. The results of this study are now published in our paper: *Cooperative dynamics of PARP-1 zinc-finger domains in the detection of DNA single-strand breaks* [120].

After gaining insights into how PARP-1 interacts with the idealized hairpin-DNA structure, I expanded the focus to the more complex system of the nucleosome, providing a better representation of DNA damage in the nuclear environment. Our goal was to understand the evolution of SSBs and DSBs, both within the isolated nucleosome (Chapter 3), and within a dinucleosome structure, with

varying linker-DNA lengths (Chapter 4). I employed Steered Molecular Dynamics (SMD) to explore the dynamics of the dinucleosome when damage occurs localized in the linker-DNA region. This approach allowed to investigate how different types of DNA damage lead to structural deformations within the nucleosome (Chapter 3), likely facilitating the recruitment of DDR proteins to the SSB site. Chapter 4 provides a comparative analysis of the mechanical deformation caused by SSBs and DSBs, offering a quantitative assessment of how these lesions affect nucleosomal integrity and dynamics. Part of these results are submitted in a joint paper: *How repair proteins identify DNA damage in the nucleosome*, under review in the journal *Biomolecules* (preprint available as Ref.[121]).

The final Chapter 5 shifts focus from the nucleosome level to a broader examination of the entire DNA damage repair chain, in the modern framework of *systems biology*. Here I employed two complementary approaches to build mathematical models of the single-strand break repair (SSBR) pathway, and notably the role of unrepaired damage. Firstly, I used the COPASI software with Michaelis-Menten, Hill or mass-action kinetic models, to explore the cell decision-making processes between repair, apoptosis and senescence end-points, involved in the damage response. The second approach involved the GEANT4-TOPAS Monte Carlo simulation framework to simulate the whole chain of radiation-induced damage, including SSBs and DSBs, with the help and expert guidance of Dr. Nicholas Henthorn from the PRECISE group at Manchester University. These models provided two key insights: (1) how competition between different types of DNA damage influences the decision-making process of the DNA damage response (DDR) pathway, and (2) the possibility of integrating the SSBR pathway into the existing DaMaRiS code for a more comprehensive understanding of DNA repair.

Hence, the ensemble of the thesis is primarily focused on computational simulations and mathematical modeling of protein-DNA interactions at the molecular scale, with an extension to modeling of protein network pathways, and supplemented by experimental details where possible. The focus on the dynamics of DNA damage, on the early molecular steps of damage recognition by a typical repair protein such as PARP-1, and the DNA damage response (DDR) pathway, should make this approach a valuable study also in the more general context of biological studies of the complex mechanisms of DNA repair.

2

INTERACTION OF PARP-1 WITH SINGLE-STRAND BREAK IN A MODEL DNA-HAIRPIN

2.1 INTRODUCTION

The PARP-1 protein (Figure 12a) is a well studied member of the PARP family. It is a chromatin-associated protein consisting of at least six functional domains: three DNA-binding Zinc-finger N-terminal domains named Zn1, Zn2 and Zn3; one BRCT domain; one WGR domain; and one catalytic C-terminal domain, including a helical subdomain (HD) (Figure 4c). When PARP-1 is not bound to DNA, Zn1 and Zn2 behave as flexible independent domains [122]. PARP-1 binding to damaged DNA activates a complex sequence of allosteric and cooperative effects between the different domains, which are not yet completely elucidated. In most experimental studies to date only fragments of the whole protein were used, often leading to divergent models, especially whether PARP-1 recognizes and binds to damaged DNA as a monomer or a dimer. Zn1 and Zn2 are known to specifically recognize DNA breaks (notably, both SSB and DSB). Zn1 from one PARP-1 copy may also cooperate with Zn2 from another PARP-1 protein to form a dimeric module that specifically recognizes DNA breaks [123]. On the other hand, Zn3 mediates as an inter-domain contact and is required to confer with PARP-1 to regulate chromatin structure [124]. The BRCT domain acts also as a DNA binding domain, but of low affinity, and is able to bind to only intact DNA without concomitant catalytic activation. The BRCT-DNA interaction mediates DNA intra-strand transfer of PARP-1 (the so-called "monkey-bar mechanism") that allows rapid movements of PARP-1 through the chromatin [125]. By analogy with PARP-2, it is assumed by many authors that the WGR domain of the sister protein domain of PARP-1 can bridge two nucleosomes with the broken DNA ends aligned in a position suitable for ligation. Such bridging induces structural changes in PARP-1 that signal the recognition of a DNA break to the catalytic domain of PARP-1. This promotes the recruitment of Histone PARylation factor 1 (HPF1) and subsequent activation of PARP-1, followed by licensing serine ADP-ribosylation of target proteins [126, 127]. HD prevents effective NAD^+-binding in the absence of an activation signal. However, after binding to damaged DNA, the autoinhibition is relieved, HD unfolds, and PARP-1 becomes able to bind NAD^+, thus starting PARylation [128, 129].

Some studies already shed light on the structural features of Zn1 and Zn2 domains interacting with DNA breaks. In the earlier studies,

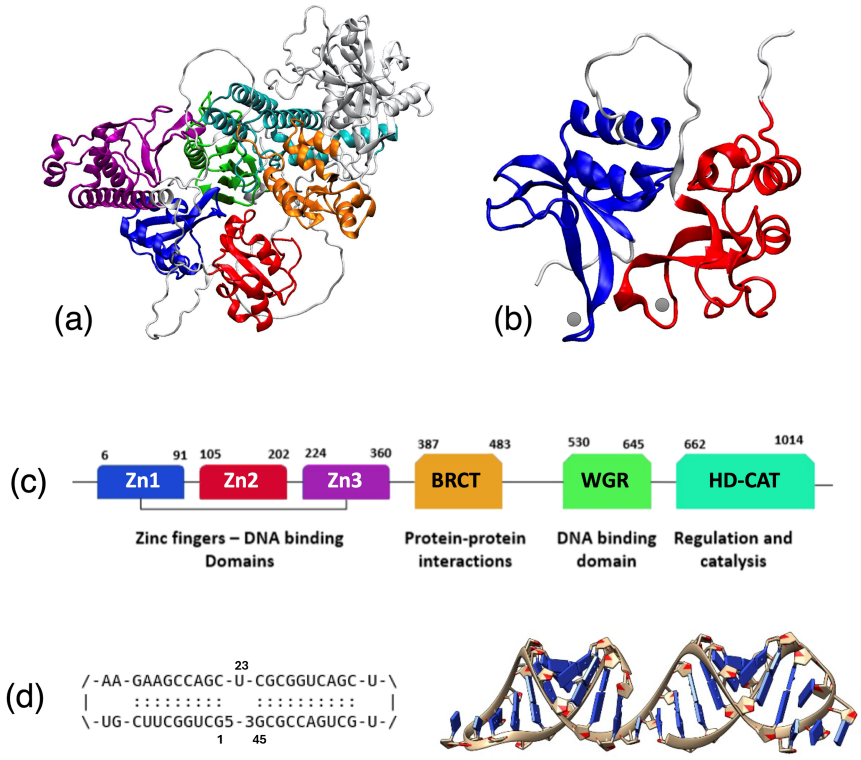


Figure 12: Molecular structures of the PARP-1 protein and of the DNA hairpin. **(a)** Structure of full length PARP-1 as modeled by AlphaFold2. Domains: Zn1, blue; Zn2, red; Zn3, purple; BRCT, orange; WGR, green; PARP-α-helical, cyan. **(b)** Details of the two zinc-finger domains, Zn1 (blue) and Zn2 (red), that recognize and interact with the DNA single strand breaks. The position of the Zn atoms (grey spheres) is tentatively indicated. **(c)** Domain structure of PARP-1; the numbers above the diagram indicate the amino acid positions. **(d)** Blueprint of the RNA sequence used as input in ROAD along with the nucleotide positions 1,23,45, and the resulting 3D-structure of the straight DNA hairpin, with a missing nucleotide to simulate a single-strand break.

single PARP-1 fragments (with either Zn1 or Zn2) were co-crystallized with short dsDNA fragments, ideally mimicking DSBs with blunt ends [130]. The results showed that both Zn1 and Zn2 can make contact with the DNA ends through interaction of Phe44 and Lys151 respectively with the 5′ and 3′ ends of DNA. However, Zn2 showed a higher affinity than Zn1 in binding to DNA ends. A subsequent study, in which dsDNA terminated with overhangs at both ends were used [123], and co-crystallized with separate Zn1 and Zn2 domains (ideally representing the N-terminals of two adjacent PARP-1 molecules), found that Zn2 interacts with *both* the 5′ and 3′ ends, tending to exclude Zn1 from the direct interaction but requiring the cooperation of both Zn fingers for the efficient tagging of the damaged DNA *in vivo*. Eustermann et al. [128] reported a high-resolution structure of PARP-1 Zn-finger domains in complex with a SSB in the hairpin DNA. The authors used a hybrid nuclear magnetic resonance and X-Ray

diffraction (NMR-XRD) technique to produce a detailed atomic structure of paired Zn1-Zn2 domains in interaction with a 45-nucleotide DNA hairpin containing a SSB. They showed how Zn2 initiates the recognition of the SSB by first binding to the 3′ end, followed by Zn1 (from the same PARP-1 monomer) making contact toward the 5′ end. This coupled action appears to lead to a drastic bending of the DNA around the SSB site, by ~130°, an effect that was also observed in older experiments by electron microscopy [131]. Very recently, the same group also performed single-molecule Förster resonance energy transfer (smFRET) analysis on the same DNA-hairpin with SSB interacting with Zn1 and Zn2 [132], showing that just the co-presence of Zn1-2 (even excluding the rest of the protein) is sufficient to bring about a kink in the DNA at the break site; it appears that Zn2 recognizes the damage by producing a small kink, and additional binding of Zn1 induces a sharper kink in the DNA.

Even though these structural studies helped in understanding many details of the binding of PARP-1 to damaged DNA and its consequent conformational changes, there is no clear-cut answer about the mechanism by which the change in conformation of DNA occurs. Moreover, the initial steps of the damage recognition are not established, since the biochemical techniques used until now only give access to final-state conformations. By combining molecular docking and molecular dynamics (MD) simulations, I tried to understand the pristine interactions of PARP-1 with the damaged DNA, and the dynamics of kinking-bending that leads to the stable binding between the enzyme and the damaged DNA site.

Strand breaks in DNA are the result of a complex sequence of biochemical modifications, starting from the simple breaking of a bond and proceeding in several steps with the progressive disruption and final removal of the damaged nucleotide(s). Hence it is crucial to understand how, and at which step, the early SSB detection is activated by the sensor proteins that continuously search the DNA for structural anomalies. In the following, I used the same DNA-hairpin structure of Ref.[128] as the reference system, to study the interaction of PARP-1, both in monomeric and dimeric form, with a model-SSB represented as a whole missing nucleotide in the center of the DNA hairpin. Such a structure is representative of the direct damage to the DNA, as can be obtained as a result of the rapid chemical evolution of the initial bond-breaking typically produced at the C4′ or C5′ carbon by radical attack.

To elucidate the molecular interactions between the zinc-finger domains of PARP-1 and the DNA damaged by a SSB, I designed a complete simulation protocol, starting with (i) the design of the atomic structure of the DNA and protein fragments, then (ii) docking of the rigid fragments (protein to the DNA) with an approximate free-energy functional, followed by (iii) all-atom molecular dynamics sim-

ulations and analysis with a state-of-the-art empirical force field, and finally (iv) comparison with the available experimental results. The details of the computational methods (for this, and the next Chapters) are provided in Appendix B, Section B.1, also including a description of Zn-finger proteins and the re-parametrization of the AMBER force-field that was necessary to include the corresponding topologies.

Since, the reported NMR experimental structure only presents a final state with a strongly deformed hairpin, as a result of the interaction with the protein, I constructed also a straight DNA hairpin, with the same nucleotide sequence, SSB damage and 2D contacts (See figure 12d). By using this at the initial configuration, it was also possible to follow the process of SSB recognition from the earliest stage of the interaction. To design the DNA hairpin in the different conformations I used the RNA-Origami Automated Design tool (ROAD), a computer-aided design software to build automated 3D-structures of nucleic acids [133], as described in the protocols Appendix B. The ensembles of DNA-proteins were solvated in a relatively small water boxes, of size 15x15x15 nm^3, with about 106,500 water molecules, and about 365,000 atoms.

2.2 Stability of the isolated DNA hairpin including an SSB

Initially it was necessary to test the thermal and mechanical stability of the model DNA hairpin including a simulated SSB. In particular, I wanted to verify whether the DNA fragment could bend spontaneously into the experimentally observed shapes [128], under the simulated thermodynamic conditions, and without the participation of PARP-1. For this purpose, I performed 2μs-long MD simulations for three variations of the model DNA: **(1)** the 45-nt hairpin from Ref.[128], with a missing nucleotide at its center representing the SSB, with the same nucleotide sequence and terminations as the experimental structure (5′-deoxyribosephosphate, dRP, and 3′-OH, Figure 4d); **(2)** hairpin DNA with hydroxyl terminations on both 5′ and 3′ ends a configuration that may be representative of the SSB gap after the pol-β removal of 5′-dRP, and could be interesting to compare the effect of a wider opening of the gap; **(3)** a long stretch of free DNA with a missing nucleotide at one strand, including the same SSB gap as in the hairpin DNA.

The MD simulation of the isolated hairpin **1**, shows that this DNA structure does spontaneously open up at the level of the SSB, and can bend substantially, over very short time frames. Even though the DNA largely fluctuates in the course of the simulation (Appendix A, Figure A1a,b), the average stable structure after cluster analysis (cut-off 0.6 nm, average RMSD 0.49 nm) remains comparable to the initial straight configuration (Figure A1c). We measured the gap opening

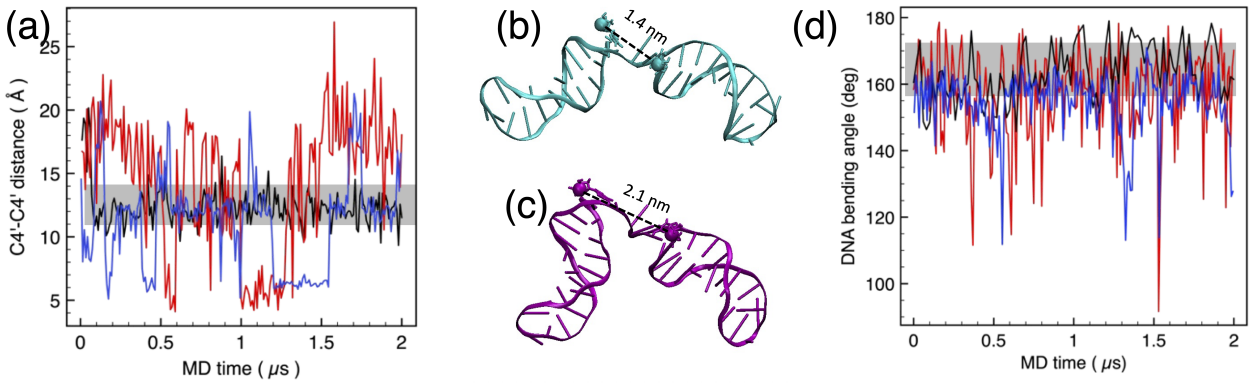


Figure 13: **(a)** SSB opening, measured as the distance between the C4′ atoms of the 5′ and 3′ guanines: straight long DNA (black), 45-nt hairpin with 5′-dRP (blue), and with 5′-OH (red). **(b,c)** Sample snapshots of the hairpin configuration at times 0.5 and 1.77 μs, showing typical frames of the intermediate (b) and wide-open conformation (c). **(d)** Bending angle, measured between the P atoms of nucleotides 12-24-34 (center and loop ends). To facilitate the comparison, the grey shaded areas highlight 95% of the fluctuation for the straight long DNA (black).

at the SSB by taking the distance d_g between the C4′ atoms of the guanine 1 and 45, respectively at the 5′ and 3′ terminals. The bending angle θ_b is measured by the relative position of the P atoms of the nucleotides 12, 24 and 34, at the two ends and approximate center of the DNA. Figure 13a,d displays the plot of these values for the 2-μs trajectories, the blue, red and black colors referring to the structures **1**,**2**,**3**, respectively.

The SSB gap displays large jumps between three different states: a "wide open" at $d_g \geqslant 2$ nm, an intermediate one at d_g ~1.5 nm, and a sort of "closed" state with a gap distance d_g=0.5-0.7 nm. The DNA bending angle experiences large fluctuations between θ_b=110° and 160° (Figure 13d), with instantaneous peaks at even smaller angles. The coupling between opening and bending indicates large fluctuations accompanied by a (less visible) twisting of the compact DNA structure. In particular, the peaks appearing at times t=1.58, 1.77, 1.83 and 1.87 μs, are correlated with the corresponding jumps in the RMSD (Figure A1b); the snapshots of the DNA hairpin in Figure 13b,c show an extensive bending and twisting of the structure in both the intermediate and wide-open configurations.

The comparison of structure **2** with dephosphorylated 5′-OH end (see red plots in Figure 13) is entirely similar to the structure **1**, and displays an even wider range of fluctuations both in the opening and in the bending angle. The RMSF-by-atom shows a distinct periodicity, with larger values of about 0.6 nm for the two 5′ and 3′ ends, the nucleotides 11-12 and 34-35 at the loops (up to 0.8 nm for the 11-nt), and the unpaired nucleotide 23 facing the gap. This indicates that the

two end loops in this particular experimental construct could help the spontaneous opening-up of DNA, both before and during the interaction with the PARP-1 protein.

For the sake of comparison, I extended the same stability test to the free-DNA structure **3** that could be representative, e.g., of a nucleosome linker (see next Chapter 4), or a transcriptionally exposed chromatin stretch. The chosen DNA length of 200 bp (about 70 nm) is longer than the typical DNA persistence length (~50 nm, [134, 135]). This structure includes a gap at the centre of the double strand (resembling an SSB) and no end loops. Evidently, such a long DNA displays long-wavelength fluctuations at the simulation temperature of 310K, as it can be observed by the time plot of the RMSD and the RMSF-by-atom, which fluctuate about values of several nm. However, when we looked at the gap opening and the DNA bending angle (see again Figure 13a,d) an important difference is observed between the free DNA (black plots) with respect to the same quantities measured on the 45-nt hairpin (blue and red plots). The gap opening in the free DNA remains at a constant value of about 1.2 nm with a small dispersion of $\pm$0.1 nm. The average bending angle in the free DNA fluctuates between θ_b=160-175°, that is practically straight.

Such results suggest that the large opening of the SSB gap apparently necessary to accommodate the PARP-1 fragment in the experiments using the DNA hairpin [128], could be due to - or at least be substantially helped by - its spontaneous thermal fluctuations. However, our results indicate that the wild bending and gap opening of the hairpin seems confined to shorter (tens of ps) times, while the average values remain close to the initial conformation.

2.3 BINDING OF PARP-1 MONOMER TO DNA HAIRPIN

The next step is to study the binding of PARP-1 to DNA containing a SSB. As detailed in the Appendix B, Section B.1, I used molecular docking with empirical free-energy functional (HADDOCK) to identify the candidate interacting conformers, followed by an all-atom MD at finite-temperature, to unveil the binding dynamics of PARP-1 and conformation fluctuations. As a first test to establish the validity and quality of the simulation protocol, I tried to retrieve the experimental interaction from Ref.[128]. For the sake of clarity, in the following protein residues are indicated by their 3-letter code and residue number, and DNA nucleotides by a single letter-number pair.

2.3.1 *Retrieving the experimental structure*

Two separate PDB structures were generated by isolating the protein and the DNA hairpin from the RCSB PDB entry **2N8A**, which were then reassembled by docking with HADDOCK. The active residues

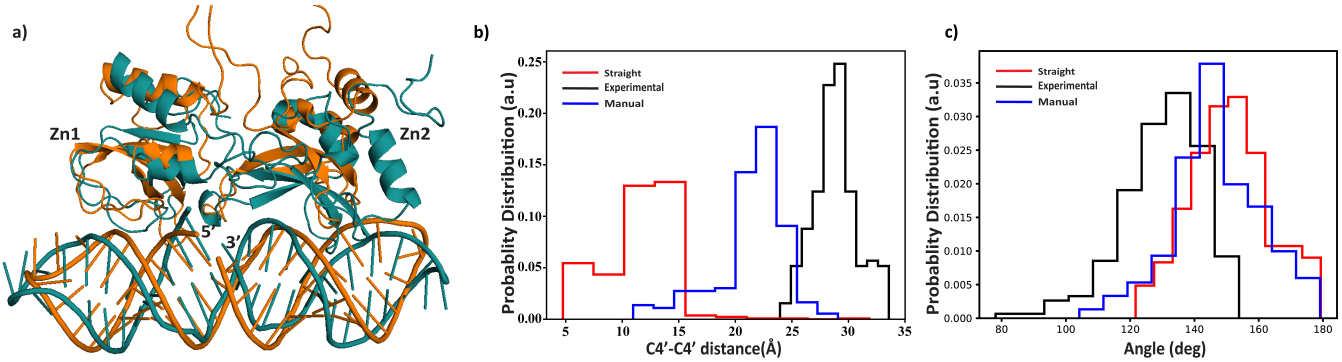


Figure 14: **(a)** Straight hairpin before (orange) and after (cyan) 200-ns MD simulation. **(b)** Distance between the C4′ atoms of the $3' - 5'$ ends: DNA starting from experimental structure (black) and free PARP-1 docking. DNA starting from the straight hairpin (red) and free PARP-1 docking. DNA starting from the straight hairpin (blue) and manual fitting of PARP-1. **(c)** Distribution of DNA hairpin bending angle measured between the P atoms of nucleotides 12-24-34; same color code as in (b).

to bootstrap the docking were: the residues Phe-44 and Leu-151 of the protein (belonging to Zn1 and Zn2 respectively), and the nucleotides G1 (5′ end of the hairpin), G45 (3′ end), plus the T23 (unpaired base facing the SSB gap) of the DNA. This resulted in docked clusters having some of the interactions that were already established in the experimental NMR structure. Then, based on the RMSD value, the best docking poses were selected as initial configuration, and performed a longer 0.5-μs MD simulation at T=310K (Figure 14a). The resulting distribution of the 5′-3′ distance d_g (i.e. the SSB gap) as measured by the distance between the C4′ atoms, (Figure 14b, black line) has a peak at $d_g \simeq 3.1$ nm, compared to 3.02 nm for the experimental NMR structure [132]. The distribution of the DNA bending angle, shown in Figure 14c (black line), extends between about θ_b=80-160°, peaked at about 135°, and rather skewed toward the lower values. By comparison, the distribution of experimental values obtained by FRET [132] is nearly Gaussian, peaked at $90 \pm 14°$.

In the simulated complex, a very close proximity between Leu-151 from Zn2, and the DNA nucleotide G45 at the 3′ end can be observed; and the Phe-44 from Zn1 in a position that is very close to the opposite DNA 5′ end, making contact also via Gln-40 and Pro-42 (a schematic of the contacts is provided in Figure A2a). Also, two hydrogen bonds within a cut-off distance of 0.3 nm were observed between Arg-18 of Zn1 and thymines at position 3 and 4; while Arg-122 of Zn2 makes hydrogen bonds with C41 and C42.

Some of these 'trademark' interactions were already described in the NMR study of Ref.[128]. Notably, it has been known for quite a long time [136, 137] that the interaction between phosphate groups of DNA and lysine, arginine and phenylalanine residues of TATA-BOX-binding proteins, helps in producing and stabilizing a sharp kink in the double-helix, and open up the DNA conformation, by inserting

two deformed α-helices in the minor groove. Similarly here, PARP-1 makes contact via Gln-150, Arg-122, Arg-18 and Phe-44, found in the Zn1 and Zn2 domains. Moreover, Zn1 digs in the SSB, and reaches for a contact with the unpaired T23 on the opposite DNA strand, using its base stacking loop(mainly Met-43 & Phe-44) (Figure A2a). All these interactions, along with the trademark interactions, appear to assist and maintain DNA bending and kinking [128, 132, 138].

Overall, the average conformation of PARP-1 after the long MD simulation is able to get to a structure rather close to the reference experimental one, as far as the contact regions with DNA are concerned. Also, the DNA maintains its strongly bent conformation, as shown from the distribution of bending angle in Figure 14c (black line). The global orientation of Zn1-Zn2 with respect to the DNA is slightly rotated by ~17° about the DNA main axis, compared to the experimental reference. However, this has minimal impact on the contacts between the protein and DNA.

2.3.2 *Starting from the straight hairpin*

The next step was to repeat the docking and simulation, but starting from the unperturbed straight conformation of the hairpin, an attempt to mimic the stage of PARP-1 randomly approaching the DNA region containing an SSB. The docking and selection of binding candidates were performed in the same way as above. However, in this case the docking of the whole protein fragment, that is, the two Zn1-Zn2 zinc-fingers connected together by a short stretch of amino acids, did give complex results. In the best docked configurations, Zn2 is found in close contact with the 3′ DNA end (G45), but Zn1 remains far from interacting.

Moreover, the relative position of Zn1 and Zn2 is such that Zn1 lies on the same side as the 3′ end, and therefore to get into a proper interaction with the 5′ end, as experimentally observed, it would require a large rotation by about 180° of the whole protein about the main DNA axis. Such results could be due to an HADDOCK limitation or an artifact, or could be indicative of a real physical-chemical issue in the interaction with the straight hairpin, which at this stage is difficult to assess. It is worth noting that the recent FRET experiments [132] using the same DNA hairpin configuration, identified a high affinity also for the Zn2 binding alone - albeit with somewhat lower values than for the Zn1-Zn2 complete pair - and vanishing affinity for Zn1 binding alone.

Hence, I reduced to select the best docking configuration having the "right" orientation of Zn1-Zn2 with respect to the SSB ends, even if this is not the one giving the best HADDOCK score and RMSD value. This configuration was then used as the starting input structure for a MD simulation of 0.5 μs at T=310K. This starting configuration has

Zn2 interacting with nucleotides G1, at 5′, and T23, the unpaired nucleotide facing the SSB. After the long MD run, it evolves into a more correct configuration (Figures 14a), with Zn1 approaching the 5′ end of the SSB. This is already a better result, albeit the closest residues to 5′ are Ser-41 and Asp-45, rather than experimentally-observed Phe-44. Moreover, Zn2 is penetrating further in the SSB opening, and Leu-151 interacts with T23. On the other hand, Zn2 has still a rather weak interaction with the 3′ end, via its Pro-149 and Gly-152 (Figure A3).

Clustering of the whole MD trajectory was performed, to extract the best representative conformations of the interacting system. A comparison of the opening d_g and bending θ_b of the SSB is given in the probability distribution histograms of Figure 14b,c, between the two simulations: (i) starting from the experimental reference (black), vs. (ii) starting from the straight hairpin (red). After MD simulation, the DNA still appears nearly straight ($\theta_b \sim 180^o$) and the gap closed ($d_g \sim 1$ nm), with little or no room for Zn2 and Zn1 to make contacts with the SSB ends, as it can be observed in the plots in Figure 15b,c. To understand if the initial orientation of Zn1-Zn2 could play a role in the SSB opening and kinking starting from the straight DNA, I therefore proceeded with manually adjusting the orientation of the protein into different interacting configurations, as explained in the next Section.

2.3.3 *Manually fixing the right DNA-PARP-1 contacts*

Given the difficulty of obtaining a starting docked structure corresponding to the experimental observations, I eventually tried to manually adjust the Zn1-Zn2 fragment to the DNA hairpin, starting from the docked structure and using CHIMERA manipulation tools. Initially Leu-151 from Zn2 was close to G45 at the 3′ end, while the Phe-44 was placed somewhat close to G1 at the 5′ end of the SSB.

Starting from this adjusted configuration, the first 150 ps of thermal equilibration were followed by a MD run of 200 ns. The best cluster, shown in the Figure 15a, has RMSD = 0.542 (cutoff = 0.65). This initial configuration resulted in a slightly bent DNA, somewhat closer to what observed in the experiments, along with an increased average distance between the 3′ and 5′ ends. In Figure 14b,c above, the blue histograms show the bending angle spanning between θ_b=130-170°, while the gap is more open, at between d_g=2-2.5 nm. In addition, Figure 15b displays the plot of distances between reference atoms methyl-C of Leu-151 vs. C2 of G45 (black trace); and methyl-C of Met-43 vs. C4′ of G1 (red trace). The MD simulation reveals that the movement of Zn2 adjusts over time, fluctuating between 0.5-1 nm by the end of the first 130 ns; however, Zn1 moves away from the 5′ strand by the end of 100 ns, and fluctuates between 2-3 nm (red trace). The significant structural transition of Zn2 and Zn1 could be due to

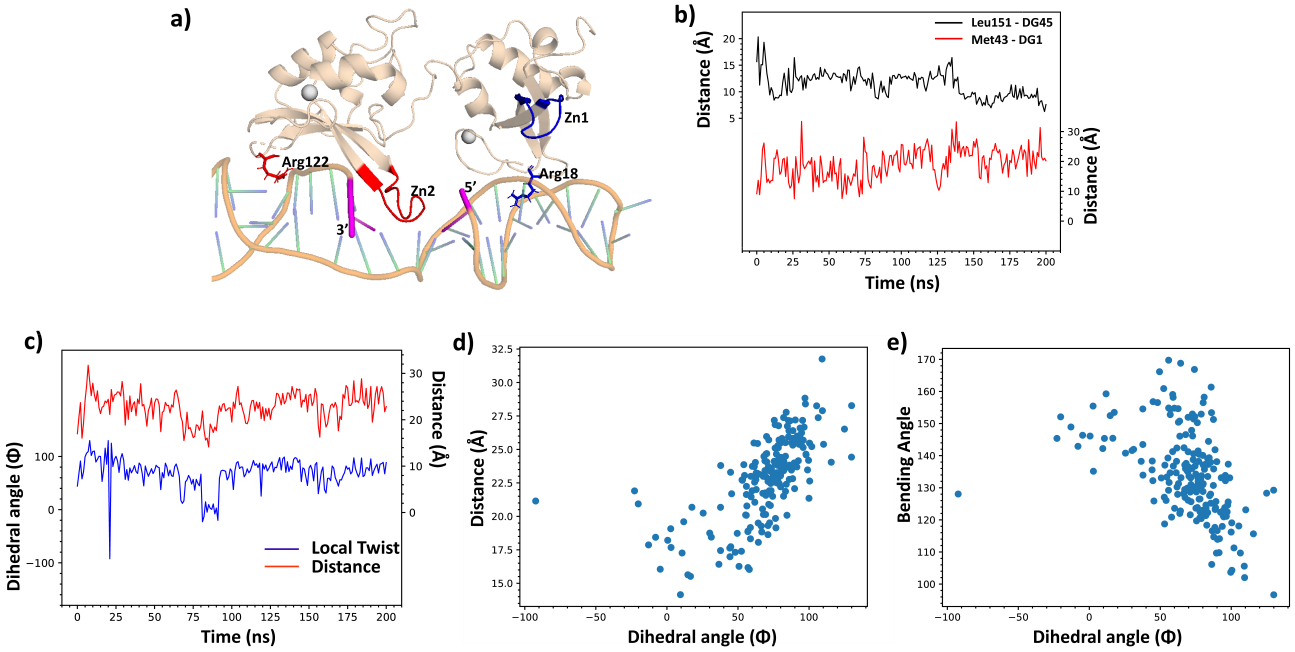


Figure 15: **(a)** Best cluster of the manually-adjusted structure after 0.5 µs MD simulation. **(b)** Plot of the C-C reference distances as a function of MD time, for the Leu151-G45 contact at 3′ (black trace) and the Met43- G1 contact at 5′ (red trace). **(c)** Plot comparing the local twist and the distance between the 3′-5′ terminals. **(d)** Scatter plot comparing the evolution of distance between the 3′-5′ terminals w.r.t local twist. **(e)** Scatter plot comparing the evolution of bending angle w/r to local twist

the presence of Arg-122 and Arg-18, which clamp to the major/minor groove of DNA (see again Figure 15a), bringing about no more that the slight bending in the DNA.

The degree of SSB opening may be correlated to the larger or smaller bending angle of the DNA hairpin. However, the distance d_g could increase also because of the local under/over-twisting of the major groove with respect to the DNA main axis, or because of a combination of bend and twist. To understand this point, we measured the local twist, defined as the dihedral angle Φ formed by the P atoms of the four nucleotides (1-22-24-45) that, covering two adjacent bp each carrying ~30-34°, should have a reference value of ~60-70°. Smaller values of Φ correspond to local under-twisting (Figure 15c), and a more closed SSB gap. It is observed a roughly linear correlation between Φ and d_g, the increase in the former (overtwisting) resulting in a larger gap (Figure 15d). On the other hand, the correlation plot of Φ vs. θ_b (Figure 15e) indicates as well a linear relationship, where lower bending angles θ_b (i.e., a more bent hairpin) are coupled with overtwisting and gap widening. During the whole simulation the bending angle spans the interval between θ_b=130-150°, but certain instances witness a very sharp decrease to θ_b=100-120° (see Figure A4). Such an occasional, sharp bending is clearly due to the clamping of the two

PARP-1 loops centered at Arg18 and Arg122, which fit in the groves on either side of the SSB, along with the mechanical constraints of DNA due to its peculiar hairpin configuration. Overall, these results show that large openings can be the result of either bending or overtwisting, or a combination of both effects.

2.4 COMPARISON BETWEEN MONOMERIC AND DIMERIC PARP-1

To ascertain whether the dimerization of PARP-1, and the consequent *independent* action of Zn1 and Zn2 (in fact a long-standing question in the literature [123, 130, 139]) could possibly play a role in the sensing of the SSB, I finally performed a series of docking+MD simulations with the two Zn fingers disjointed. At first, the 13-aa stretch that links the two moieties in the monomer was simply removed from the structure. In this case, we observed that the simple removal of the 13-aa linker from the same configuration used in Section 2.3 (direct docking to the straight DNA hairpin) does not affect significantly the PARP-1-SSB contact. After a 200-ns MD simulation, the comparison of RMSD and RMSF between the two simulations (Figure A5, see data with- vs. without-linker) displays only minor differences such as the sharp spikes around atoms 375 and 1099 of the DNA, corresponding to nucleotides T11 and T35, in the loops at the opposite ends of the hairpin.

Then, I tested the independent docking of the two Zn2 and Zn1 fragments to the straight DNA hairpin, either simultaneously or sequentially. The rationale for using again the straight hairpin, was to ascertain whether the Zn fingers have the ability to spot a pristine SSB, at least in this interaction mode.

For the **simultaneous** interaction, the 'three-molecule' docking option of HADDOCK was used. Zn1 (active residues Met43, Phe44) and Zn2 (active residues Gln150, Leu151) were freely docked to active sites G1 and G45 of the hairpin, without further specification of the mutual interactions. In practice, however, no meaningful starting structures could be obtained, since all the best HADDOCK scores corresponded to configurations in which Zn2 gets in the SSB, filling the gap between the 5′ and 3′ ends, and Zn1 is placed somewhere close but way outside the SSB, in a weakly interacting position. This is likely another proof supporting the notion that the closed SSB has no enough space for both Zn fingers, and Zn2 is the highest-affinity species to start the interaction, while Zn1 remains peripheral.

Then, for the **sequential** interaction, I started with docking Zn2 alone against the initially straight DNA hairpin. Since the conformation of Zn1 and Zn2 are now independent, there are many more possibilities for the interaction. For the sake of argument, I focused the attention on two, nearly symmetrical orientations of Zn2 with respect to the SSB, namely: **orientation (1)** resembling the experimentally observed orientation of monomeric-Zn2, and **orientation (2)** with Zn2

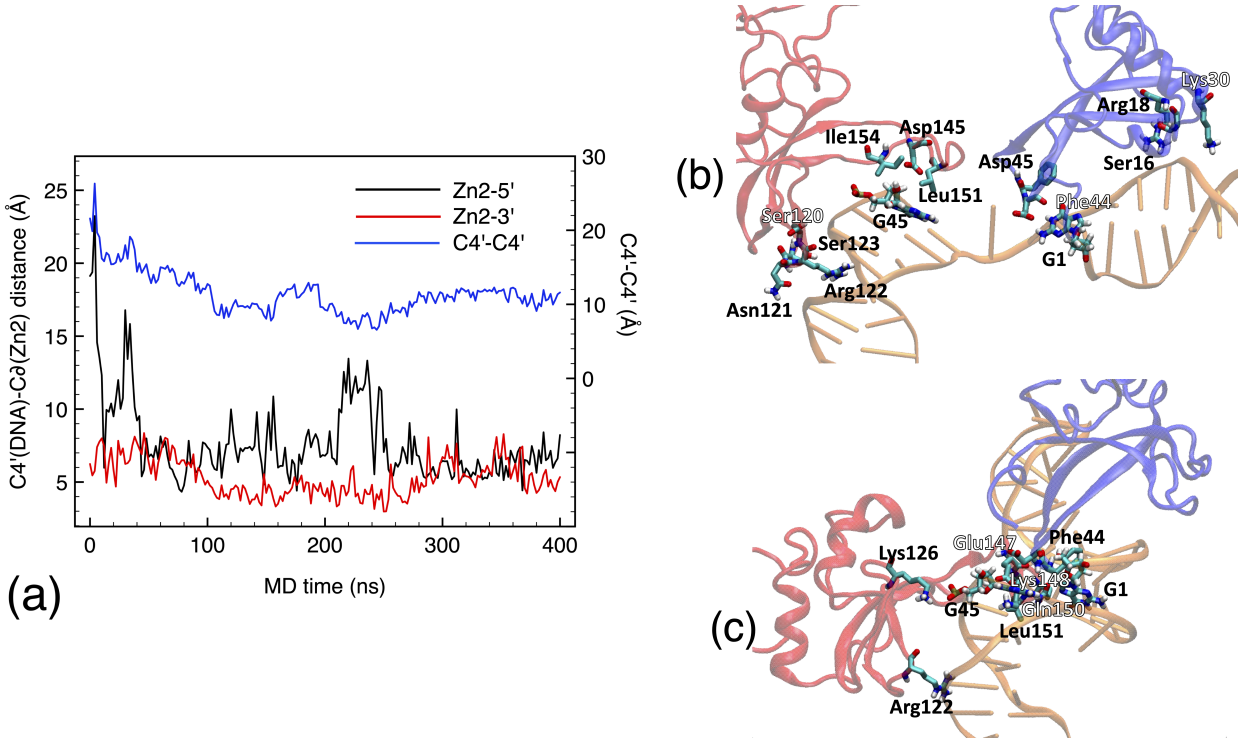


Figure 16: **(a)** Interaction of Zn2 with the pre-bent DNA hairpin starting from **orientation (1)**(see text). The black and red plots (left Y-axis) represent the time evolution of the distance between the Cδ of Leu151, and respectively, the two C4′ carbon atoms of the 3′ DNA terminal (black), or 5′ terminal (red); the blue plot (right Y-axis) represents the C4′-C4′ distance indicating the opening width of the SSB gap. **(b)** Initial configuration of Zn1 (blue), Zn2 (red) and DNA (orange), after independent docking and subsequent relaxation to the open DNA hairpin (at time t=2ns). **(c)** Same as (b), for the (ii) closed hairpin (at time t=400ns). These two structures (b,c) will be used in the analysis of Section 2.4.1 (see also Figure 17).

mirror-rotated by about 180° about the vertical direction (perpendicular to the main DNA axis). For each case (see the following subsections), firstly an equilibration of Zn2 will be studied, and secondly Zn1 will be added, based on the experimental affinity data from which it consistently appears that Zn2 is the PARP-1 moiety that establishes the strongest contact.

2.4.1 *Dimeric PARP-1: the "experimental" Zn2 orientation*

For the **orientation (1)**, the MD equilibration for 400 ns at T=310K shows that with this initial conformation, Zn2 has a considerable reactivity towards *both* the 3′ and the 5′ ends. As shown in Figure 16a, Zn2 is initially close to 3′; however, after about 300 ns of simulation it begins approaching also 5′, and for the remainder of the simulation it maintains a steady distance of about ~4 nm from either DNA ends. At the same time, the C4′-C4′ distance d_g (SSB gap, blue plot in the Figure 16a), starting from values as large as 2-2.5 nm, settles to a much

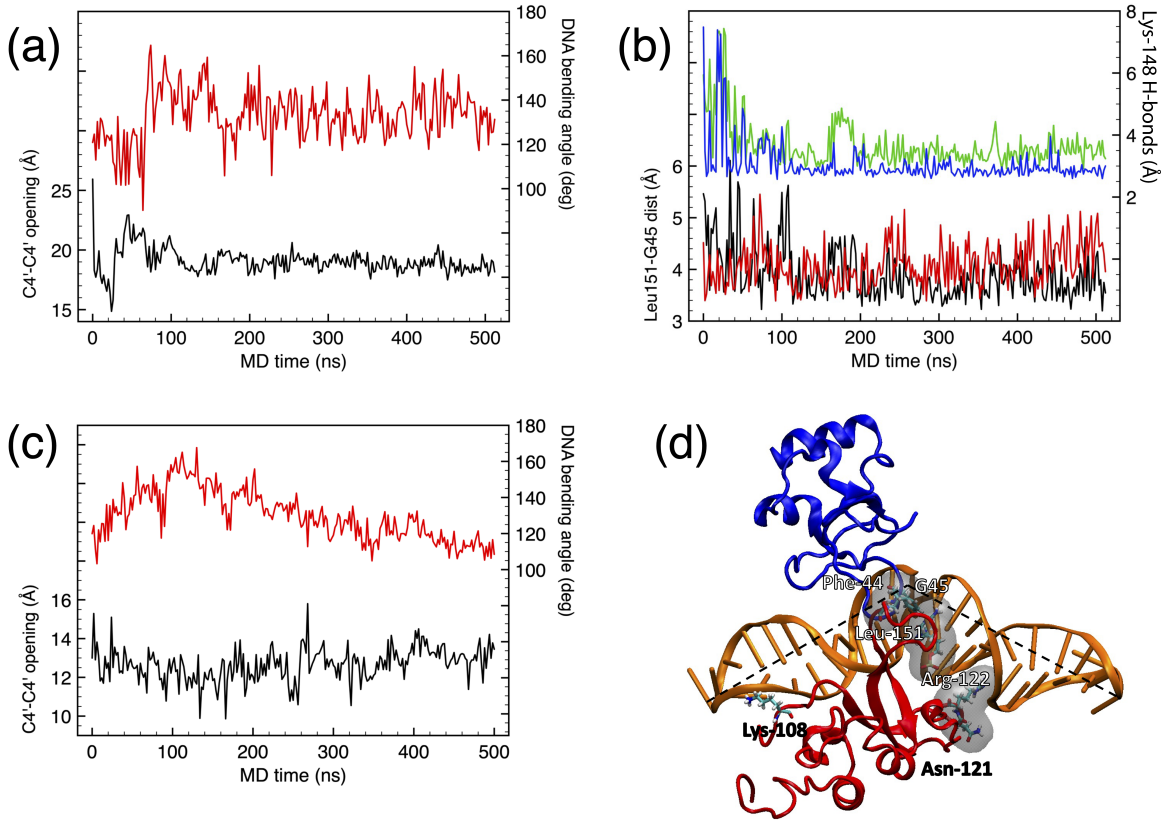


Figure 17: **(a) t=2ns-docking structure**. Evolution of the SSB gap opening (black plot, left Y-axis) and of the DNA bending angle (red plot, right Y-axis). **(b)** For the same structure, interaction of Zn2 with the 3′ DNA end: distance between the two Cδ of Leu-151 and the C6 (black) and C8 (red) of guanine-45; H-bonds between Lys-148 Nζ and thymine-23 O8 (blue) and N3 (green). **(c) t=400ns-docking structure**. Evolution of the SSB gap opening (black plot, left Y-axis) and of the DNA bending angle (red plot, right Y-axis). **(d)** For the same structure, best cluster for the (ii) conformation, represented as ribbons (DNA=orange, Zn1=blue, Zn2=red); the interaction regions are depicted as shaded surfaces; the dashed line represents the instantaneous DNA bending angle.

shorter d_g ~1 nm at about the same time, indicating the rapid closing of the gap, and with a conformation of the hairpin that turns more and more straightened, compared to the largely bent initial state.

Next, I performed several distinct HADDOCK docking experiments, in which Zn1 was approached to the 5′ end of this Zn2-DNA complex at different time frames of the MD simulation. In the following, two representative situations were selected, at times t=2 ns (Figure 16b) and t=400 ns (Figure 16c), respectively characterized by a very open, vs. a rather closed hairpin conformation. For both initial dockings, two subsequent 500-ns MD equilibration runs were performed. Detailed bonding and solvation free-energy information was deduced from the analysis of the best cluster for each simulation.

For the ***t*=2ns docking**, the SSB gap always remains quite open at about d_g=2 nm, as shown in Figure 17a (black plot), and the bending

angle remains constantly about θ_b=130° (red plot). The MD trajectory shows a moderate interaction of Zn2 with the 3′ DNA end. Leu-151 and Gly-152 still contribute about half of the total hydrophobic contact with G1. However, the total ΔG=-2.2 kcal/mol is in large part contributed by the residual interaction of Lys-148 with the middle-exposed DNA base T23, between which a H-bond at 0.195 nm is also formed (Figure 17b, blue plot). Zn1 now appears to interact more strongly with the 5′ end, compared to the starting configuration. Phe-44 still maintains a strong hydrophobic interaction, where the phenyl group is mostly facing the G45 heterocycles, and Lys-47 makes an 0.198 nm side-H-bond with the ribose O5′. The total (solvation) + (H-bond) ΔG is -6 kcal/mol. The direct Zn1-Zn2 interaction is maintained by a strong pair of salt bridges, formed by the Nζ of Lys-148 and the two carboxyl oxygens of Asp-45. This interaction is likely to distract Zn2 from a possibly closer interaction with the DNA, by holding it closer to the Zn1 loop, rather than to the 3′ end.

Concerning the ***t*=400ns docking**, the initial arrangement of Zn1 and Zn2 appears to remain little altered for the whole 500-ns MD; because of their mutual interaction, both zinc fingers seem to have a reduced interaction with the SSB gap. However, an interesting information obtained from the plots of this simulation in Figure 17c includes the DNA bending angle rapidly evolving towards quite smaller angles, about 120° in the second part of the trajectory, while at the same time the SSB gap remains quite closed at 12-13 nm. Given the little interaction of both Zn1 and Zn2 with the SSB gap, and the contact of Asn-121 and Arg-122 apparent in Figure 17d (already in slight contact also at the beginning of the simulation, see Figure 16c), such a bending appears to be the result of spontaneous thermal fluctuation. In fact, only at much later times a third contact (Lys-108, see figure) will arise on the opposite side of the DNA hairpin, thus stabilizing the opening angle.

2.4.2 *Dimeric PARP-1: the 180°-flipped Zn2 orientation*

For the **orientation (2)** of Zn2 with respect to the straight hairpin, the best docking pose resulted in a cluster with an RMSD 0.128 nm, all the poses of the cluster having the base loop of a.a. 149-152 of Zn2 very close to the 3′ DNA end. After 400-ns MD equilibration of the top-scoring pose, a conformation with a large opening of the SSB gap was obtained, with Zn2 reaching its stable position along the 3′ end. While this conformation approaches an opening angle similar to that observed in the experiments, it should be equally possible that the kink is due to a spontaneous fluctuation of the DNA hairpin, and/or by direct action of the Zn2 finger penetrating in the SSB gap. By comparing the θ_b bending angle of the phosphate backbone, and

the SSB opening d_g, it is observed that the DNA gradually unfolds with θ_b fluctuating between 120-130° and d_g=20-25 nm.

Next, this average structure obtained as the best cluster from the MD simulation was used as the starting Zn2-DNA complex, to further dock the Zn1 monomer. This combined Zn1-Zn2-DNA configuration is therefore our best candidate for further equilibration at finite temperature and pressure. After 1 μs of MD simulation, the trajectory is clustered and representative configurations are extracted.

It is observed that Zn2 penetrates in the SSB gap, with a total ΔG=-3.9 kcal/mol. The main interactions included in this ΔG are displayed in Figure A7: (a) direct interaction with the G45 at 3′ end (Figure A7a);(b) hydrophobic interaction of the 119-122 a.a. loop, (c) Arg-122 making two H-bonds with the O2 oxygen of C42 (0.207 nm) and (d) with O4′ ribose oxygen of C23 (0.249 nm) at the minor groove (Figure A7b); (e) Gln-150 making contact with C22 and T23 at the opposite side of the SSB (Figure A7b).

Concerning Zn1, Figure 18 shows a summary of the results of MD simulation, focusing on the special binding configuration of Zn1 at the 5′ end (Zn2 removed from the view for the sake of clarity). Zn1 strongly interacts with G1, by making a strong H-bond of Pro-49 with the ribose O5′ and a weaker one with C4′ (Figure 18b), plus a very peculiar contact of Phe-44 (see below). At a later time it also establishes a strong contact with the DNA groove, the Arg-18 making two H-bonds with the O6 oxygen of G4 and G6. Its overall ΔG at this later stage is close to -8 kcal/mol, largely coming from the 43-45 a.a. loop and the 5′ terminal.

The interaction of Phe-44 with the aromatic cycles of guanine G1 is of a special "edge-to-face", or T-type, which was shown to be one characteristic feature of protein-DNA interactions [140]. In such π-bonding configuration, the phenyl of Phe-44 makes an angle with the plane of the G heterocycles typically between 60-70°, rather than coupling in a side-by-side stacked arrangement. Phenylalanine is found to be the most likely candidate for such T-type bonding with DNA, twice more frequently with thymine compared to purines [140]. Such a bonding geometry for our case is shown in the central panel of Figure 18, where the π-π angle is also defined. The evolution of the Phe44-G1 angle is shown in Figure 18a, which fluctuates between 60 and 80°, a figure in good agreement with the literature findings [140].

By looking at Figure 18c,d there seem to be no apparent correlation between the fluctuations of the T-bonding of Zn1, and either the DNA bending angle (Figure 18c), or the SSB gap opening (Figure 18d). The former evolves from a rather bent shape in the first ~500ns, with θ_b=90-120°, to a more closed shape in the second half of the MD trajectory, with θ_b=140-160°. However, the relative arrangement of Zn1-Zn2 is such that the SSB gap (Figure 18d) remains always very

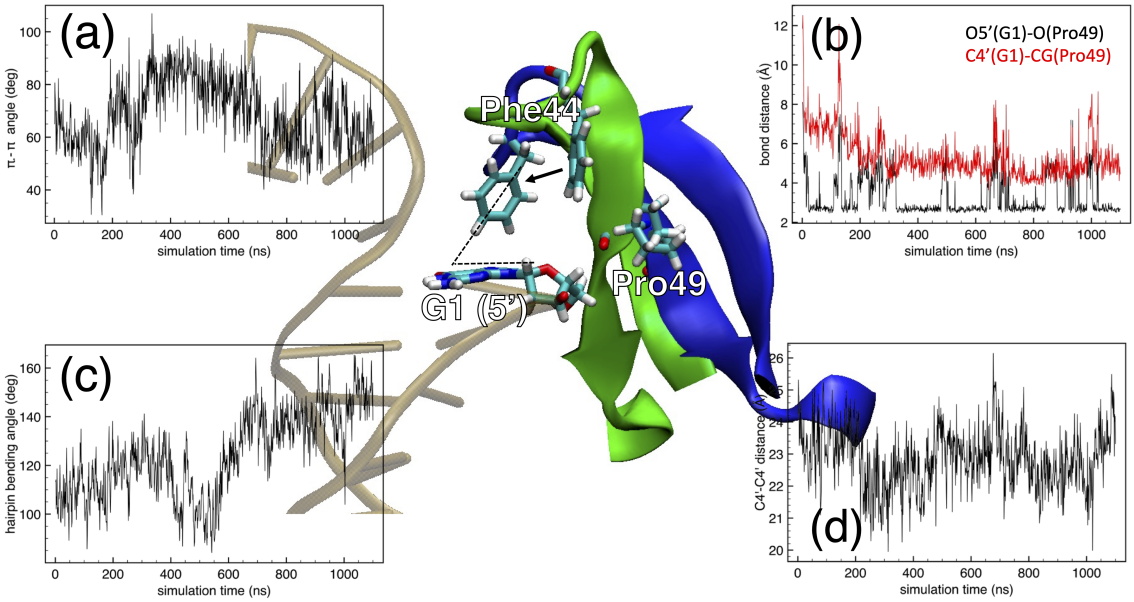


Figure 18: Interaction of the Zn1 with the DNA 5′ terminal after 1.2 µs (blue ribbon), starting from the manually docked configuration (green ribbon). Only the contact loop is depicted. Part of the DNA sketched as a brown ladder in the background. **(a)** Time plot of the $\pi-\pi$ angle between the planes containing the aromatic cycles of Phe44 and G1 (sketched by the dashed angle in the central panel); the hexagonal ring of Phe44 moves from the initial to the final position as indicated by the little black arrow. **(b)** H-bonding distances between G1 and Pro49. **(c)** DNA hairpin bending angle, measured between the P atoms of nucleotides 12-24-34. **(d)** Opening of the 5′-3′ gap in the DNA hairpin, measured by the distance between the respective C4′ atoms.

open, at d_g=2.2-2.4 nm, independently on the straightening of the hairpin.

Clearly, such results for **orientation (2)** are rather interesting when compared to those obtained with the **orientation (1)**, albeit in the (2) the Zn2-Zn1 orientation is someway "opposite" to the (1) (in turn, closer to experimental observations). They seem to suggest that, if a dimer made of two copies of PARP-1 independently interact with the SSB (a possibility not yet fully explored in the experiments), a wider range of interaction possibilities arises. Of course, some of these configurations would be excluded in the realistic situation with two full copies of PARP-1 interacting, instead of using just the Zn fingers, since the whole proteins could lead to steric clashes of the dimer. Overall, the ensemble of simulations of this Section 2.4 point out that the role of Zn2 in pre-determining the opening of DNA is relevant, but might not be as crucial as early suggested. It is rather the particular structure and sequence of the DNA hairpin, which possibly allows the SSB gap to open up spontaneously by frequent, free thermal fluctuations. Only when the fluctuation spontaneously opens up

the hairpin, Zn2 could make its interaction, stabilize the DNA kink, and subsequently make room for Zn1.

2.5 SUMMARY AND OUTLOOK

To summarize, in the present Chapter 2 I investigated by all-atom molecular dynamics simulations the competition-cooperation of two Zn-finger domains of PARP-1, in recognizing an SSB in a model fragment of DNA. Since one main purpose was to make a link with recent structural biology experiments [128, 132], I used the exact same 45-nucleotide ssDNA hairpin designed with a sequence such that the structure makes two near-symmetric loops, and closes back with a missing nucleotide between the 5′ and 3′ terminations (the guanine-1 and guanine-45, facing the central, unpaired nucleotide thymine-23). Both in the published experiments and in these MD simulations, the gap between the two terminal nucleotides were taken as representative of a fully-formed SSB with proper 5′-dRP and 3′-OH terminations, and also a configuration with both OH-terminated ends for the sake of comparison. Such situations can be representative of an enzymatically-cut SSB or the final stage of processing radiation-induced SSBs, just prior to the start of the repair phase. The PARP-1 protein was represented in full atomic resolution only for the regions actively used in the co-crystallization experiments, specifically, the two Zn fingers Zn1 and Zn2. In some of the MD simulations we kept them physically connected by their 13-a.a. linker, while in other simulations we made them act independently, to simulate a PARP-1 dimer. In this way, widely different configurations could be explored, ranging from the structure with the DNA hairpin straight and the gap "closed", to the experimentally identified configurations in which the hairpin is strongly bent and the SSB gap is wide open. The main findings of this Chapter can be summarized as follows:

1. The highly peculiar hairpin conformation of the DNA structure allows for a large conformational variability under thermal fluctuations, compared to straight, long DNA harboring a similar SSB at its center. While the straight DNA can experience broad curvatures, especially when its length approaches and surpasses the persistence length ($\gtrsim$50 nm), the atomic structure surrounding the SSB remains relatively unaffected. It is therefore possible that, also in the experiments, the DNA hairpin could spontaneously open up to extreme deformations (such as, bending angles much smaller than 120-130°, SSB gap openings of 2.5-3 nm or more), thanks to its special conformation, and even without the initial help from the PARP-1 contacts.

2. The experimental structure of the interacting complex [128, 132] can be readily retrieved in the MD simulations, starting from

the separate fragments initially taken in the respective experimental conformations. This means that the experimental structure represents indeed a global thermodynamic minimum of the DNA-Zn1-Zn2 complex also for the molecular simulation, thereby providing a strong support to the reliability of the simulation protocol, and of the subsequent results.

3. The interaction of PARP-1 with the DNA hairpin in the straight conformation is not able to display the direct transition to the "final" state, represented by the experimental structure. It was observed that Zn2 is indeed able to force itself inside the SSB gap and make contact with the 3′ end, attaining values of bending angle and gap opening not far from the experimental. However, the position of Zn1 in these instances is not correct and lies far from the opposite 5′ DNA end, unlike what is seen in the experiments. A series of MD simulations in which both Zn1 and Zn2 were forcefully placed in a better initial position with respect to the straight SSB, demonstrated a steady contact between Zn1-5′ and Zn2-3′. However, in this case the SSB gap does not open and the hairpin remains practically straight (i.e., the DNA-protein contact remains "external" to the SSB).

4. Resistance to large bending is apparently due to the extra contacts that both Zn fingers establish, by pushing their side loops (the a.a. 17-19 for Zn1, and the a.a 120-122 for Zn2) deep in the DNA grooves flanking the SSB gap, which forbid further bending of the hairpin. It may anyway be speculated that, over a much longer time scale than the μs accessible to MD simulations, such 'side contacts' could eventually lead to a forced bending. Overall, a good correlation existst between the SSB gap opening with the DNA bending angle, as well as with the degree of local over-twisting of the double helix. From a kinetics point of view, it appears that the DNA bending is mostly due to spontaneous, free thermal fluctuations until, at some later moment, the side protein loops contact a DNA groove, and "lock-in" the conformation, be it a straight or a bent hairpin. On the other hand, the SSB gap opening is initially correlated with the DNA bending and twisting, but it is ultimately determined by the arrangement of the Zn fingers, primarily Zn2 which has the tendency to penetrate deep in the gap, and interacts with the nucleotides opposite to the SSB.

5. The question whether PARP-1 operates as a monomer, using the Zn1 and Zn2 from the same molecule for detecting the SSB, or as a dimer, cooperatively using Zn1 and Zn2 from two different molecules, remains an open one. I performed computational experiments by using opposite arrangements of the two Zn fingers, either acting in parallel or sequentially. It is found that Zn2

displays the highest affinity for the SSB open ends; it attacks preferentially the 3′ end, coherently with the experimental findings, but it can also attack both ends, especially if starting from a relatively more closed SSB gap. The role of Zn1 is therefore secondary, as shown in the experiments (which however used only monomeric PARP-1). Also in the dimeric form, Zn1 must fight its way to the closed SSB, and it is only when the SSB is already substantially open, that it may join Zn2 in recognizing the SSB.

By putting together the findings of Section 3.2.3 about PARP-1 interaction with the straight and closed SSB, next to the observations of Section 3.2.1 about the easier flexibility of the DNA hairpin interacting with the Zn2 alone, and the observations of Section 3.1, about the large opening and bending fluctuations observed for the isolated hairpin in the absence of PARP-1, an interesting interpretation of the dynamics of recognition of the SSB by PARP-1 can be reconstructed. A possible picture that emerges from the present study – in agreement with the *final* structures found in experiments – can be that the Zn2 first interacts with the 3′ end of the SSB, thereby helping the spontaneous DNA bending; after which, Zn1 can start interacting with the 5′ end, leading to the final observed kinking; the fact that the direct docking of the Zn1-Zn2 couple on the pristine closed SSB gives a good interaction, as far as the adhesion free energy, but with little or no opening of the SSB, because of the extra contacts along the DNA sides, adds support to the need for spontaneous thermal fluctuations of the DNA hairpin. The main objection to such a picture is, then: why the experiments can observe Zn1-Zn2 interacting with a wide-open SSB, and a largely bent DNA hairpin?

The answer can be found in the free fluctuations of the isolated hairpin *prior* to interaction: it is possible that a spontaneously fluctuating DNA, offering every now and then the SSB in a wide-open conformation, has a larger affinity for Zn1-Zn2 compared to Zn2 alone (the so-called 'bend-then-bind' action), and this is what has been interpreted as Zn1-Zn2 forcefully opening the SSB ('bind-then-bend'). Notably, from our preliminary comparison of the interaction of PARP-1 with a free DNA strand, it appears that this mode of interaction is not entirely specific to the peculiar construction of the hairpin. While in the hairpin the terminal loops greatly facilitate ample structural oscillations, also the free DNA shows enough bending upon contacting the protein, in this case associated with its long-wavelength fluctuations. An interesting way to prove experimentally this idea could be to try to use a hairpin construct but with much longer flanking sequences, so as to keep the closing loops as far as possible, left and right of the SSB gap, thus approaching the free DNA condition.

In Chapter 4 I will study the mechanical behavior of SSB and DSB damage in the linker DNA. Unfortunately, the time constraints did

not allow to complete the study by also including the DNA-PARP-1 interaction in the context of linker DNA joining pairs of nucleosomes. One can speculate that the free fluctuations of the DNA observed here, in that case will be severely constrained by the presence of massive nucleosomes at both ends; as it will be seen, the relative displacements of the nucleosomes induce further mechanical deformations by coupling bending, twisting and kinking modes [141]. Moreover, it should be considered that, at least in radiation-induced SSBs, the gap terminations are far from clean $5'/3'$ ends at the stage of initial defect detection. Therefore, the ability of PARP-1 to recognize a "dirty" terminated SSB remains another key issue to be elucidated.

3

INTERACTION OF PARP1 WITH SINGLE-STRAND BREAKS IN THE NUCLEOSOME

3.1 INTRODUCTION

In the previous Chapter 2, I studied the interaction of PARP-1 Zn-finger domains with SSB in a DNA shaped as a short hairpin loop. It was seen that the hairpin-DNA can undergo conformational changes when the SSB interacts with the Zn-finger domain of PARP-1. This could be either due to one loop of the Zn-finger contacting the backbone of DNA, thereby creating hydrogen bonds; or due to the highly special structure of the hairpin, in which a few unpaired nucleotides are flanking near the loop (see Figure 12d). Anyway, depending on the ordering of the elementary steps of the process, such an interaction could represent either a "bind–then–bend" (i.e., the enzyme actively deforms the defect site after finding it), or rather a "bend–then–bind" (the defect is spontaneously modified prior to binding) sequence [142–144]. In our laboratories, we are actively pursuing the idea that spontaneous mechanical deformation and stress-field concentration at defect locations could be a key to exalt the visibility of the DNA damage, someway "attracting" the repair enzymes in a more efficient way.

This issue becomes further complicated when DNA is wrapped around the nucleosome, because of the steric hindrance and extra charges provided by the histone proteins, and notably histone tails. Numerous studies showed that upon interacting with free DNA fragments, PARP-1 displays a strong affinity for 5′-dRP end near a SSB, or with the blunt ends near the duplex that, in a way, are taken to mimic a dissociated, blunt DSB [128, 132, 145]. The few experimental studies of the interaction of PARP-1 with isolated mononucleosome particles have repeatedly shown PARP-1 bound to the blunt ends of DNA that are found at the entry/exit of the nucleosome [146–148]. Recent studies also demonstrated that two or more molecules of PARP-1 can bind to both nucleosome free ends, again mimicking DSB damage [146].

The binding of PARP-1 to the nucleosome brings about allosteric changes not just in the enzyme itself, but it also may partially unfold the nucleosomal DNA, e.g. by reducing the mobility at certain regions [147]. This could also be possibly due to destabilization of the intra-nucleosomal interactions, between histones & DNA or among the core histones [147]. Thus, it is surmised that the interaction of PARP-1 with nucleosomes could drive some form of reorganization of the nucleosome itself, and the magnitude of local changes depends

upon the strength (and number) of PARP-1 attachment(s). It is indeed observed that multiple molecules of PARP-1 bring about extensive rearrangement and unfolding of the nucleosome particle, while a single copy of PARP can only bring about local changes at the binding site [146]. In certain cases, PARP-1 binding kinetics is accelerated by the presence of a H2A.X histone variant, and may therefore contribute to PARP-1 accumulation, chromatin reorganization, and PAR synthesis [149]. Moreover, the study by Luger et al. [148] showed a large reduction of the binding affinity for nucleosomes lacking the linker-DNA fragment (i.e., the bare nucleosome core particle, NCP).

There are no theoretical or experimental studies of PARP-1 interacting with single-strand breaks at nucleosomes.

Given the above experimental landscape, it is however worth reminding that there are no studies that focused on the molecular details of PARP-1 interacting at the site of a SSB located on the body of nucleosomal DNA. On the other hand, computer simulations have already provided interesting insights about the protein-DNA interactions in the context of damage and repair of different types of defects [150]. Such a situation provides a critical motivation for our investigation by molecular simulations the interaction of PARP-1 with DNA damage directly in the nucleosome body, and in particular the role of SSB, which is traditionally much less studied than the more dangerous DSB. Hence, in this Chapter 3 of my thesis I addressed the molecular dynamics of interaction between the SSB localized at different positions on nucleosomal DNA, and the Zn-finger domains of PARP-1 already used in the previous Chapter. Further, I designed multiple SSB in the backbone of the nucleosomal DNA, while ensuring that these SSB can interact also with the histone tails, with the main aim of answering to the following questions:

1. How PARP-1 interacts with SSB at different positions in the nucleosome?
2. Can the presence of PARP-1 other than at entry/exit points of the nucleosome induce nucleosomal reorganization?

Figure 19b depicts the initial configuration of the nucleosome containing three SSBs at distinct positions: **site1** is within the nucleosome, at superhelical position-5 (SHL-5); **site2** is at the dyad, position SHL-0; and **site3** is just near the entry/exit of the nucleosome, SHL-7. All of the SSBs are far apart from each other, and lie in same DNA strand, so that they cannot evolve into a kind of DSB, even if at a long distance. For each SSB site, the free DNA termini are terminated by 3′-OH and 5′-dRP ends. This is because PARP-1 display a greater binding affinity to the "dirty" ends created by the breaks [128], therefore this could help in establishing a comparison by analogy with the experimental and simulation results on the free DNA in Chapter 2 [120, 128]. I used again the HADDOCK software to dock the PARP-1 fragment including the two Zn-finger domains, to the SSB; the base-stacking loop in Zn2 was always designated as an active residue for

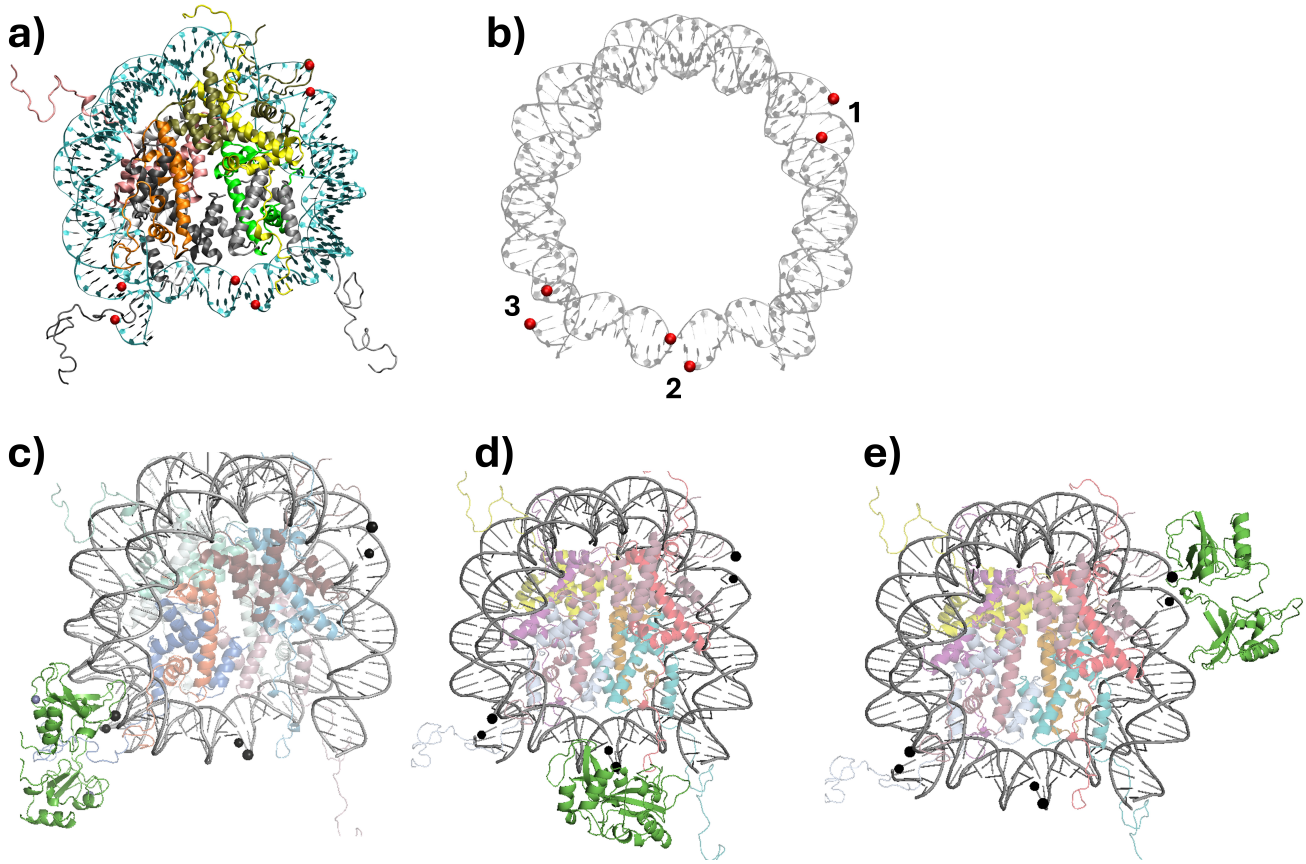


Figure 19: **(a)** Top view of the **1KX5** nucleosome; histone proteins represented as color ribbons; DNA in cyan cartoon representation. **(b)** Nucleosomal DNA with the positions of the SSB numbered 1 to 3; the free SSB ends are represented by colored spheres. Note that site 3 is at the end of the DNA double-loop, next to the lower entry point of the NCP. **(c-e)** The initial configurations of the three SSB positions around the 1KX5 nucleosome with the docked PARP-1 (Zn1-Zn2 represented in green ribbons); (c,d,e) respectively depicting SSB site 3, 2 (dyad) and 1.

interaction with the 3′ end of the SSB. After the docking phase, to understand how the interaction would evolve between the Zn1-Zn2 and the SSB, I performed MD simulations according to the protocol described in Appendix B, Section B.1. The system nucleosome+PARP-1 is solvated in a water box of size 25x25x25 nm^3, with about 497,000 water molecules and a total of about 1,530,000 atoms.

3.2 Stability of SSB at different positions in the nucleosome

As an initial check prior to introducing the PARP-1 interaction, I firstly tested the thermal and mechanical stability of the model nucleosome including the three SSBs. In particular, I wanted verify whether there could be observed a spontaneous "opening" of the SSB free ends just by thermal fluctuations, and any spontaneous mechanical defor-

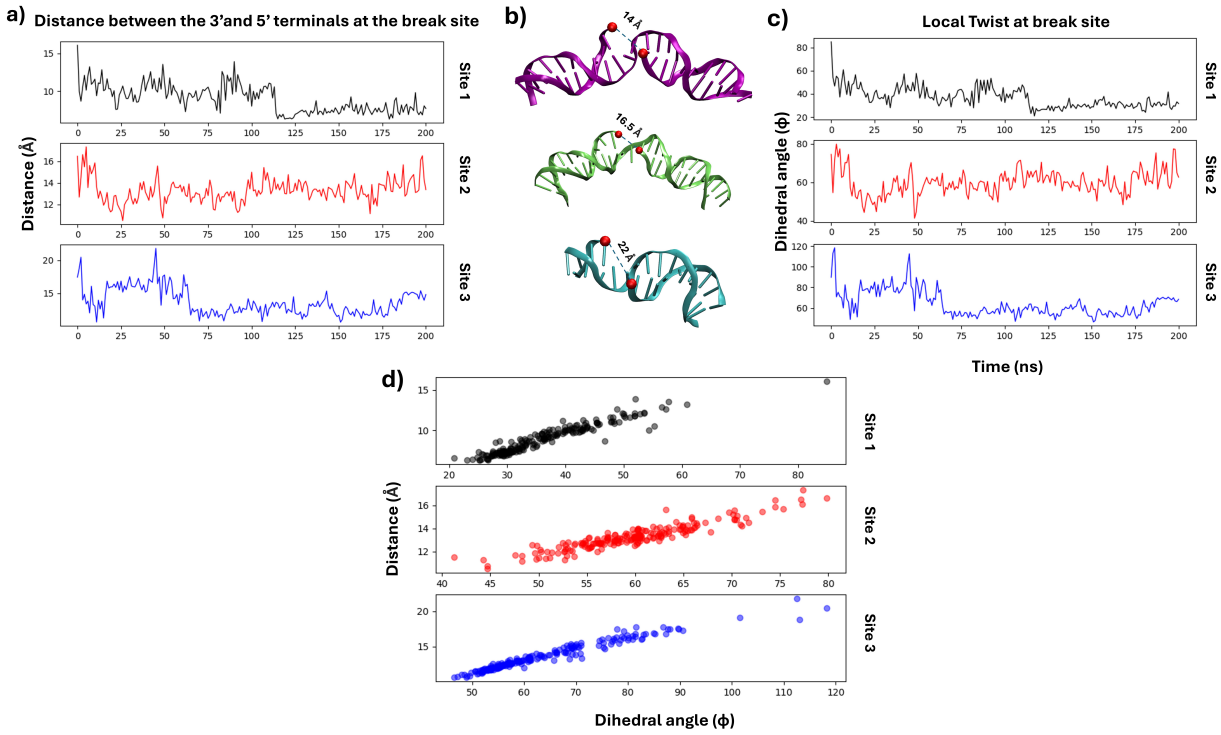


Figure 20: **(a)** Gap opening distance between the 3′ and 5′ ends of the each SSB break site in the free nucleosome (without any PARP-1). **(b)** The corresponding snapshots of each SSB at the maximum opening. **(c)** Evolution of local twist Φ at each SSB site over 200 ns of molecular dynamics. d) Correlation of the gap opening d_g with the dihedral angle Φ.

mation near the SSB site, without the help of PARP-1 [120, 128]. For this purpose, a first series of 200-ns MD simulations were run for the nucleosome system with the three SSB and no PARP-1.

As usual, the SSB gap distance d_g is calculated between the instantaneous positions of the C4′ atoms in the 5′ and 3′ ends. After 200-ns of MD simulation, it is observed that the SSB opening between the free ends fluctuates between d_g=15 to 20 Å, as shown in the plot in Figure 20a. Upon comparing individual SSB sites, it is observed that **site1** and **site2** display very minimal fluctuations in the d_g value, with **site1** closing back to d_g ~8-10 Å in the second half of the trajectory. On the other hand, **site3** (close to the entry point of the NCP) does open up to as much as 22 Å, due to thermal fluctuations. For comparison, in Chapter 2 the SSB in the hairpin-shaped DNA – but not in the long free-DNA – was observed to open up spontaneously due to thermal fluctuations, even in the absence of PARP-1.

The next question was the attempt to understand whether mechanical deformation in the SSB site affects the opening and closing of the gap. For this purpose, I calculated the local twist at the break site, defined as the value of the dihedral angle Φ between the C3′ atoms of the two base-pairs flanking either side of the unpaired nucleotide. Given that the DNA helix is twisted by about 34° at each base-pair,

the normal value of Φ should be about 60-70°. Figure 20c,d respectively show the evolution of the local twist Φ over 200 ns, and the correlation between the gap opening distance d_g and the angle Φ. The last plot (d) demonstrates that the local twist and the gap distance are linearly related, where an increase in the SSB gap opening is associated with a large local over-twisting of the double helix. For **site1**, the closing of d_g to ~10 Å corresponds to an under-twisting down to Φ ~35-40°. The dyad **site2** displays a nearly unperturbed configuration, in both the d_g and Φ values. Interestingly, the larger values of d_g and Φ for **site3** compared to the other two sites should be due to its position at the entry of the nucleosome (see Figure 19b). This particular SSB position leads to the formation of a short stretch of 5-nucleotide dsDNA at the entry of nucleosome, right before the SSB, which is a highly unstable configuration quite prone to opening and deformation under thermal fluctuations (see Appendix 1 Figure A8), even during the initial few ns of the MD trajectory. Also, it is to be noted that the nucleosome model derived from the 1KX5 structure includes full length histone tails, which also play a major role in the mechanical deformation at the break site [141, 151]. This aspect will be explored in more detail in a next subsection.

3.3 Analyzing the Zn-finger domain contacts with SSB

Experimentally, SSB identification by PARP-1 Zn-finger domains is predominantly characterized by a phosphate–backbone loop and base stacking-loop of the protein fragment, interacting with the break site [128]. This latter smFRET experimental study showed that the Zn2 domain has a higher affinity to the damage site, compared to Zn1 [132]. While Zn2 identifies the break, and brings about conformational changes in the DNA, Zn1 would further add more interaction with the break, which eventually helps the allosteric changes in the whole PARP-1 molecule. The major interaction occurs between Leu-151 of Zn2 and the 3′ DNA end, which induces PARP-1 to undergo some allosteric shift while strengthening its binding to the SSB. Zn1 then enhances these interactions, by adding further contact points at the DNA 5′ end mainly via its Phe-44 residue, thus stabilizing the binding, to start PAR–chain synthesis and chromatin restructuring [128].

In this part of the study, I attempted to replicate the specific contacts mentioned above, between PARP-1 zinc-finger domains and the SSB sites in nucleosomal DNA, similar to the protocol and observations already seen in Chapter 2 for the DNA hairpin structure. The initial structure was designed by docking three individual copies of the Zn1-Zn2-finger domains (PDB ID: **2N8A**) to each of the three SSB sites around the nucleosome, with the protocol discussed in the Appendix, Section B.1 using HADDOCK and Chimera public-domain

softwares. I merged both these methods together, due to the difficulty in achieving a good quality of the contacts between the Zinc-finger domains and SSB, comparable to the observations of previous Chapter 2. In retrospect, this may be due to the limitations of HADDOCK in dealing with too complex and too large molecular assemblies. Also, it was seen in Chapter 2 that the "right" contacts are necessary for bringing about the subsequent opening and bending of the SSB. In fact, the docked clusters obtained from HADDOCK had some of the interactions observed in the experimental NMR structure for the DNA hairpin; from these clusters, the best docking pose were selected based on their RMSD values, as the initial configuration; then, Chimera was used to further improve the adjustment of the interacting moieties, and get the final structures as shown in figure 19c-e. These structures were then used as initial configurations for 200-ns MD simulations at 310 K (Figure 14a).

The resulting MD trajectory was first analyzed with the GROMACS clustering algorithm, with a RMSD cutoff of 0.45 nm and GROMOS method. The most representative clusters for each SSB interaction site are shown in Figure 21, the Zn1 and Zn2 domains depicted in blue and red ribbons, respectively. Clearly, at each of the SSB **site1, 2, 3** it is evident that the Leu-151 of Zn2 is very close to the 3′ terminus, and Phe-44 is away from the 5′ terminus. At each site, a number of hydrogen bonds and Van-der-Waals interactions between Zn12 and DNA backbone are observed. Similar to what was observed in Chapter 2, such interactions should be responsible for stimulating gap opening, thus helping Zn2 to establish the pristine interaction with the SSB site.

In the initial analysis, I looked at the SSB gap–opening distance d_g between the 3′ and 5′ termini flanking each SSB site, to determine whether Zn2 could stably interact with these regions. During the 200-ns MD simulation, the distance d_g for all the three SSB sites fluctuates between $12 \leqslant d_g \leqslant 16$ Å, close to their initial separation of 12 Å. However, all SSBs repeatedly open the gap to larger values during the MD trajectory, up to d_g=20 Å, which was not observed during the 200–ns MD simulation of free SSBs without PARP-1 (see Figure 20a), for which the d_g remained at 14-16Å. Especially for **site3** the gap has large fluctuations to $d_g > 18$ Å throughout the MD simulation, in the presence of Zn1-Zn2. Altogether, it can be concluded that the presence of PARP-1 brings about a substantial change in the opening distance at each SSB, to values rarely observed in the absence of Zn1-Zn2. This might seem an indication in favor of the active modification of the damage site by the enzyme ("bind–then–bend"). However, note that the opening d_g does not seem yet sufficient to accomodate also full interaction with Zn1.

The next step was to understand if and how the interaction with Zn1-Zn2 can influence these changes, and to check whether the "trade-

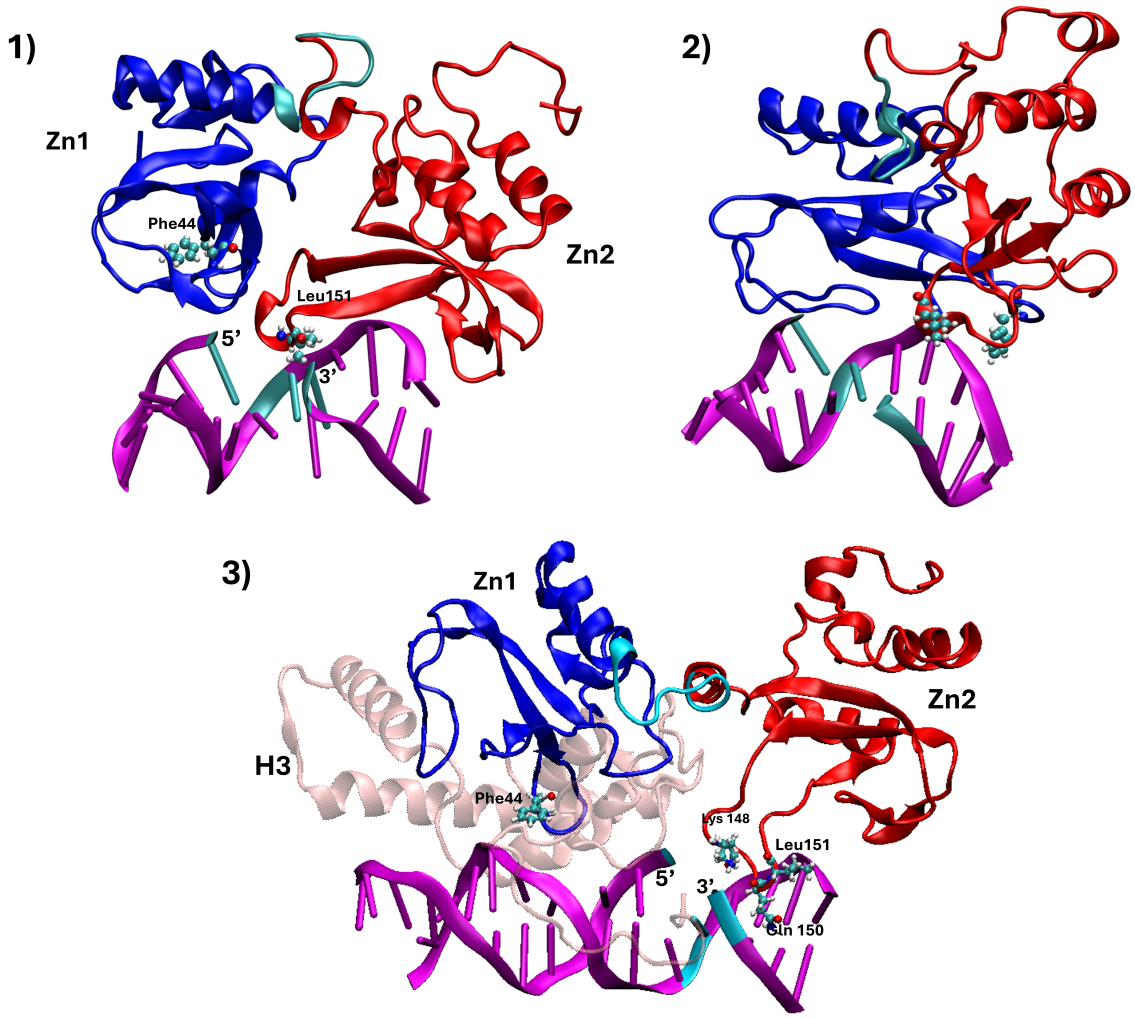


Figure 21: Configurations depicting the interaction of the Zn1 (blue ribbons) and Zn2 (red ribbons) with the DNA fragment including the SSB (purple), for the three sites. The best clusters are obtained from the 200–ns MD trajectory, by setting a RMSD cutoff of 0.45nm. For SSB **site3** also the interaction with histone H3 (depicted in pink ribbons) is displayed.

mark" interactions are still identified in the nucleosomal DNA, within such a limited time frame of 200 ns. As discussed in Chapter 2, the affinity of Zn2 to the 3′ terminus of the SSB site is notably higher than that of Zn1 towards the 5′ terminus. To validate this finding, we calculated the distance between Leu-151 of Zn2 and the 3′ terminus of the SSB sites in the simulated complex. Our results show that Leu-151 maintains close proximity to the 3′ end, suggesting a stronger binding affinity across all observed break sites. As illustrated in Figure 22c, the Leu151–3′ distance at **site1**, and **site3** fluctuates between 4-7 Å; by contrast, at **site2**, Zn2 exhibits significant fluctuation in distance, ranging between 10-20 Å; however, it eventually achieves close contact around 5 Å after 100 ns of MD, and even shows a tendency to decrease further, at longer times. This behavior could also be attributed to a somewhat non-optimal initial docking configuration,

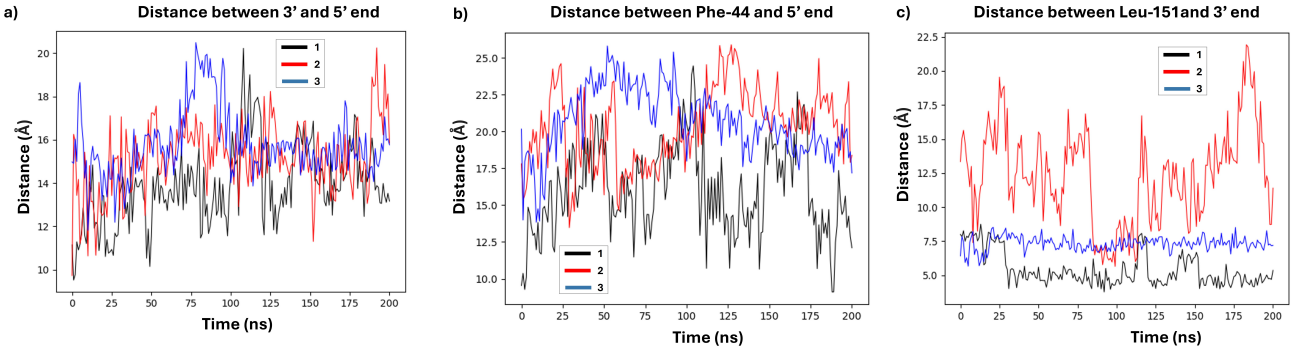


Figure 22: **(a)** SSB gap opening distance d_g between the C4′ atoms of the 5′ and 3′ termini in the nucleosome with docked Zn-finger domains of PARP-1. **(b)** Distance between Phe-44 and 5′ terminus. **(c)** Distance between Leu-151 and 3′ terminus. Black, red and blue lines respectively depict SSB sites 1 (middle), 2 (dyad) and 3 (end).

where Zn2 was positioned a little too further away from break site, and the longer MD run could help in stabilizing the contact to the final values.

The other characteristic interaction observed in the NMR study of free DNA is the parallel stacking of Phe-44 against the 5′ end of the DNA. As noted in the previous study of the DNA hairpin in Chapter 2, this alignment is achievable only when there is a substantial opening d_g between the SSB termini, allowing Zn2 to establish the first contact with the break site. This contact subsequently facilitates the subsequent Zn1 interaction with the 5′ terminus. In the case of a too narrow gap opening, which may be more typical in the more constrained conditions of the *in vivo* chromatin, these specific interactions are more challenging to observe. Hence, just like the previous study, it was more difficult to bring up these interactions with the damage sites in the nucleosome as well. As depicted in Figure 22b, all the SSB sites experience a quick movement away of Zn1 within the initial few ns of the MD simulation; even though at SSB **site1** the Zn1-5′ distance comes down to 10 Å, this does not appear as a stable interaction. These findings support the hypothesis that a very effective Zn1 interaction with the 5′ end requires a larger gap opening between the DNA termini, with Zn2 positioned closely to the 3′ end. This also suggests that, while Zn2 and Zn1 can individually and independently interact with SSB termini regions, the stability and accessibility of these interactions are influenced by site-specific fluctuations and structural adjustments, which may vary within the context of the nucleosome compared to the free DNA of the inter–nucleosome linker segments.

In addition to the role of Leu-151 and Phe-44 in their contacts with the 3′ and 5′ termini of the SSB just described, numerous hydrogen bonds are observed between Zn1-Zn2 and/or histone tails, and the DNA backbone at the SSB site. Table 2 summarizes the major hy-

SSB Site	Donor (Residue)	Acceptor (Residue)
1	ARG-122	ASP-413
	SER-418	DA-19
	LYS-22	DG-23
	ARG-122	DG-132
	LYS-148	DT-129
	SER-368	DA-27
	ARG-318	DA-134
	THR-124	DT-131
	THR-313	DC-133
	THR-313	DC-133
	LYS-126	DC-130
	DG-21	LEU-151
	TYR-404	DT-20
	ARG-156	DC-130
	LYS-148	DA-128
	ARG-314	DC-133
	ARG-18	DT-120
	SER-376	DA-27
	LYS-386	DC-127
	LYS-377	DC-26
	TYR-402	DG-123
	LYS-131	DG-132
2	THR-605	DG-71
	THR-118	DG-71
	SER-360	DC-69
	VAL-117	DG-71
	ARG-180	DT-70
	SER-20	DG-75
	ARG-18	DC-76
	SER-16	DC-76
	ARG-34	DG-75
	SER-41	DC-77
3	GLN-150	DC-3
	SER-5	DC-85
	SER-5	DC-85
	ARG-156	DT-147
	ARG-167	DC-77
	ARG-167	DC-76
	SER-5	DA-86
	ARG-8	DG-139

Table 2: Hydrogen bonds formed by the Zn12 fragment (blue) and histone tails (red) with the DNA backbone at the SSB site, during 200–ns MD simulation. Out of more than 100 H-bonds spanning the entire MD simulation, only those with more than 15% occupancy rate are reported. The three-letter codes represent residues in Zn12 (blue) or histones (red). The two-letter codes (DG, DC, DA, DT) correspond to the nucleotides.

drogen bonds observed at each SSB site over the course of 200 ns MD simulation (only H-bonds surviving for more than 15% of the trajectory, as computed with the H-bond analysis tool of VMD, are

reported). Significantly, Arg18 and Arg122 within the Zn1-Zn2 loop form stable H-bonds with the DNA backbone for **site1**, and Arg18 alone for **site2**. This same effect was observed in Chapter 2 to play a key role in inducing the bending of the hairpin–DNA structure, and is also observed in the experiments [128]. However, in the context of the nucleosome, and particularly at the more flexible **site3**, we could not observe a similar role of Arg-18 and Arg-122, likely because of the more strict mechanical/structural constraints arising in the nucleosome: for this "entry" site of the NCP, the Arg-122 loop should find a binding position a few nucleotides away from the 3′ end but the abrupt termination of the nucleosome DNA (missing an extended linker) makes such a binding position unavailable. It is therefore possible that such an interaction could be more stable in the extended nucleosome. On the other hand, Arg-18 of the Zn1 loop moves away from the DNA backbone during the simulations. However, there are numerous other H-bonds made by Zn1-Zn2 either with histones, or with the DNA backbone, which help the interaction and its stronger affinity towards the break.

3.4 NUCLEOSOMAL REORGANIZATION BY PARP-1

PARP-1 attachment to the nucleosomal DNA including a strand-break can bring about reorganization, as well as unfolding of the nucleosome structure. Experimental studies by Maluchenko et al.[146] performed on isolated nucleosomes, showed that interaction of the free DNA termini with two or more PARP-1 copies can bring about a complete reorganization of the nucleosome, while sub-stochiometric PARP-1 binding can at most bring about local changes in the DNA. Also, smFRET studies [147] showed that PARP-1 attached at the linker DNA can unfold the nucleosome, as well as induce mechanical deformation, thereby helping DSB identification (however, note again that a "DSB" in such studies is represented just by the blunt ends of the 147-bp DNA in the NCP, the nucleosome core particle, and not as a true bond break on opposite strands). Similarly, in my computational study of this Chapter including multiple SSB in the nucleosome bound to Zn-finger domains, I observed different degrees of mechanical deformation and nucleosome reorganization. As a quantitative measure of the DNA deformation, I introduced the "vertical" distance d_w between the P atom in the unpaired nucleotide in each of SSB **site1** and **3**, and the P of the closest nucleotide in the DNA winding just adjacent to it (see Figure 23d). In practice, this is a measure of the vertical distance between the two loops of DNA going around the NCP. For the sake of comparison, a similar quantity was measured between a pair of P atoms at two nucleotides in a SHL position faraway from any SSB, marked as "reference" site in Figure 23d.

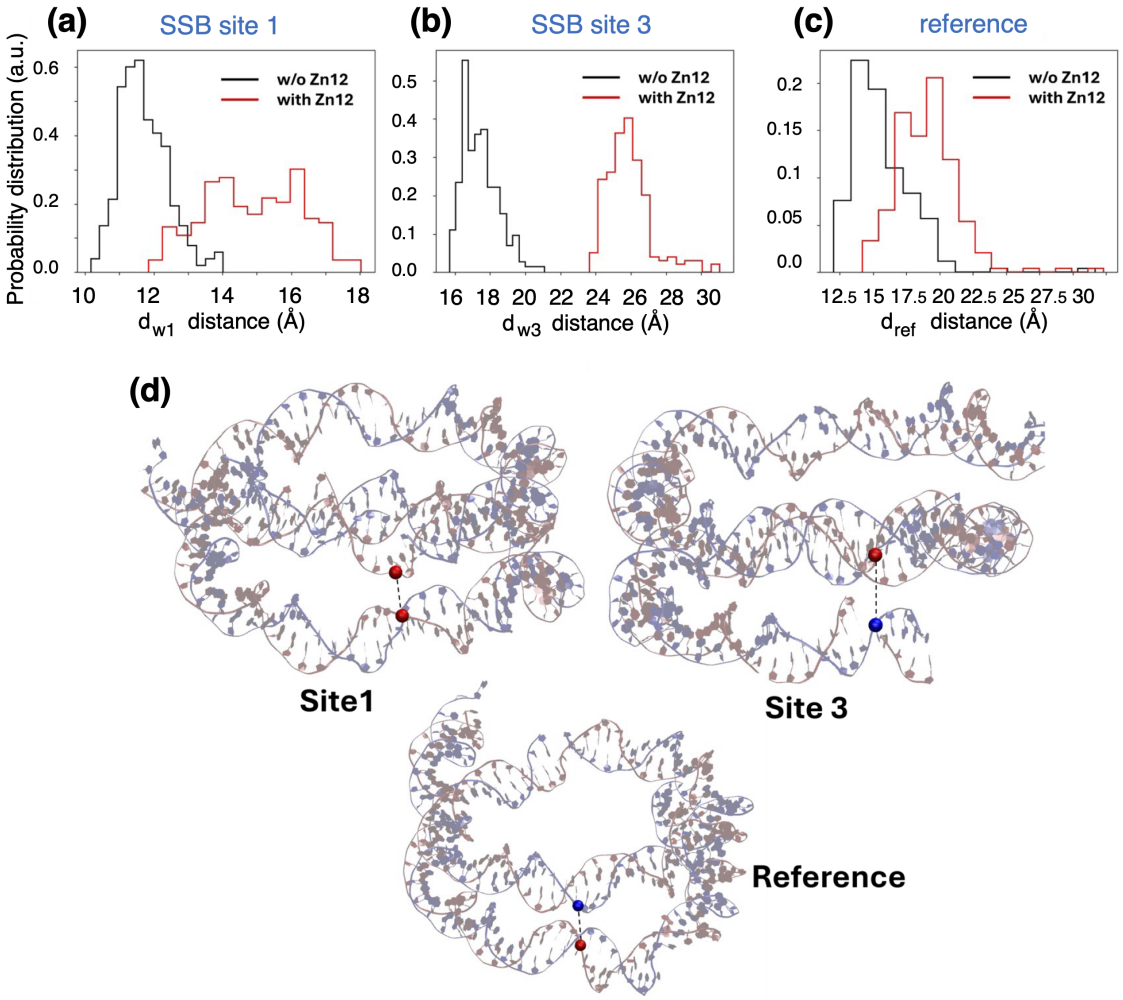


Figure 23: **(a,b)** Mechanical deformation calculated as the distance d_w between the upper and the lower windings of nucleosome DNA at SSB sites **1,3**, and **(c)** at the "reference", unperturbed site. **(d)** Definition of d_w for the three configurations, as the distance between the pairs of colored spheres along the phosphate backbone. (see text for more details).

Figure 23a-c plots the probability distribution of the distances d_{w1}, d_{w2} and d_{ref}, between the SSB site and the DNA winding just above/-below, with and without PARP-1 (respectively, red and black histograms). It can be observed that there is a huge difference in the distribution, when Zn1-Zn2 is bound to the nucleosome w/r to when it is not there. For instance, the distribution of d_{w1} at **site1** without Zn1-Zn2 is narrow and centered at 11.5 Å, while with the Zn1-Zn2 attached we observe a much broader distribution, bimodal with maxima at d_{w1} about 14-16 Å. A similar drastic change is seen for the SSB at **site3**, in which case the distribution is narrower and centered at ~16-17 Å without Zn1-Zn2, while it becomes centered at 21 Å with Zn1-Zn2 bound, the width of the histogram extending up to $d_{w1} \gtrsim 30$ Å. (Of course, I did not measure d_w at **site2** due to its location at the dyad, which has no secondary DNA winding above/-

below.) To make sure that Zn1-Zn2 is actually responsible for such a mechanical distortion or deformation of the nucleosome structure, I measured the same quantity at the reference unperturbed site, which is located opposite to the dyad between SHL-3 and SHL-4, and therefore quite far from either one of the PARP-1 copies. Interestingly, the results for the d_{ref} reference distance display nearly similar values with and without Zn1-Zn2 docked, where both distributions fluctuate between 12-22, and 13-23 Å, respectively, only occasionally visiting larger values. Hence, these results constitute a first proof that the interaction with PARP-1 can bring about partial opening as well as structural deformation of the nucleosome. Maluchenko et al. [146] observed that the distance between adjacent DNA windings could be increased up to 32 Å with the presence of an entire PARP-1 dimer interacting with the nucleosome DNA ends (that is, at SHL–/+7). In our case, it is interesting to note that just the presence of the partial Zn1-Zn2 domains (and not the complete PARP-1, that is just 20% of the total protein body) is already capable to move the two nucleosomal DNA loops away for more than 1-nm difference, and moreover well inside the NCP structure, instead of just at the DNA ends. Even though there is no complete opening of the nucleosome, it is notable that just a single Zn1-Zn2 domain was sufficient to bring about enough local deformation in the nucleosome, similar to what reported in [146]. It could be speculated that interaction with whole PARP-1 could even lead to partial or total disassembly of the nucleosome, to initiate the SSB repair (in a more complex framework, old biochemical studies had indeed shown that PARP-1 could facilitate nucleosome disassembly, by PARylating histones [152]).

To further characterize the local deformation at the SSB sites, with vs. without Zn1-Zn2, I also looked at the local twist deformation. Similar to Section 3.2 above, the local twist is measured all along the MD trajectory by the dihedral angle Φ, which could highlight the under-/over twisting of the nucleotides flanking either side of the unpaired nucleotide. Indeed, given the large opening deformation, a significant change in the local twist values could also be expected, with vs. without Zn1-Zn2. However, the variations between the two conditions displayed in the plots of Figure 24, are in this case more subtle. The SSB **site1** displays the largest variation, with the probability distribution of Φ going from a skewed gaussian centered at $\Phi \sim 30^{o}$ without Zn1-Zn2, to a broader gaussian centered at a larger over-twisting of $\Phi \sim 70^{o}$ when Zn1-Zn2 is interacting at the SSB; SSB **site3** has a similar behavior, however with a smaller variation between the maxima, from $\sim 55^{o}$ to $\sim 75^{o}$, and a similar shape evolution of the Φ probability distribution. At the SSB dyad **site2**, the difference in Φ twist angle with vs. without Zn1-Zn2 is yet less evident, the two symmetric distributions going just from about 60 to about 70° peak value. Overall, these results confirm that the presence of Zn12 brings about

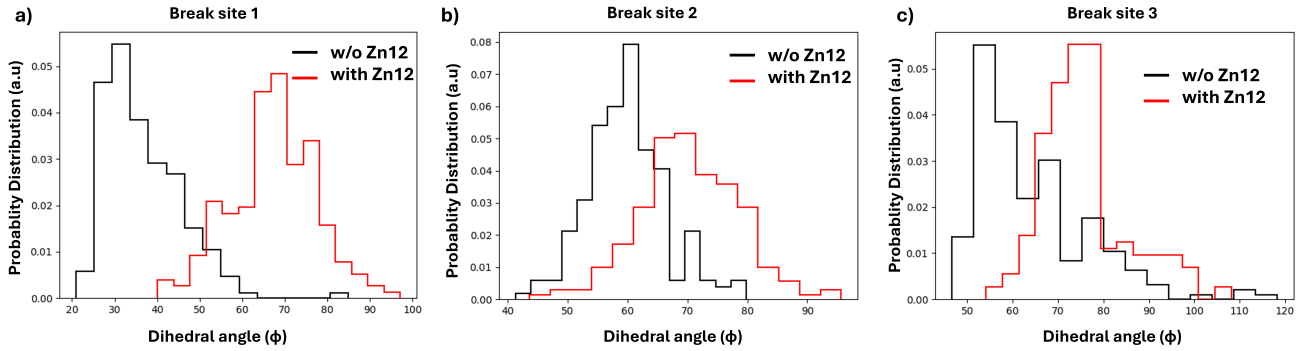


Figure 24: Probability distributions of the dihedral angle Φ for each break site, calculated during 200-ns MD trajectories, with vs. without Zn1-Zn2 bound.

a local conformational change in the nucleosome, which might help in damage identification by PARP-1 [147].

3.5 MECHANICAL DEFORMATION AT SSB SITES DUE TO HISTONE TAILS

At SSB **site1** and **site 3** it is observed a gradual mechanical deformation developing during the MD, along the nucleotides near the 3′ strand, which was absent at **site2**. To investigate this further, I analyzed the structural arrangement and energetics of the ensemble of 12-15 nucleotides surrounding the SSB + the Zn1-Zn2, and histones H2A, H3 and H4, using the PDBePISA software. Figure 25 provides a graphical illustration of the details of hydrogen bonds established between aminoacids and nucleotides near the two SSB sites. It is worth noting that histone H2A makes H-bonds with a stretch of nucleotides opposite to the SSB; whereas, the N-tail of H3 makes H-bonds with the Zn1 finger, and the H4 "bridges" Zn2 with the 5′ SSB terminus.

It can be concluded that Zn2 induces significant structural shifts in the DNA backbone, potentially facilitated by the dynamical action of histone tails. This leads to substantial gap openings d_g between the 3′ and 5′ ends of the SSB, aligning with the observations in Chapter 2. Such opening allow Zn2 to establish interactions, and subsequently recruit Zn1 to the break site. These findings suggest that DNA unwinding should be essential for PARP-1 to perform PARylation, while Zn-finger domains facilitate the reversible unwinding necessary for this process [147]. Interestingly, the behavior observed at SSB **site1, 3** was unique and not replicated at the dyad **site2**. It is however possible that further exploration, especially with longer simulation trajectories and similar contacts, may yield comparable outcomes for other SSB sites.

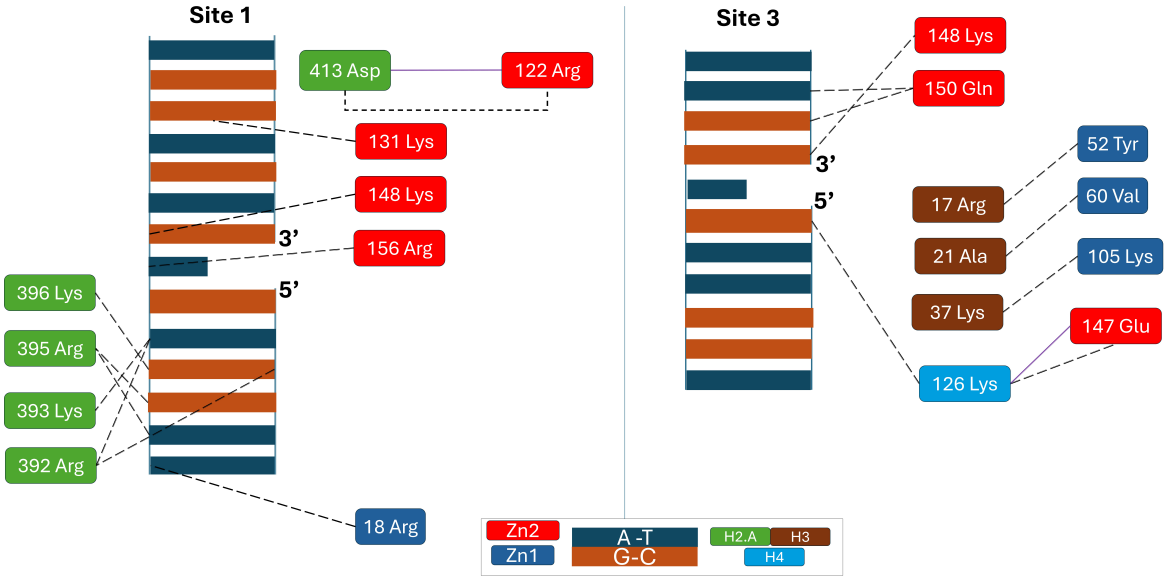


Figure 25: Graphic showing major H-bonds between DNA vs Zn1-Zn2 and histones for **site1** and **site3**. Dashed lines represent H-bonds, while solid lines represent salt bridges. The color code provided in the lower-right corner helps in labelling aminoacids in the two Zn-fingers (red/blue) and in the histones (green/sky-blue/brown); and A-T, G-C base-pairs in the DNA; the unpaired nucleotide in each panel identifies the SSB location, also indicated by its $5' - 3'$ termini.

3.6 SUMMARY AND OUTLOOK

In conclusion, in this Chapter it was found that PARP-1 interacts with nucleosomal DNA by binding to the SSB sites, in a way similar to what observed in free DNA. However, it also reorganizes the local DNA and nucleosome structure, aligning with experimental results that show its importance in compacting the chromatin, and histone chaperone activity during damage repair [148, 153, 154]. These studies show that PARP-1 bound either to an isolated (mono)nucleosome on its linker DNA, or at a tri-nucleosome aggregate with free linker DNA, can reorganize and compact their structure. Even though several experimental studies report a much reduced affinity of PARP-1 to nucleosomes without the DNA linker, it remains open the question about how the identification of damage could happen, when the DNA damage (notably, SSB or DSB) is located within the nucleosome structure, that is, someplace on the DNA wrapped around the histone core. Hence in this part of my study I decided to investigate the interaction of PARP-1 zinc-finger domains (Zn1-Zn2) with single-strand breaks (SSBs) in nucleosomal DNA, focusing on the attending structural and mechanical consequences.

The results revealed that SSBs induce local opening of nucleosomal DNA, with fluctuations in gap distance driven by both thermal and mechanical effects. The presence of Zn1-Zn2 enhanced these open-

ings, particularly at the break sites, facilitating critical interactions such as the proximity of Leu-151 to the 3′ end, and hydrogen bonding with the DNA backbone. Compared to the same interactions observed in the experimental studies on free DNA hairpins, achieving the PARP-1 "correct/trademark" interactions in the context of the nucleosome poses substantial difficulties. Also, there are no experimental literature that could support that Zn1-Zn2 interacting with nucleosome can bring about the same interactions as observed in DNA oligomer experiments. Most studies support the bending propensity of the free DNA/linker DNA, and the higher affinity of Zn1-Zn2 [128, 147]. The bending propensity is out of scope for this study because I used a mononucleosome with no free linker extending near the entry/exit points. However, I verified the high affinity of Zn2 for the 3′ SSB ends, and that interaction with Zn1-Zn2 can indeed bring about an increase in the gap opening at the SSB site, helping it to interact with Zn2 domain. Muthuraj et al. in their studies [154] raised the hypothesis that the Zn-finger domains could not be possibly involved in terminal DNA-interaction in a tri-nucleosome complex. It was also difficult to observe the "trademark" interactions in my model system, but still there were numerous H-bonds implicating Zn1-Zn2, the histone tails and the DNA backbone, which sufficiently brought the SSB to open up, and constitute an integral part of SSB identification by PARP-1. Along with many intricate interactions, the presence of Zn1-Zn2 brought about structural reorganization in the nucleosome, by increasing the splitting between adjacent winding about the strand-break site. Altogether, these results highlight the unique structural contributions of Zn1-Zn2 to nucleosomal remodeling, underscoring its relevance as a potential focal point for understanding the PARP-1 broader roles in genome stability and damage response pathways.

4

DAMAGE IN THE CHROMATIN: LOCALIZATION TO THE LINKER DNA

4.1 INTRODUCTION

The very first step of any attempt to repair damage that is highly localized to one specific position on the DNA, most often to just one nucleotide, is a formidable task: finding this site amid billions of base-pairs of undamaged DNA. Pure 3D-diffusion of proteins, randomly searching DNA stretches for a possible damages has long been thought to be insufficient for the DNA-repair process. Repair proteins must be extremely efficient to inspect all the positions in the genome ($\sim 3 \times 10^9$ base pairs), and detect and repair damaged bases (around 1 in 10^6 base-pairs per cell daily). Given that for each protein there could be $\sim 10^4 - 10^5$ copies in the nucleus, each such protein (excluding redundancy) should be able to sample around 70,000 base pairs of DNA in every cell cycle (about 12 to 24 hours, see e.g. Ref.[155]). The fact that, in practice, repair proteins can achieve their job much faster, largely defying any "facilitated diffusion" model [156–158], points to the existence of better than purely stochastic search mechanisms.

It was proposed a number of decades ago that the search of a protein for a specific site on DNA (be it a repair protein finding lesions, a transcription factor binding to promoters, or any other DNA-binding protein that needs to associate, with a unique and specific DNA position) could be drastically sped up, by combining 3D diffusion with short stretches of 1D-search along a DNA fragment adjacent to the binding site [159, 160]. In such a picture, the protein would randomly bind to a position on DNA, then scan around the DNA in a sort of random walk-like fashion, and subsequently dissociate to rebind at another position, and repeat this cycle until some target site is found. The idea behind this, is that 1D diffusion is in principle quite faster than 3D, albeit for this to be achieved the protein must "climb" the physical/chemical energy barriers that develop all along the DNA backbone (whereas 3D diffusion occurs essentially in the solvent).

Among the various modes of possible protein-DNA interactions, both sliding and hopping along DNA represent a kind of 1D-diffusion. However, their mechanisms are fundamentally different and this has led to some debate in the scientific community. Proteins undergoing 1D sliding must maintain contact with DNA continuously, through a sort of corkscrew motion along the helical path [161, 162]. In contrast, hopping should involve microscopic dissociation from, and rebinding to, the same piece of DNA, but without fully dissociating into

solution [160, 163]. Some of these mechanisms have been actually identified experimentally [164–166], however this complex problem is still quite open.

Compared to glycosylases such as UDG or OGG1 implicated in the early steps of the BER and NER pathways, or to Ku70/80, MRE11, Rad50 and NBS1 implicated in DSB recognition for NHEJ, the PARP1 protein is a very big object, with its >1000 aminoacids for a total molecular weight of >100 kDa.The hopping on DNA of such a large object should occur with slow diffusion times according to Stokes-Einstein's equation for the diffusion coefficient,

$$D = \frac{k_B T}{6\pi\eta R}, \tag{1}$$

in which the molecular size R is at the denominator.

In an interesting series of experiments by atomic-force microscopy (AFM) of long DNA oligomers deposited on a surface [167], was seen to bind to nicks, free ends, and abasic sites, while conducting damage search mostly through 3D diffusion. In the experiment, co-localization of APE1 with PARP-1 on DNA was found capable of inducing 1D diffusion, while excess APE1 also facilitated the dissociation of DNA-bound . However, the diffusion coefficients measured are in the range 10^{-11} cm^2/s, still very low for a purely random search of the defects. On the other hand, such an experimental configuration is interesting as a coarse model of chromatin, since relatively long stretches of dsDNA were used, about 400-500 bp in length, suspended between silica beads in a "tightrope assay", that could mimic the DNA linker joining pairs of nucleosomes (albeit the typical linker lengths should be rather in the range of 20-100 bp).

Linker DNA, the stretch of DNA between two nucleosomes, plays a crucial role in chromatin condensation. By using large-scale MD simulations, Kenzaki and Takada demonstrated that variations in linker DNA length between trinucleosomes significantly affect their folding into zig-zag structures or distributed nucleosome stacking [168]. Longer linker DNA lengths (45-55 bp) facilitate more efficient chromatin folding and nucleosome stacking in 0-start, 1-start, and 2-start fibers compared to shorter linkers [169]. Additionally, studies show that longer linker DNA reduces the energy cost of deforming the DNA during nucleosome stacking, since the mechanical stress is better distributed across longer stretches of base-pairs [169].

The connection between DNA damage and the attending mechanical reorganization may be a key point, still overlooked by the biological literature. Studies about chromatin deformation from DNA damage demonstrated that chromatin reorganization following damage impacts nuclear mechanics, potentially altering how cells respond to, and repair DNA lesions [170]. These studies offer insight into how DNA damage can interplay with mechanical changes within the chromatin and nucleus, affecting both cell structure and function. How-

ever, despite the existing understanding of how mechanical forces and DNA damage interact, there remains limited research on the intricate dynamics of DNA strand breaks, particularly in relation to mechanical deformation in the nuclear context [141]. One interesting possibility to get around the problem of how repair proteins identify the damage sites among the billions of undamaged base pairs, instead of purely random searching (be it either 3D or 1D-diffusion), could be that **damage sites undergo important mechanical deformation**, and such deformed DNA structures could be more easily targeted and recognized by repair proteins, compared to "normal" DNA. The local deformation at the defect sites would make the damage more "accessible", by inducing a mechanical stress field along the DNA backbone that facilitates target identification by the repair proteins [120, 141, 151]. For example, a recent study on DNA mismatch-repair dynamics Ref. [171], highlighted that a sequence of mismatches (such as CC-C/CCC and TTT/TTT) in the oligonucleotides trigger conformational changes by thermal fluctuations. While it is well known that mismatches induce local DNA deformation [172], with measurable impact on the binding free-energy [173], in this case laser temperature-jump spectroscopy combined with FRET demonstrated that the rapid fluctuation dynamics of DNA "guide" Rad4/XPC protein to interact with the specific damage sites. Their study suggests that structural deformation of the DNA may indeed act as a special pathway for the damage sensing by repair proteins. In the previous Chapters 2 and 3, it was shown that PARP-1 requires a large deformation to open up at the site of the SSB, in order to recognize it as a defect, and attach to it via its Zn1 and Zn2 units.

Mechanical deformation of DNA makes the damage sites more "visible" to the repair proteins

In Chapter 2, simulations of single-strand breaks (SSBs) in free, long DNA showed that the break does not open up spontaneously widely enough, to allow to interact with, and identify the defect. However, a chunk of free long DNA cannot be a good representation of the chromatin environment, mainly because of the lack of mechanical binding at its ends, which restricts the entropic freedom and modifies the free-energy landscape (other factors being, e.g., the role of proteins and histone tails). In the reality of the cell nucleus, DNA is organized into chromatin, with nucleosomes and linker DNA subjected to various physical stresses such as thermal agitation, as well as compression, sliding, and stretching from the intrinsic chromatin dynamics or due to mecanical stresses suffered by the cell

In this Chapter I will investigate the more realistic scenario of a linker-DNA stretch, flanked by two nucleosomes, as the minimal mechanical configuration that could represent a model of chromatin fragment. While in the previous Chapter 3, I showed interacting with nucleosomes at various key sites around the structure, here I focussed on the ability of linker-DNA carrying localized breaks (SSB and DSB) to induce localized mechanical deformation. As it will appear in the

foregoing, preliminary results suggest that thermal fluctuations alone are sufficient to induce enough opening and twisting at DSBs, to potentially attract . However, for SSBs the thermal fluctuations alone seem to be insufficient; instead, mechanical fluctuations arising from chromatin dynamics can facilitate the necessary deformation. Clearly, the final cherry on the cake would have been to include simulations of interacting with such defect configurations also in the linker-DNA, however the time constraints of the thesis did not allow to complete this final step, which is left for the immediate future developments.

Hence, this Chapter 4 aims to answer the following questions:

1. How do DNA strand-breaks in linker regions between nucleosomes behave under mechanical stress?

2. How/to what extent the different types of mechanical deformations could facilitate defect recognition by repair proteins?

For the purpose of this study, I performed a series of all-atom MD simulations on a basic dinucleosome configuration. The molecular structure was designed in such a way that there are two equal copies of a nucleosome (PDB ID **1kX5**) along with different lengths of linker DNA. The linker length varies from 24-bp(short-linker) and 42-bp(long-linker) between the two nucleosomes. Both nucleosomes include full histone tails, so as to ideally mimic the chromatin environment. Subsequently, for each of the linker length variations, I included DNA damage in the form of either a SSB, or a DSB, which was compared to the same dinucleosome without defects, thereby resulting in total of 4 variations (see Figure 26).

All the MD simulations and most data analyses were carried out using the same standard protocol described in Appendix B. To ensure structural stability of the modified configurations, I initially performed energy minimization using the steepest-descent algorithm, followed by constant-{NVT} and -{NPT} equilibration for 150 ps to bring the system to 310 K and 1 bar pressure. The dinucleosome complexes were solvated in water boxes of approximately 35x35x35 nm^3 in size, containing about $1,380,000$ water molecules, along with Na^+ and Cl^- ions to neutralize the system and maintain a physiological salt concentration of 0.15M, for a total of about 4,2 million atoms. After minimization and constant-{NPT}-equilibration, all the four variants of dinucleosome were allowed to relax for 10 ns at constant-{NVT} without restraints.

It is worth noting that the DSB-containing variants already exhibited spontaneous mechanical deformation at the damage sites, hence the MD trajectories were just followed for the required time length, of more than 100 ns each, without further intervention. Instead, the SSB-containing variants did not display substantial deformation under the only effect of thermal fluctuations, at least for the short time

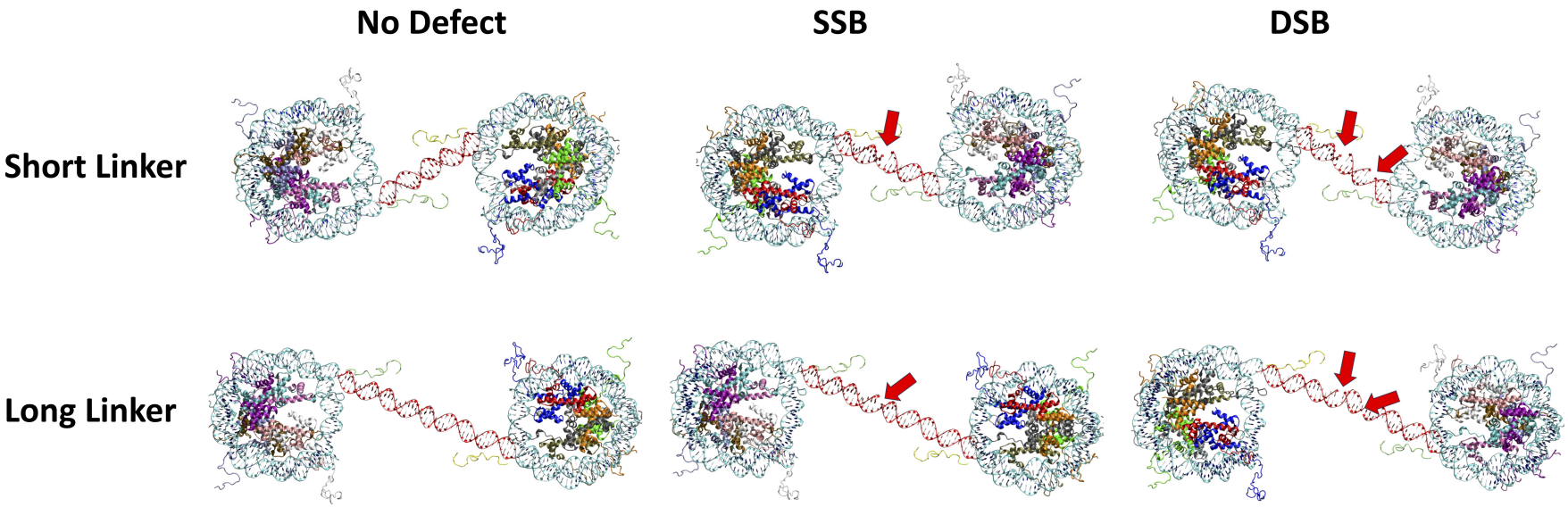


Figure 26: The 6 structural variations of the dinucleosome, categorized by defect type and linker length. Red arrows indicate the sites of phosphate breaks with a missing nucleotide. Note that for graphical purpose the scaling of the images in each row is different. Also, in the first row the image arrangement gives the impression that the two nucleosomes are interacting, however in the course of the MD simulation they remain always separated away from each other.

window explored here. Therefore, I performed steered molecular dynamics (see Appendix B.3 for the GROMACS "Pull-code") for these SSB-containing variants, by applying equal and opposite forces on both nucleosomes, such that they could slowly come closer and expose the strand break. Altogether, the ultimate goal was to induce some form of thermal and/or mechanical deformation/flexiblity in the linker DNA, such that they could expose the break, either SSB or DSB. As already said, it is our idea that exposing the break by localized mechanical deformation, and its correlated stress field, could be a key event facilitating the identification of the damage by the repair proteins, potentially accelerating the repair process well beyond the typical diffusion time-scales. Moreover, these subtle structural changes in the DNA could also play a significant role in chromatin dynamics.

4.2 Comparison of DNA linker deformation with vs. without a SSB

4.2.1 *Simulation setup and force application*

The two variants of the dinucleosome with different linker lengths (24- and 42-base-pairs, with randomly chosen nucleotide sequence) and the presence of SSB (middle column of Figure 26) were subjected to a longitudinal force on each nucleosome core, such that they could come close to each other and induce deformation in the DNA linker. This turned out to be necessary since even after a rather long MD trajectory with only thermal fluctuations, the two nucleosomes remained at a nearly constant distance, and the DNA linker between them remained in the extended conformation. We also tried in the beginning a very short linker-DNA configuration with 20-bp that was readily excluded from the rest of the study, since its extremely short linker made the two nucleosomes too close and constantly interacting. Notably, this sets a sort of "lower cutoff" length for observing linker-DNA mechanics.

For the purpose of inducing mechanical deformation, I applied a linear force using the "Pull code" in GROMACS 2020.4 package [174, 175]. The pull-code module (see Appendix B.3) allows to apply force or constraints on collective variables/reaction coordinates, using either distance, angle, or dihedral between the centre of mass of the predefined groups of atoms, called the "pull groups". These groups can also be moved in a particular direction, provided by a direction vector supplied by the user. In this study I used three pull-groups, namely one group for each nucleosome core, and one including the linker DNA and nearby histone tails (the reference group towards which the two other are moved by the applied force). Each nucleosome core group is subjected to an initial harmonic potential using the so-called

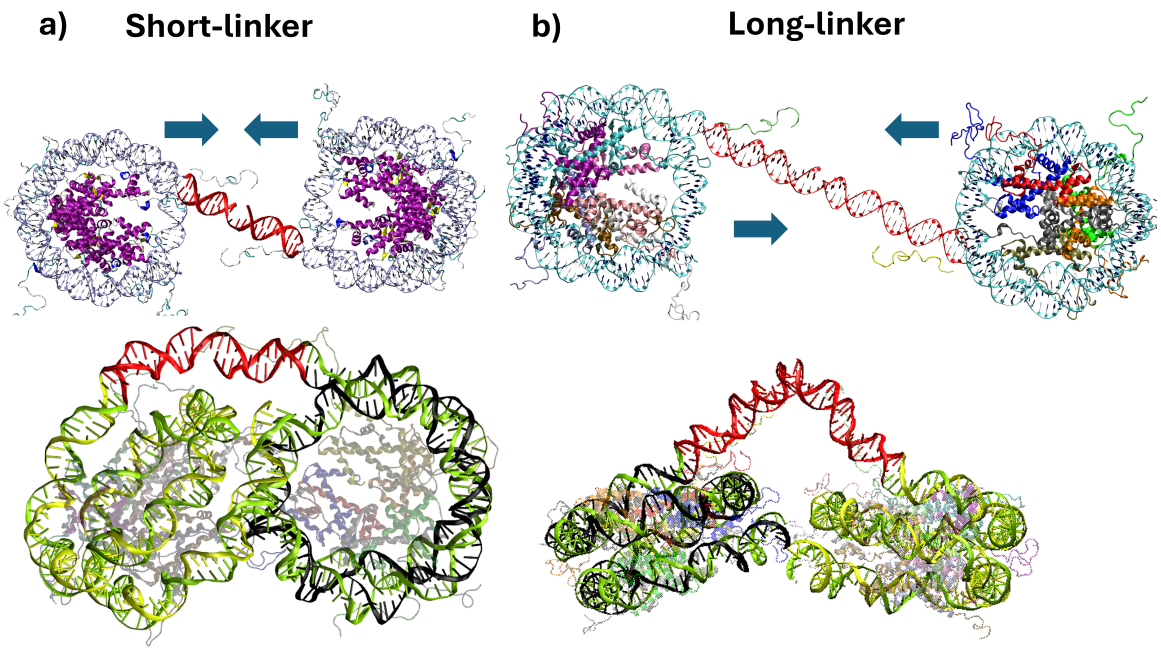


Figure 27: The upper panel shows the initial structure for the two variants including the SSB in the linker-DNA; the lower panel shows the configurations with bent linker-DNA after pushing them together. The red-colored fragments highlight the position of linker-DNA; the histone proteins are shown transparently for better view; the arrows show the direction of pulling forces.

"umbrella pulling" between the centers of mass (COM) of each pull-group, while keeping the linker-DNA as the reference group, which remains at an average constant position. To evaluate the structural changes induced by the presence of defects in the DNA, similar MD trajectories were also conducted for two analogous linker-length variations of dinucleosome but without the SSB (left column of Figure 26), using the same setup. While such a procedure is indeed artificial, however it could mimic the more realistic compressive/shearing phenomena actually occurring in the chromatin during its internal dynamics, for example during transcription, replication, compaction into heterochromatin, and so on; notably, the levels of forces used, in the 10-pN-range (see below), are compatible with typical forces exerted on chromatin by nuclear lamins, actin filaments, and molecular motors [176].

As previously mentioned, our primary objective was to observe mechanical deformation at the break site. From the study in Ref.[141], we guessed that a likely mechanical effect from the presence of the strand-break should be some form of localized bending or kinking of the DNA helix, which would geometrically bring the two nucleosomes closer to each other, and expose the break site. To achieve this in the case of SSB, the "umbrella" pulling was applied in the positive and negative Y-direction joining the two COM (see Figure 27), using

an appropriate value for the effective spring constant of the fictitious force (k) and displacement rate (dx), to bring the nucleosomes close to each other and induce a bend or kink in the linker DNA. For both the short- and long-linker variations, I used k=10 kJ/mol/nm^2 and dx=0. 5 nm/ns. The resulting apparent force applied by the umbrella potential to the fictitious spring is 110 pN. However, this is not the same level of force that is actually pulling the nucleosomes. This can be more directly estimated by looking at the constant drift regime imposed by the umbrella potential, in the approximation of linear viscous drag. In the linear regime, the applied force is equal and opposite to the viscous drag force. To apply the Stokes' formula for a spherical object moving in a fluid:

$$F = 6\pi\eta R v \tag{2}$$

we need the nucleosome velocity in the linear drift regime, which is about 0.8 nm/ns (as deduced from the plot in Appendix A, Figure A16); the nucleosome radius R=5.5 nm; and the water viscosity, which for the TIP3P model is about η=0.3 mPa·s at 310 K and 1 bar [177].[1] As a result, an average force F $\simeq$25$\pm$5 pN acting on each nucleosome is derived. Altogether, the average force applied for all the four variations (defect-free vs. SSB) is small enough to ensure that each pair of nucleosomes is gently compressed towards each other, as seen in the lower panel of Figure 27, without affecting their structure. Furthermore, it is a level of force compatible with typical values observed in the chromatin (forces directly transmitted from the cell cytoskeleton via the nuclear envelope being even higher, in the 0.1-1 nN range, [176]). Regarding the time-scale of entire simulation, each system was simulated at various time points until a desired arrangement with the two nucleosomes enough close to each other was obtained, which varied from 5-15 ns for either variation.

4.2.2 *Structural changes induced by SSB on the whole nucleosome*

As seen in the last Section, a very small force applied on the system is sufficient to bring the two nucleosomes together by progressively reducing the distance between them. To measure the relative displacement between the two nucleosomes, the distance between the center of mass (COM) of each nucleosome was measured, relative to the initial distance. The relative bending angle is obtained by measuring the

[1] Such a value is indeed quite off from the experimental value, η=0.89 mPa·s. In their work, González and Abascal show that other water models give much better predictions, for example the 0.79 for the SPC/E, or 0.84 for the TIP4P/2005. However, the AMBER14 force field is well tuned to the TIP3P water model, which is generally used by the bio-molecular modelling community. Of course, no water model can perform optimally on every observable, and viscosity is a weak spot for TIP3P, which is however quite good for many other water properties.

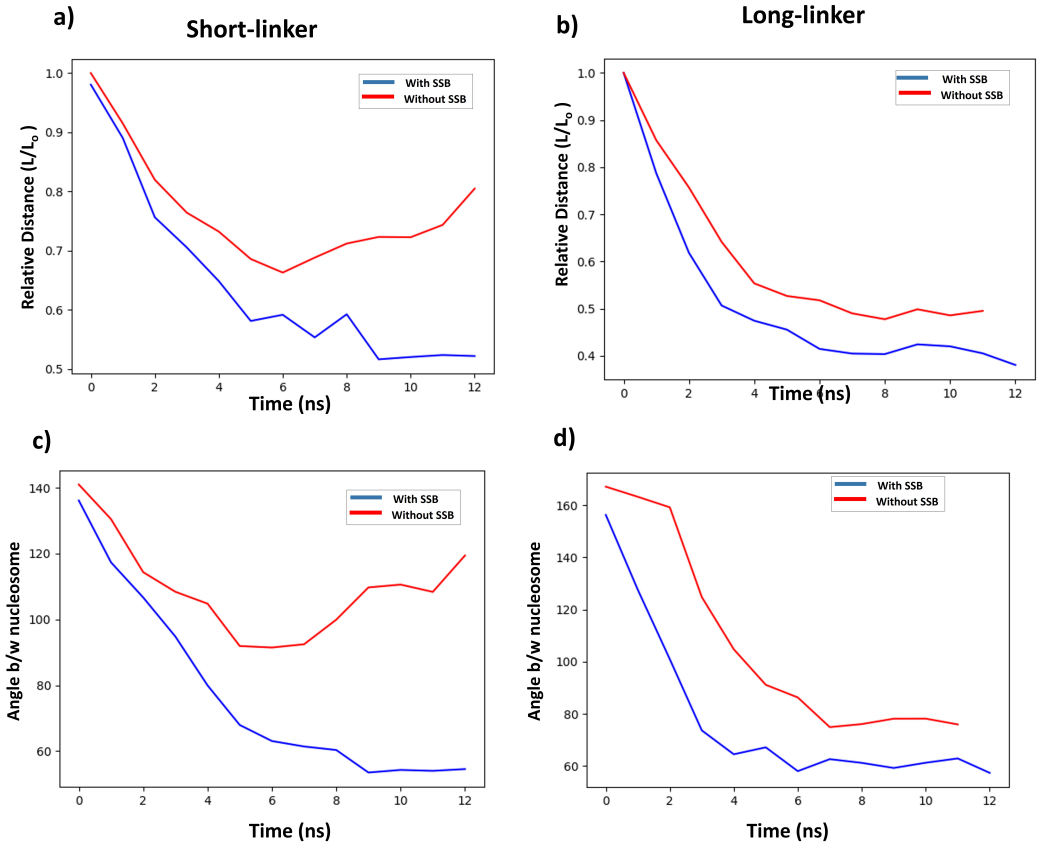


Figure 28: (a,b,c) Relative distance in the dinucleosome pair with and without SSB; L_o is the initial distance. (d,e,f) Angle between the two nucleosomes over 12ns for both variants.

angle subtended by the three COMs, with the linker pull-group comprised between the two approaching nucleosomes. For all the four variants (with vs. without SSB), Figure 28 reports the relative displacement due to the applied force (upper row, L_o representing the initial distance between the nucleosomes), and the bending angle about the linker-DNA (lower row).

With identical initial conditions, and equal initial distance between the nucleosomes, a more drastic compression in both SSB-containing variants were observed. The systems containing a SSB has a relative distance of about 0.7 and 0.6 times the initial distance, respectively for longer and shorter linker , while the same systems without the damage reached a maximum approaching distance of only 0.5 and 0.3 times the initial value. Also, assessing the bending angle of the dinucleosome reveals a similar trend for both the short- and long-linker variants. In the SSB-containing systems, the bending angle was reduced to less than 60°, compared to the system without the defect, in which the angle closed to just about 90-100° (Figure 28d). Both results for the displacement and bending angle demonstrate a higher rigidity of the undamaged DNA linker, which is substantially weakened by the presence of just one SSB approximately positioned mid-

way between the two nucleosomes. Overall, these results show that the presence of SSBs significantly alters the energy dynamics, relative compression, and bending propensity of nucleosome systems, with these effects being most pronounced in longer vs. shorter lengths. Furthermore, these analyses reinforce the idea that the presence of the damage introduces notable mechanical deformation in the DNA backbone, which may play a crucial role in the ability to recognize damaged regions effectively.

4.2.3 *Structural and mechanical changes in the linker-DNA*

The comparison of the dinucleosome structures with vs. without a single-strand break (SSB) under identical soft force conditions revealed distinct structural deformations in the linker DNA. The sequence of snapshots in Figure 29 illustrates how even under the moderate load of about 25 pN, the presence of the SSB strongly altered the structural dynamics of the linker-DNA, resulting in pronounced bending and kinking behaviors. While short-linker DNA with an SSB exhibited a more mild kinking and bending, the longer-linker DNA displayed a progressively larger and sharper kinking at the SSB. By contrast, the pristine linkers without damage accommodated the extra mechanical stress by twisting and bending, maintaining a conformation as straight as possible.

Previous studies by MD simulation, showed that under steady compression, a trinucleosome system connected by 36-bp linkers develops several types of mechanical instabilities, going from simple Euler buckling, to a kind of Brazier-like kinking, to coupled twist-bending [141] (see Appendix A, Figure A15). Such types of mechanical instabilities are commonly observed when applying a uniaxial compression to meso-macroscopic systems, from cylindrical rods to microcrystals (see, e.g., [178, 179]). Following the assumptions of that study, one can assume that compressed DNA behaves similarly to a cylindrical rod (approx. radially homogeneous) with Young's modulus E=300 MPa, length L=10 nm (that is, about 30 bp), and radius R=1.1 nm; it turns out that for Euler-like instability to occur, a minimum load P_c ~25 pN is required, from the equation

$$P_c = \frac{\pi^2 E \mathcal{J}}{L^2} \tag{3}$$

where $\mathcal{J} = \frac{\pi R^4}{4}$ is the transverse cross-section moment of inertia. That study suggested that upon increasing the compressive load above the "critical" value of ~25 pN, DNA would transition from Euler bending to some form of coupled twist-bending, in order to relax the excess stress [141]. Our results concur with these findings, in that no kinking but just a moderate bending (Euler-like) was observed in the linker DNA without damage, under a similar compression force of 25 pN (Figure 29, lower row).

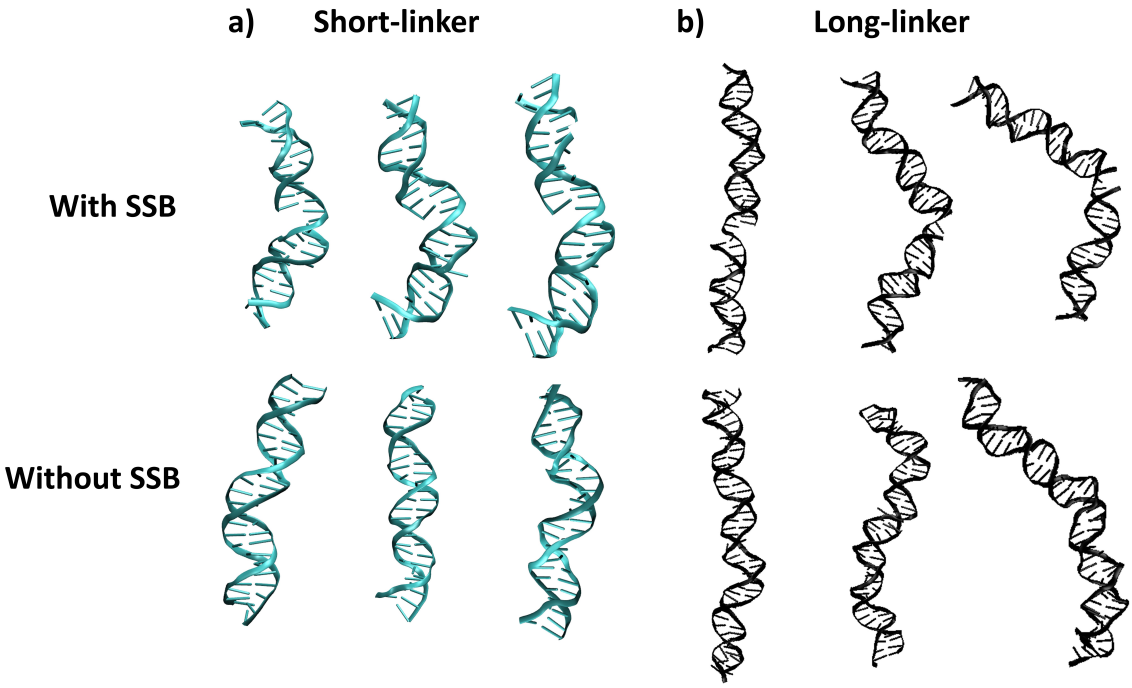


Figure 29: Molecular-scale representation of the deformation of the linker-DNA lengths at various time frames, with vs. without SSB. It may be noted that the defect-free DNA develops a pronounced twisting/bending to accomodate the excess stress (with partial extrahelical expulsion of a bp, see text), whereas the damaged DNA develops a visible kink at the SSB site, while the two segments flanking the kink remain rather straight.

By contrast, the presence of any defect, such as a hole in the cylinder, can reduce the buckling stress, thus altering the overall stress distribution throughout the cylinder. Similarly, a defect like an SSB/DSB in DNA can drastically reduce the buckling stress, suggesting that a defect near the neutral axis (imagined as the line running through the geometric center of the cylinder, in the linker DNA) significantly lowers the critical load due to imperfections, stress concentrations, or localized deformations. By looking at the molecular structures in the upper row of Figure 29, it is suggested that a sharp kinking forms at the SSB site, already under the same compressive force of 25 pN, all the more evident in the longer-linker DNA (right-side frames). In line with previous findings, these observations further support the hypothesis that chromatin deformation is heavily dependent on the presence of structural defects in the DNA, notably the linker DNA between nucleosomes.

The development of "Brazier-like" kinking can be appreciated more quantitatively, by looking at the intra- and inter-base-pair parameters of DNA, and monitoring the variation in helical parameters [180] in the linker-DNA, averaged over the final part of the MD trajectory. Such local parameters, defined in Figure 30, completely define the deformation within each base-pair (bp), and between adjacent stacked-

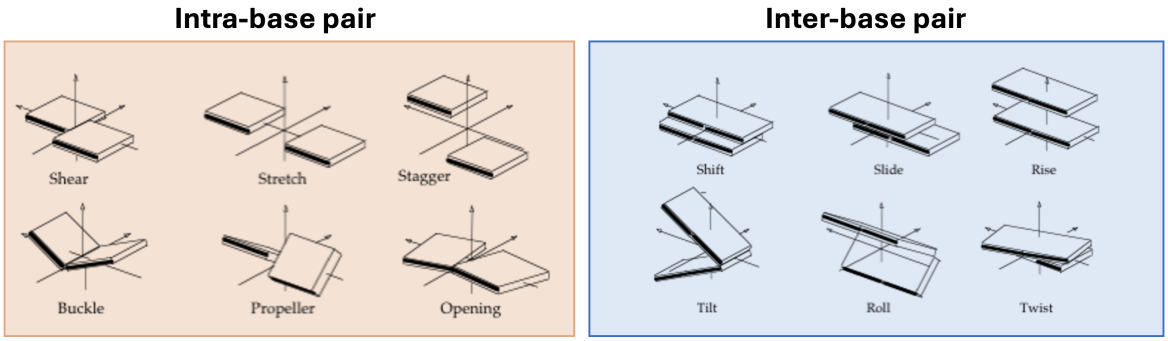


Figure 30: Geometrical definition of the structural helical parameters of DNA. Note that there are 3 rotations and 3 translations about the $\{x, y, z\}$ axes for both intra- and inter-bp configurations. **Left** Intra-base-pair parameters. Upper-row: Shear, Stretch, Stagger define relative translations; Buckle, Propeller, Opening define relative rotations between the two bases. **Right** Inter-base-pair parameters. Upper-row: Shift, Slide, Rise define relative translations; Tilt, Roll, Twist define relative rotations between two adjacent base-pairs.

bp, and can be readily computed by the CURVES+ free-software module [181]. Due to the software limitations, it was necessary to add a dummy nucleotide between the 5′ and 3′ termini where SSB is located, for all the variants including SSB; this dummy nucleotide was later removed for the analysis.

The plots provided in Appendix A, Figure A11, highlight the mechanical deformations of DNA linkers with- (blue plots) vs without-SSB (red plots). It can be clearly appreciated how structural disruptions influence the double helix. For the short-linker, the presence of the SSB promotes the flip-out of a nearby base (position 8 in the plots) that shows up clearly in the Stagger, Stretch and Opening parameters; Stretch and Shear show the deformation at the bp flanking the SSB (pos. 13-15); the deformation between adjacent bp flanking the SSB is also evident in the Shift and Rise inter-bp parameters, by the sharp up-down flipping of the two peaks with respect to the average values, and by the peak in Twist. The flipped-out bases are illustrated in the snapshot Figure 31a, with the A·T bp shown in blue-red thicker sticks.

The long-linker follows this pattern of concentrating the deformation within the (intra)-bp parameters Stretch and Shear (Appendix A, Figure A12. However, in this case the undamaged DNA displays a very interesting peculiarity: it will be noted that very sharp peaks appear in the red plots at the very same location (pos. 19-21) where the SSB is located. This linker sequence is slightly richer in A·T bp (56%) compared to G·C (44%), compared to the short linker in which the proportions are reversed; however the sequence around pos. 18-21 is a ds(TAAT), similar to a TATA-box, and therefore quite flexible compared to other parts of the sequence; at that spot, the central bp A·T at pos.20 flips out towards the major groove, followed by minor rotation also of the A·T at pos.19 and T·A at pos.21, to accommo-

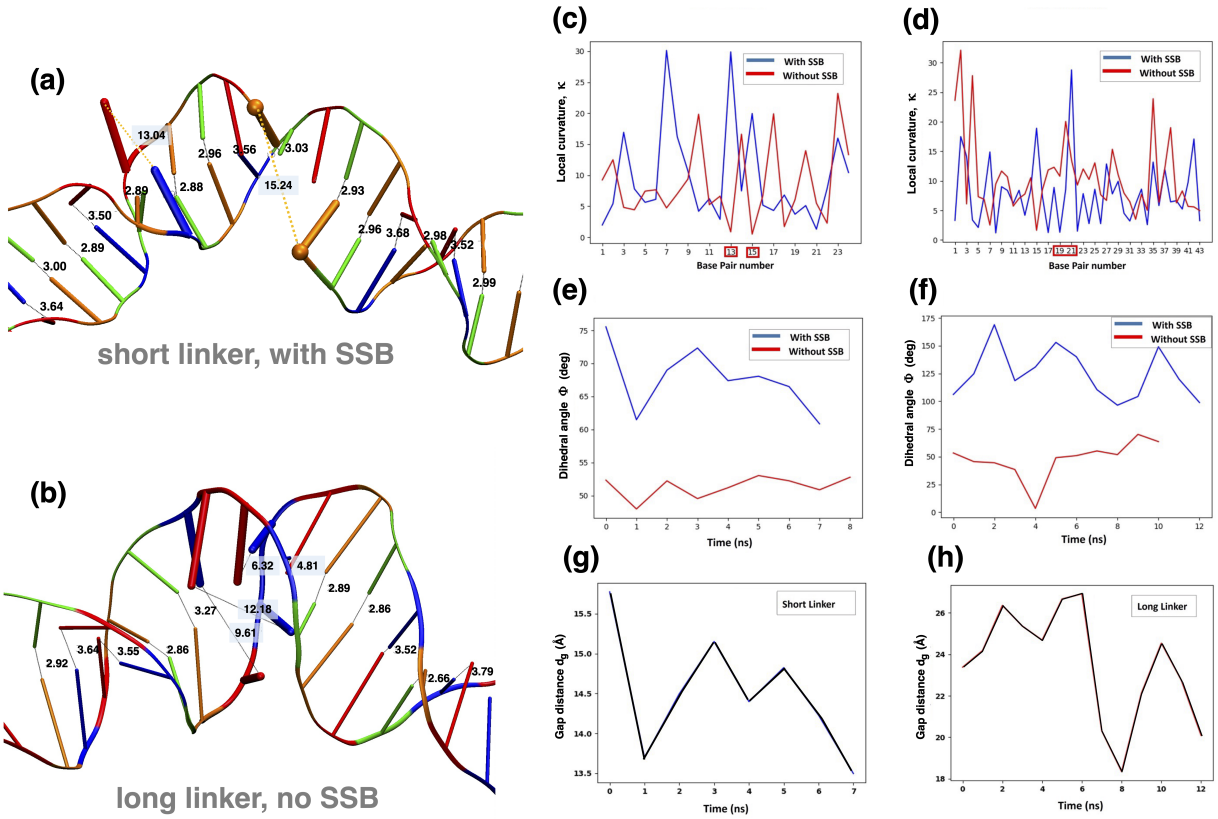


Figure 31: **(a)** Cartoon snapshot of the short-linker DNA with the SSB. G·C base-pairs are in orange-green, A·T base-pairs in blue-red; numbers indicate the N1-N3 atom distance in each bp. Note the flipped-out A·T bp (thicker sticks), with a N1-N3 distance of 13.04 Å; the SSB (thicker sticks) is flanked by two guanines, with spheres indicating the C4′ atoms and a d_g=15.25 Å. **(b)** Snapshot of the long-linker DNA without SSB. Note the large N1-N3 distances in the sequence ds(TAAT), with the central A·T bp (thicker sticks) flipped out in extrahelical positions. **(c,d)** Plot of the local curvature along the bp sequence for the short (c) and long (d) linker with vs. without SSB, calculated from Eq.(4). Red squares on the x-axis indicate the location of the SSB for the blue plots. **(e,f)** Time evolution of the dihedral angle Φ at the location of the SSB site for the short (e) and long (f) linker with vs. without SSB, during a short MD trajectory. **(g,h)** Time evolution of the SSB gap opening distance d_g for the short (g) and long (h) linker.

Localized damage in the DNA acts as a mechanical stress-concentrator

date the strong compression resulting from the bending (see Figure 31b, where the purine-pyrimidine N1-N3 distances are reported, the values for the ds(TAAT) sequence are highlighted). It may be counter-intuitive that the various inter-bp parameters in the blue plots display minor variations compared to the no-SSB plots: however, this shows that the sharper "Brazier-like" kinking at the SSB site in this case, very effectively relieves the excess stress, the two stretches of DNA flanking the SSB remaining rather straight. Such a behavior is entirely similar to what observed in Ref.[141], where sharp kinking concentrates both the stress and excess elastic energy at the damage spot, while twist-bending distributes the stress more smoothly about the zones of maximum curvature.

From the Tilt and Roll parameters, the local curvature between each two adjacent bp can be obtained, as [141]:

$$\kappa = \sqrt{\text{Tilt}^2 + \text{Roll}^2} \tag{4}$$

Note that this local value is not necessarily indicative of the overall curvature of the entire DNA strand, which is measured by the radius of curvature of the whole linker length. Rather, this quantity highlights the local bending caused by the two base-pairs stacked one above the other. The curvature plots in Figure 31c,d derived from tilt and roll parameters, further underscore the localized structural instability at the break site, with sharp peaks in regions surrounding the SSB. The two nucleotides flanking on either side of the break site (marked by red box in plots) show a very large value compared to the 'no-SSB' counterpart.

This localized deformation is another signature of the SSB gap opening and local twist in the DNA at the break site. Therefore, I computed the d_g and Φ values as already defined in the previous Chapters 2 and 3. The time-dependent plots in Fig.31e,f, show that linkers with SSB exhibits more significant variability and fluctuation in dihedral angles, reflecting enhanced flexibility and reduced structural stability near the break site. Conversely, DNA without SSBs shows smoother and more stable dynamics, indicative of more rigid base-pair stacking. The quite larger value of over-twisting observed for the longer linker also implies that the gap d_g tends to open to larger values, between 24-26 Å, while they remain around 14-15 Å for the shorter linker (Fig.31)g,h). This concurs with results observed in Chapter 3 were a positive correlation was observed between d_g and Φ

Overall, it is concluded that the presence of SSB damage in the linker-DNA brings about mechanical instabilities even under small forces applied to the system. Such bending and kinking causes the damage site to become more exposed, thereby making the damaged region more accessible to repair machinery. However, it is notable that also the undamaged DNA under forced compression may undergo extrahelical flipping of the base-pairs situated near the regions of maximal curvature.

4.3 Comparison of DNA linker deformation with vs. without a DSB

Given the interesting results of the previous Section 4.2, I now turned to the next question of whether and how the presence of DSB can mechanically deform the dinucleosome. To this purpose, I replaced the SSB with a DSB in the two configurations with different lengths of linker DNA (see Figure 32). Unlike most of the recent experimental works on mechanical deformation on nucleosome/chromatin, which

identify a DSB as the blunt end of the dsDNA at the entry/exit sites of the mononucleosome, here I constructed a proper DSB by introducing two SSB in the opposite strands of the linker DNA, separated by the 4 intact base-pairs of the ds(TAAT) sequence. More properly, this is the configuration prior to the eventual formation of a two-ended DSB breaking apart of the DNA in two fragments with hanging unpaired nucleotides (the sticky ends) .

The two systems (short and long linker) were then simulated for around 150 ns MD. As in the case of SSBs, initially a series of unbiased-MD simulation (i.e., without any applied force) were performed on each setup. Interestingly, already in these unbiased simulations a rather pronounced bending of the linker DNA was observed, making the nucleosomes coming close together spontaneously, as shown in the Figure 32. As it will be seen, this relative movement of the nucleosomes towards each other helped the DSB to be exposed, by increasing the gaps at each SSB termini; also, this brought about a consistent mechanical deformation in the linker DNA, also promoting extrahelical base flipping. Hence, the application of extra forces being unnecessary, this section deals with understanding how these unbiased-MD simulations let the dinucleosome system evolve spontaneously into a closed V-shaped configuration.

To extract the representative structure of the system, clustering analysis using the GROMOS method in GROMACS 2023.3 package [182] was performed. The best clusters obtained by averaging over the final part of the MD trajectory are shown in the figure 32a,b. Each setup, regardless of the linker length between the nucleosomes, underwent a smooth transition from a straight to a bent conformation, bringing the nucleosomes closer together. This bending propensity is more pronounced in the long-linker variation and ends up in a visible kink at one of the two SSB making up the DSB, while a smaller degree of bending is observed in the short-length linker.

The final stability of the system that reaches a lower-energy state upon approaching, is shown both in the RMSD plot (Figure 32c) and COM plot (Figure 32d). After about 70 ns of MD simulation, the RMSD of the two nucleosomes stabilized at ~3 and ~6 nm for short and long linker, respectively. The COM-distance also reduce with a similar trend, for both the short and long linker, respectively decreasing from 16 to 12 nm, and from 22 to 9 nm within the first 75 ns of the trajectory; after which the distance remains stable. Altogether, this initial analysis shows that the mere presence of a DSB can deform the linker DNA, and bring the individual nucleosomes closer together, to a configuration where the system is more stable compared to the configuration with fully-extended damaged linker.

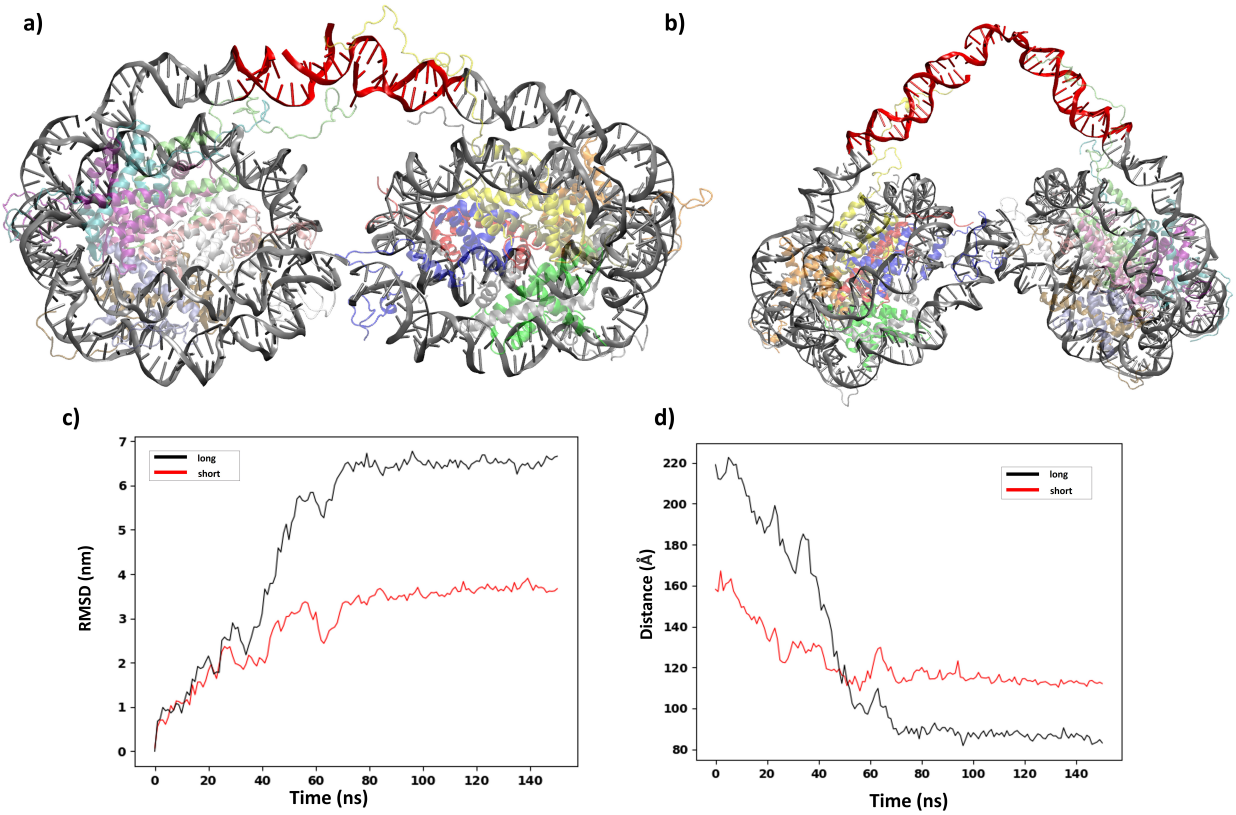


Figure 32: (a,b) correspond to the best cluster obtained after 150-ns MD simulation for short- and long-linker variants respectively (the red-colored part is the linker DNA). Note that the two nucleosomes are not touching (not even their histone tails) despite the particular projection of the figure seems to show them in contact. (c) RMSD trajectory over 150-ns for the two variants. (d) Distance between the COM of the nucleosome pair.

4.3.1 *Mechanical deformation due to DSB and its significance*

Both linker-length variants produced a bending/kinking of the DNA in the presence of a DSB. This happens gradually over the first 70 ns and then the system settles about a constant conformation, which could be due to the steric hindrance of two big nucleosomes, now very close to each other. It is also noteworthy to mention that the nucleosomes do not interact with each other, neither the histone tails nor the ends of the DNA (despite the snapshots in Figure 32a,b seems to show a contact, this is only due to the graphical projection).

Unlike the case of SSB under steered-MD, in which the the force-pulling of the nucleosomes resulted in transitioning from bend to twist-bend and kinking as seen in Figure 29c, with the DSB the transition is a very smooth Euler-bending to a sharp kink in the linker DNA. This is clearly seen in the snapshots of Figure 33a,b, and in the plot of Figure 33c: the long-linker DNA approaches a bending angle of around 100° by 50 ns, while the short-linker approaches 120°. This bending of the DNA is accompanied by the structural deformation

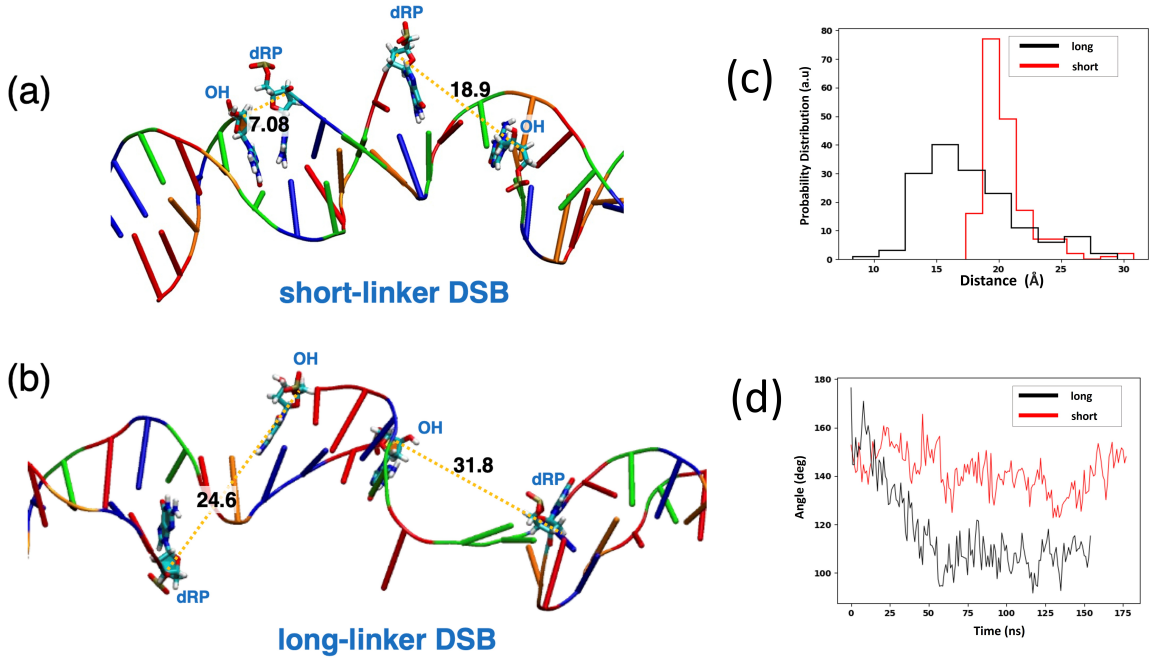


Figure 33: **(a,b)** Representative snapshots of the central portion of the sequence around the DSB, in the short and long linker, respectively. The 5′-dRP and 3′-OH terminations of the two SSB making up the DSB are indicated in licorice. The gap opening distance is measured, as usual, between the C4′ atoms of each base flanking the break. **(c)** Time evolution of the bending angle between the two nucleosomes. **(d)** Probability histograms for the gap-opening distance d_g, averaged over the 150-ns MD trajectory.

of DSB as well. The unpaired nucleotides at each SSB making up the DSB become exposed by the widening of the gap between each pair of $5' - 3'$ termini. Notably, in both linkers one of the SSBs opens up much more than the other; the gap distance d_g for the end terminals increases up to >30Å, as shown by the probability histograms of the d_g values in Figure 33d; such very large values are similar to the extreme opening distance observed when the hairpin-DNA (Chapter 2) bends upon interacting with PARP-1 protein.

Also in this case, the CURVES+ analysis allows to pinpoint some details of the level of deformation in the presence of DSB. The two plates in Figure A13 and A14 in Appendix A, display the ensemble of the calculated helical parameters for both linker configurations, on the best cluster extracted from the second part stable of the MD trajectory. The two SSB making up each DSB are located at positions 12-17, and 19-23, for the short- and long-linker, respectively. Large peaks are observed in all cases for the intra-bp parameters (notably, Shear, Stretch, Stagger and Opening): they nearly always correspond to the terminal nucleotides flipping out of the DNA helix, with ample oscillations. Also the inter-bp parameters clearly locate the maximum deformation at break sites, even much more clearly than in the case

of the isolated SSB in Section 4.2 above; in particular, the huge peaks in Roll and Tilt are indicative of an extreme local curvature at the damage sites.

Unfortunately, time constraints did not allow to complete this Chapter with the interaction between PARP-1 and the damage in the dinucleosome linker. Also, it would be extremely interesting to extend this study by comparing with DSB configurations after the full break, which would separate the linker DNA in two hanging free-ends. We will certainly develop this part of the study at later times.

4.4 SUMMARY AND OUTLOOK

To summarize the work carried out in the present Chapter 4, all-atom MD simulation of dinucleosomes with varying linker lengths, and including either a SSB or DSB in the linked DNA were performed. This study is among the first to demonstrate that complex defects within the linker DNA can independently induce significant structural alterations in the nucleosomal architecture. The key take-away points that are crucial to this study include:

1. Applying even a small amount of force to each of the nucleosomes, including a SSB in the linker DNA, brought the nucleosomes close to each other, with a twist-bend, or a sharp kink (for the longer linker) developing in the break site. The same forces applied to damage-free systems could not bring pronounced mechanical deformations.

2. Comparing the extent of structural changes occurred within the linker DNA with a SSB, a minimum length of the linker DNA of more than about 24-bp seems necessary for the mechanical deformation to develop.

3. Introduction of a DSB in systems with similar linker-DNA lengths shows that just with the thermal fluctuations and without additional forces, linker DNA underwent a dramatic bending and kinking at the defect site; very large opening of the breaks is observed, asymmetrically (one SSB opens much more than the other).

These points help us to conclude that with more complex damage, pronounced bending and structural deformation can be expected in the chromatin. Simpler damage like SSBs require some external forces to activate the mechanical deformation. These internal forces, in the 10-20 pN range, can be readily contributed by the cellular/nuclear environment via mechano-transduction, e.g. through the LINC complex [183], to bring about similar, localized conformational changes within the chromatin. From a biological perspective, this finding suggests that damage-sensing proteins, such as Ku70/80, UDG, or , may be

more readily directed to these localized areas of structural instability, rather than requiring an exhaustive random search of the entire chromatin. The inherent changes in the damaged linker DNA could act as signals or markers, simplifying the damage recognition and repair processes. Also, the role of the mechanical stress and elastic energy redistribution around the damage should be explored, however the time constrains did not allow to pursue also along this interesting direction.

Overall, these results raise intriguing questions about the potential role of multiple defects within the nucleosome or linker DNA in driving large-scale chromatin remodeling events. If just a single damage can induce such pronounced bending and deformation, it is conceivable that clusters of damage may trigger yet more extensive structural reorganization of the chromatin. This could impact processes like gene expression, replication, or DNA repair and potentially contribute to cellular responses such as senescence or apoptosis. While the exact mechanisms underlying such chromatin remodeling will require experimental verification, this computational study highlights the need for further investigation into how damage at the molecular level can influence chromatin dynamics and higher-order chromatin organization. Understanding these processes may not only shed light on fundamental biological mechanisms but also provide insights into developing targeted therapeutic strategies for diseases linked to defective DNA repair, such as cancer.

5

MATHEMATICAL MODELING AND MONTE CARLO SIMULATION OF THE DNA DAMAGE REPAIR PATHWAYS

5.1 INTRODUCTION

Biological systems are impervious to rigorous quantitative analysis and mathematical approaches, being composed of interconnected systems that result in complicated, non-linear dynamics between the system components [184]. Cells live in a complex environment, and can sense many different signals including physical, chemical or biological parameters. The response to such signals is the production and activation of appropriate proteins, which perform actions on the internal state or the external environment of the cell. Some messages result in the activation of transcription factors that bind to DNA to regulate the rate at which specific target genes are read (i.e. trascribed into mRNA then translated into proteins). For example, *E. coli* works with an internal representation of about 300 transcription factors that regulate the activity of its about 4,000 different genes. Over the last half of the XXth century, the molecular biology community has continued to uncover the details of this molecular domain. The painstaking effort involved in these nano-scale experiments necessitated a so-called 'reductionist' approach, in which research projects often addressed individual molecules or molecular interactions.

As experimental techniques progress and provide more and more complete picture, the ensemble of interactions between transcription factors and genes can be better summarized and represented by genetic, or more generally *biological system networks*. This integrated vision of the cell activity has given rise, at the beginning of the new century, to the novel domain of **systems biology** [185, 186], which aims at identifying general biological principles, possibly underlying the complex networks of thousands of interactions. The search for these principles is far from complete, but it is made possible by advances in experimental technologies (the so-called 'high-throughput', and 'post-genomic' approaches) that provide a vast amount of elementary data about the individual components of the networks and their interactions.

Such a more quantitative approach to molecular biology allows network diagrams to be extended to mechanistic mathematical models. These models serve as working hypotheses: they help us to understand and predict the behavior of complex systems. Reliance on computational models is a key character of modern systems biology: the

Human Genome Project [187] could not have been completed without advances in bioinformatics that allowed the processing and interpretation of vast amounts of sequencing data [188]. A complementary mathematical approach to systems biology is the investigation of intracellular processes as *dynamic systems*: the corresponding models essentially mimic the behavior of intracellular networks, albeit by making a number of simplifications [186].

The mathematical modeling of a biologic dynamical process may be broken down into three main steps: 1) conceptualization; 2) mathematical formalization; and 3) management and optimization. There are many approaches to conceptualize a biological network, including differential equations, Bayesian equations, stochastic systems, agent-based modeling, etc [189]. The choice of the method depends on the hypothesis/concept that has been put forward. Mathematical modeling can have a predictive capability, which helps experimental researchers to simulate a particular metabolic pathway, or an ensemble of connected pathways under different conditions. This can be exploited in every field of experimental research, and has recently led to the development of powerful tools to build mathematical models, and formulate and validate the biological hypotheses.

One of the most utilized type of mathematical modeling is through systems of **ordinary differential equations** (ODE), with either a top-down, or a bottom-up approach [190]. In this approach, the concentrations of each chemical species are followed in time by a dedicated differential equation; the equations form an algebraic system in which the rates of production and consumption of the different species are mutually linked. If spatial diffusion of the chemicals is also taken into account, such as in a multicellular model [191], the concentrations can also be made dependent on the spatial variables, e.g. in 2D or 3D, and turn into partial differential equations (PDE). Solutions are usually obtained by discretization over Eulerian grids in space, and typically use forward algorithms for the time-dependent solutions (e.g., Runge-Kutta, predictor-corrector etc.).

Monte Carlo methods (MC) are an example of stochastic simulations. MC is a well-established ensemble of techniques to carry out various computational studies, with many applications also to biological problems [192, 193] . All the MC methods rely on repeated random sampling of probability distributions of the independent variables, to estimate numerical results by generating large ensembles of possible outcomes based on predefined rules. The method enables the study of distributions, rare events, and uncertainties within complex biological phenomena. The approach typically selects an initial configuration state and iterates through one or more steps, exploiting a great flexibility and efficiency in sampling configurations. This property makes it suitable for modeling biological events such as DNA damage-repair pathways, metabolism of proteins, molecules etc.

Building on the discussions in previous Chapters, which explored the microscopic interactions between DNA damage and repair proteins, highlighting how these interactions can induce dynamic changes within DNA, the nucleosome and chromatin, this final Chapter shifts the focus to a comprehensive analysis of the single-strand break repair (SSBR) pathway. In the first Chapter, the connection between DNA damage, cancer, and cellular senescence was discussed, and the basic ingredients of the damage-repair pathways were laid out. Extending that perspective, it is essential now to investigate how the out-of-field radiation dose can induce additional single-strand breaks (SSBs) in the margins of the treatment volume, where normal cells lie, as demonstrated in [13]. As described in Chapter 1, these additional non-repaired SSBs can lead to cellular senescence and, over time, potentially result in secondary sarcoma.

To understand the decision-making processes underpinning this phenomenon, two complementary approaches were adopted in the following:

1. Differential-equations Modeling: I employed different kinetic models in COPASI software, to represent the SSBR pathway under out-of-field radiation exposure, and briefly explain how one can utilize this tool to understand the cell decision-making process.

2. Monte Carlo Simulations: using the GEANT4-TOPAS Monte Carlo framework, I initiated a preliminary setup for modeling the SSBR pathway. Since the current framework in TOPAS lacked the SSBR pathway model, my work focused on establishing the foundational aspects to include the SSBR pathway, which could be expanded in future studies.

Through these approaches, this Chapter aims to provide a more general understanding of the complex interplay between radiation-induced DNA damage and repair pathways, with implications for cellular responses.

5.2 Mathematical modeling of the SSBR pathway triggered by out-of-field radiation.

Both the SSBR and BER pathways have been mathematically modeled using a deterministic approach by various groups [194–196]. In this Chapter I followed a straightforward approach, in which the entire cascade pathway triggered due to out-of-field dose radiation was mathematically modeled using a system of ODEs, reproducing for the various chemical species either the Hill kinetics, Michaelis-Menten kinetics, or mass-action kinetics (the detailed explanation for each kinetics model is given in Appendix B section B.4.1).

This entire ODE–system was developed in COPASI. COPASI or COmplex PAthway SImulator [197], is a free software package for modeling and simulation of simple to complex biochemical pathways. The software has pre-built numerical algorithms, to perform time-course simulation, parameter estimation, stoichiometric analysis, stability analysis, and stochastic simulation. COPASI software allows the user to tweak tiny details, e.g., initial concentrations, reaction conditions, reaction rates etc. The major step is initializing all the transition from one state to another in form of simplified relationships

$$\mathrm{reactants} \rightarrow \mathrm{products}$$

as depicted in Figure 34b. For each relationship, this is followed by adding the reaction kinetics, represented by an ODE, as in Figure 34c, and reaction coordinates; explicit values for the rate-constants and initial concentrations of each species can be fixed, either to study arbitrary conditions, or to reproduce an experimental set up. As a result, features such as parameter estimation, time-course simulation, and sensitivity analysis allow the user to fit experimental data and estimate reaction constants effectively. Appendix section B.4 briefly gives an overview about the software.

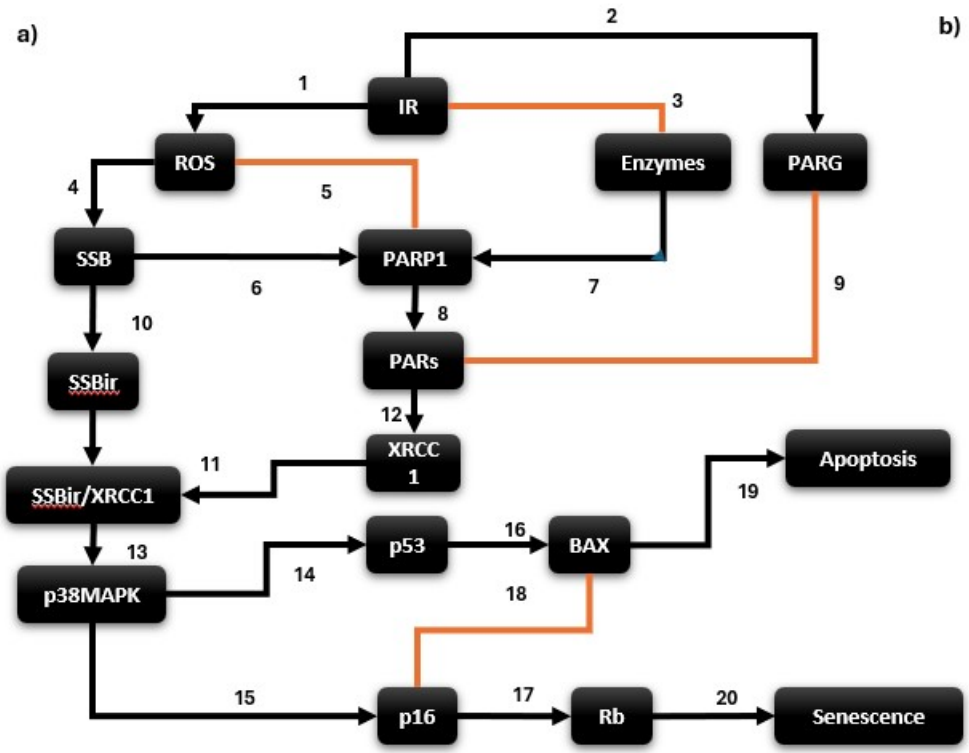
The workflow of the SSBR pathway was already described in detail in Chapter 1, Section 1.5. The flowchart depicted in Figure 34a simplifies the entire cascade pathway: each transition from one reaction complex to another, is represented by a reaction kinetics and reaction rates (k_n). Arbitrary values of reaction coordinates were chosen in practice, for those parameters for which not enough experimental data are available to validate the model. At this stage, the complete model set-up and the corresponding time-course analysis provide only theoretical results; however, they represent a solid foundation to explore the SSBR pathway once enough microscopic data will become available.

5.2.1 *COPASI simulation of the SSBR pathway*

We conducted several COPASI time-course simulations to explore how reaction kinetics influence each factor over time. As illustrated in Figure 34a, the pathway begins with a radiation dose imparted to the cell (**IR**), leading to the production of reactive oxygen species (**ROS**). These species induce single-strand breaks (SSBs) at a defined rate, some of which are recognized by **PARP-1**, initiating the SSBR repair process. However, a fraction of SSBs may remain unrepaired (**SSBir**) and instead activate the **p38MAPK** path toward either apoptosis or cell senescence end-points.

The rates of reaction for different species were simulated over 100 minutes with an initial radiation dose of 2 Gy, for which Figure 35 shows a set of typical results for the concentrations of the various

a)



b)

No:	Reaction	Reaction kinetics
1	IR -> ROS	Mass Action (irreversible)
2	IR -> PARG	Mass Action (irreversible)
3	IR -> Enz; IR	Hills Inhibition
4	ROS -> SSB	Mass Action (irreversible)
5	ROS -> PARP1	Hills Inhibition
6	SSB -> PARP1; PARP1	Complex
7	Enz -> PARP1	Hills Activation
8	PARP1 -> PARS	Hills Activation
9	PARG -> PARS	Hills Inhibition
10	SSB -> SSBir	Mass Action (irreversible)
11	SSBir + XRCC1 -> SSBirXRCC1; XRCC1	Complex
12	PARS -> XRCC1	Mass Action (irreversible)
13	SSBirXRCC1 -> p38MAPK	Hills Activation
14	p38MAPK -> p53	Mass Action (irreversible)
15	p38MAPK -> p16	Mass Action (irreversible)
16	p53 -> BAX	Mass Action (irreversible)
17	p16 -> Rb	Mass Action (irreversible)
18	p16 -> BAX	Inhibition
19	BAX -> Apoptosis	Mass Action (irreversible)
20	Rb -> Senescence	Mass Action (irreversible)

c) $$\frac{d[p38MAPK]}{dt} = \left(k_{13} * \frac{[SSBir_{XRCC1}]}{(1 + SSBir_{XCCC1}]}\right) - (k_{14} + k_{15}) * [p38MAPK]$$

Figure 34: **(a)** The final COPASI scheme of the SSBR cascade pathway for the cells in the margin, where red arrows show the inhibition/down-regulation, and black arrows represent the transition from one state to another. **(b)** The table represents the different reactions and the kinetics each of them follow. **(c)** The example differential equation represents the reaction kinetics for p38MAPK. All the equations for the rest of the species can be found in Appendix B, section B.4.2.

species. Our analysis revealed that the increase of concentration of ROS over time contributes to downregulating the expression of PARP-1 in the system. [ROS] reaches an equilibrium point (by ~30 minutes in Figure 35a, according to the chosen reaction-rate parameters). But, once [ROS] reach this point, the presence of increasing SSBs stimulates PARP-1 production in the system (Figure 35b). Now, in the model a fraction of SSBs is supposed to remain unrepaired (Figure 35e). The biological hypothesis behind such unrepaired damage (see also Chapter 1, Section 1.5), is that for different reasons (e.g., impoverishment of the NAD^+ substrate, that is the "Enzymes" block in the flowchart) PARP-1 starts being down-regulated, and produces non-functional (too few, and/or too short) PAR-chains (Figure 35d), whose concentration peaks before [ROS] saturation, at ~20 minutes; this should hamper recruitment of CK2 to the damage site, and the consequent phosphorylation of XRCC1. In such conditions, the SSB repair may remain incomplete, or ineffective (see e.g. [118]). (Note that in the model only the unphosphorylated fraction of XRCC1 is being followed, assuming that the "normal" functioning of the pathway leads to properly-repaired SSBs, which are excluded from the study.)

The concentration of unphosphorylated-XRCC1 should theoretically reduce, once the SSBir/XRCC1 foci are being formed. However, the actual concentration evolution depends on the actual rate of production of unrepaired-SSB, (which keeps increasing up to a maximum, till t ~40 minutes) and hence the rate of association of the SSBir/XRCC1 complex. In this case, the production rate of XRCC1 is sufficiently high, that both effects combined only leads to a slower synthesis of XRCC1 (Figure 35c).

The formation of SSBir/XRCC1 foci triggers the p38MAPK pathway, which drives a critical decision between activating p53 or p16 paths [198]. Its concentration (Figure35f) peaks at about t ~30 minutes, that is the approximate saturation time of [ROS] and before the peak in [SSBir]. This critical decision-making step depends on (1) the reaction rate for choosing either one of the paths, and (2) the reaction rate of p16 inhibiting BAX. Playing with the reaction parameters highlights the competition between the two paths, and the corresponding end-points. For instance, keeping the same reaction rate for both p16 and p53 the outcome is nearly equally split, with just a slight increase in the cells choosing p16 pathway (this is because of the inhibitory action of p16 towards BAX, Figure35g). Upon reducing, p16 reaction rate to 10% of p53 (Figure 35h) we see increasing [BAX] (less inhibition) and p53 becomes slightly favored with respect to p16; if we assume p16 reduced to 50% of the initial value (and keeping the same inhibitory reaction rate p16$\neg$BAX), p53 is being chosen by the model over p16 (Figure 35h); we verified that for a large imbalance between p53/p16, even a maximum inhibition rate of [BAX] is not enough to make p16 prevail. The behavior in 35g is roughly consistent with

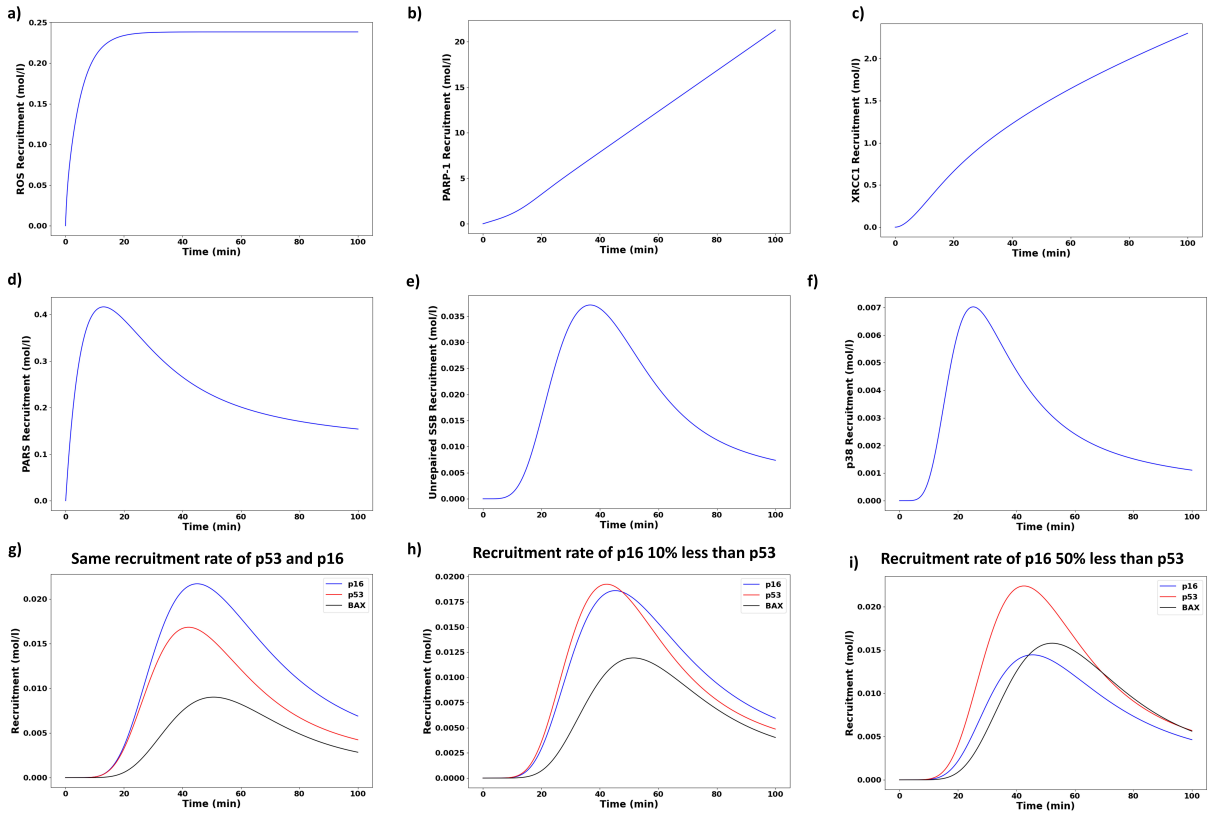


Figure 35: Example of COPASI results for a time-course simulation of 100 minutes, with arbitrary (but reasonable!) choice of the free parameters. **(a-f)** Time-evolution of the concentrations of, respectively, ROS (from input radiation dose), PARP-1, XRCC1, PAR-chains (also PARS), unrepaired-SSB (SSBir), p38MAPK recruited to the system. **(g,h,i)** Concentration of p16, p53 and BAX recruited to the system, at different settings of the respective reaction rates. See text for details.

experimental results, particularly in fibroblasts, where cells tend to favor the p16→Rb senescence pathway over apoptosis. Unless the reaction rate for the p16 path goes to zero, the cells will always have chance to choose the senescent pathway. This is definitely to be attributed to the inhibitory effect of p16→BAX⊣ p53 on the apoptotic pathway, which will have to be assayed in biological experiments.

Furthermore, we found that the choice of kinetic modeling equations significantly impacted the outcomes. When mass-action kinetics was used instead of Hill's equation, the rates of reaction species were altered, and cells could show a preference for the apoptotic pathway rather than the senescence pathway, contrary to experimental observations. Additionally, our analysis highlighted the sensitivity of the model to parameter values. For this study, arbitrary parameter values were used due to the lack of experimental data, but their influence on the results highlights the importance of accurate parameter estimation in future work.

5.3 MONTE-CARLO SIMULATION OF THE SSB REPAIR PATHWAY

The previous Section provided a deterministic approach to depict how cell decisions can be affected, by tuning the concentrations of each protein species, and the corresponding reaction rates in the pathway. However, the lack of experimental data to validate the model still poses a major limitation to get meaningful results, and only very qualitative conclusions can be drawn. On the other hand, when experimental data are available, the mathematical modeling can fully demonstrate its valuable contribution to understanding complex phenomenology. This is the case of the work I carried out in adapting the SSB repair model to the GEANT4-TOPAS simulation environment, for which a large amount of experimental data about PTV irradiation with energetic particles is available. In this framework, we simulate by Monte Carlo techniques a beam of ionizing particles (protons) through a computational model of the cell, to explore the damage pattern, and study the repair pathways being triggered. GEANT4-TOPAS tool with TOPAS-nBio extension [199, 200] was used for this purpose. The Monte-Carlo simulation kit, which already included a complete description of the DSB repair, was used to quantify the DNA damage introduced in the chromatin model after proton irradiation. During the time I spent at Manchester University, in the PRECISE group, my original contribution was to model the entire SSBR pathway triggered by the SSBs and include it in the complete simulation software. The upcoming section will introduce the tools used, outline the model framework, and discuss the results along with the future perspectives of this study.

5.3.1 *Introduction to **TO**ol for **PA**rticle **S**imulation (TOPAS)*

Improvements to radiation therapy and medical imaging techniques can be made by understanding the path of particle beam through the equipment and tissue. These precise calculations are carried out in Monte Carlo (MC) simulations. Monte Carlo simulation is a mathematical technique that helps find the possible outcomes of an uncertain event. It is a type of algorithm that utilizes repeated random sampling between fixed values to produce a large number of likely outcomes, with event likelihood weighted by some prior measurement. For radiotherapy physics, MC simulations is used to model dose deposition through track structure simulation. This has become the golden standard and is frequently used as a comparator to analytical dose models.[201]. They provide a comprehensive description of the physical and chemical transport of radiation through matter on an event-by-event basis. Hence, these could provide details that are challenging to measure during *in vitro* experiments [202]. MC-track simulation codes have been prevalent in the scientific and medical fields

for decades to understand the physical-chemical effects of ionizing radiation at the subcellular level. Examples of such codes are MOCA8b [203], DBREAK [204], PARTRAC [205], KURBUK [206], Geant4-DNA [207, 208], RITRACKS [209], DaMaRiS [200], and TOPAS-nBio [199]. The main goal of these codes is to simulate DNA damage under different radiation conditions, thereby examining the nature of DNA strand breaks and the subsequent DNA repair. This thesis chapter uses the TOPAS-nBio framework to simulate the induction of DNA damage (SSB and BD) and the mathematical modelling of DNA Damage response through the SSBR pathways via extending the DaMaRiS model.

The TOol for PArticle Simulation (TOPAS) is an extension of Geant4 simulation toolkit, to make advanced MC simulation in radiotherapy. They are designed to be user friendly so that the user requires no advanced programming expertise. TOPAS can simulate x-ray and particle therapy treatment heads, reconstruct patient geometry from CT images, and score parameters such as dose and fluence. It can also save and replay phase space data to further interrogate a particular "event" or interaction. It offers advanced graphical capabilities, operating in a four-dimensional (4D) framework to both observe a treatment and to account for temporal variations in beam delivery or changes in patient geometry during treatment. One of the main feature of TOPAS is its parameter control system. Users can input parameters in one or more text files to control simulation properties. Further to text file input, the user can write bespoke C++ code to interact with the TOPAS engine at any point in the simulation. This versatility offers the 'expert' user to extend capabilities in a convenient way to modify TOPAS for their particular needs, e.g. custom scorers, advanced geometry, physical interactions etc. Examples of parameter specifics can be extensively found in the documentation by TOPAS as well as by Perl. et al. [https://www.topasmc.org/] [210]. Along with dose parameters, TOPAS collaboration have also incorporated linear energy transfer (LET) scoring, relative biological effectiveness (RBE) models for proton therapy, and outcome models like tumor control probability (TCP) and normal tissue complication probability (NTCP) [211, 212].

TOPAS has been effectively utilized in research on radiation therapy physics and macroscopic biological systems, including organs and cells. However, advancing our understanding of radiation effects requires more fundamental research to explore mechanisms such as oxygenation, intracellular signalling, and drug-induced radiation sensitization or resistance. Additionally, variations at cellular and subcellular levels in both tumors and surrounding normal tissues must be accounted for, highlighting the need for detailed in silico simulations. Hence, the TOPAS-nBio extension aims to combine detailed physics and chemistry simulations with realistic representations of bi-

ological structures, such as cells and nuclei, thereby enabling a mechanistic link between sub-cellular energy deposition and observable biological outcomes. It was developed as an extension of TOPAS and designed to simulate radiobiological experiments at nanometer resolution by utilising advanced physical interaction models down to the very low energy relevant to the DNA scale. It retains TOPAS's user-friendly design and reliability while offering advanced Monte Carlo simulations at the nano-scale. It makes sophisticated computational tools accessible to researchers seeking to enhance their experimental designs or data analyses with more detailed physical, chemical, and biological modelling [199].

5.3.2 *DNA Mechanistic Repair Simulator (DAMARIS) for SSB repair*

The DNA Mechanistic Repair Simulator (DaMaRiS) framework, developed at the University of Manchester [200], served as the in silico platform for modeling in this study. Built using the Geant4-DNA Monte Carlo simulation toolkit (version 10.4), DaMaRiS (version 0.3) requires input data on the distribution of DSB and SSB damage following irradiation. For this study, we incorporated the chromosome structure in the form of Hi-C data to Monte Carlo radiation track structure simulations of DNA damage [213], producing an output in the Standard DNA Damage (SDD) Data Format [214]. Designed for versatility, DaMaRiS allows users to define, construct, and test DNA repair models, with each parameter having a clear biological definition. This mechanistic approach enables modification of the models by incorporating or removing biological mechanisms as they are discovered, allowing the framework to represent entire repair pathways, which primarily include NHEJ and HR pathway for repairing DSBs[215]. The versatile approach also allows the user to test competing theories of DNA repair pathways. Several versions of a pathway can be simulated and compared to experimental data to uncover the "most likely" pathway [216].

For details on Hi-C and SDD format, please check Appendix B.

Each step in the model reflects how proteins interact to process DNA breaks, from recognition to repair completion, providing temporal insights into how damage is resolved within repair pathways. The DaMaRiS models in this study were extended to include BER and the initial stages of SSBR, with plans to expand the framework to capture the full DNA damage response (DDR) at the molecular level. The framework also examines critical points where pathway choices are made—where DNA repair may transition between pathways—and evaluates alternative pathways that may be activated in cases of repair failure [215].The basic algorithm used in this study is structurally similar to that of the algorithm utilized by Ingram et.al.; in their study analyzing the DSB repair pathway. The schematic below illustrates the steps followed.

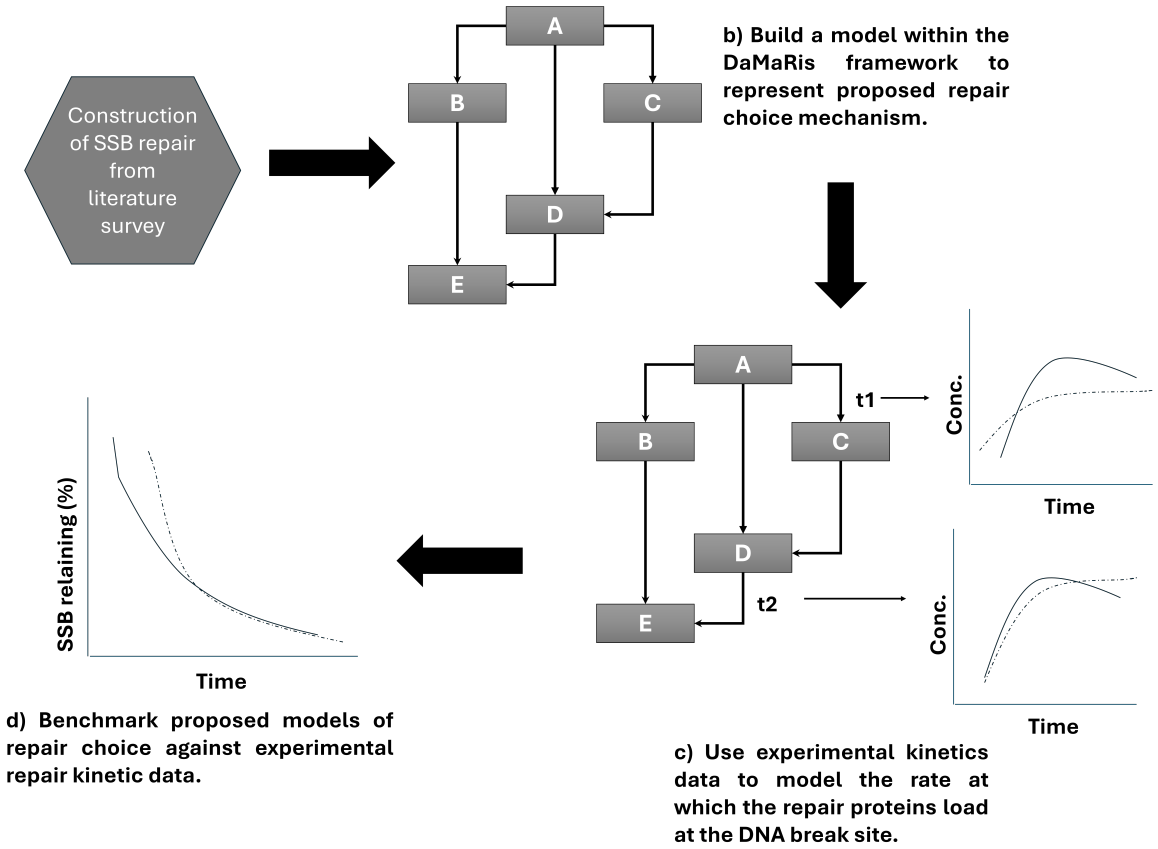


Figure 36: **(a)** A comprehensive literature review was conducted to identify potential mechanisms influencing DNA repair pathway choice. **(b)** These mechanisms were integrated to construct coherent scenarios describing the decision-making process for repairing DNA double-strand breaks. The proposed scenarios, along with their associated protein recruitment steps, were modeled using the DaMaRiS framework. **(c)** The scenarios were parameterized to incorporate protein recruitment time constants, reflecting recruitment kinetics consistent with experimental data from the literature. **(d)** The parameterized models were then utilized to simulate DNA repair processes at the single-cell level following radiation-induced damage, enabling the measurement of overall repair kinetics for each scenario.

This part of my thesis adapts this algorithm to evaluate the DDR pathway. Hence it covers majorly two parts:

1. Quantifying the different type of DNA damages due to irradiation; which includes SSBs, DSBs, BD, and clustered lesions

2. Analyzing how these damages are being repaired by DNA damage repair pathway and validating with various experimental results; for this purpose the algorithm shown in Figure 33 was used. 36.

5.3.3 *SSB repair pathway model: Framework, Assumptions and Limitations*

DaMaRiS has an existing DSB repair pathway model, which can simulate the motion and reaction of DNA double strand ends produced when a DSB is created and the repair proteins. The repair model parses the data from the SDD format to obtain information about the position of 2 strand breaks of the two free ends of the DSB which are perceived as 2 separate objects by the model, further the repair proteins are recruited to the damage site where the DSB ends are processed [200, 215, 216]. Therefore to incorporate SSB inside the DNA damage model, the DNA damage model was altered to produce a SDD output file in such a way that all the isolated backbone damages are treated as SSB site which would be recognized by the damage model as an individual object. Unlike DSB ends, each SSB site is treated individually to recruit repair proteins at it's damage site. The motion of SSB are trapped in the space unless a user defined repair protein forms a complex with them, once formed they move together. The motion is defined by continuous time random walk (CTRW) diffusion model of sub-diffusive motion [217]. By definition, CTRW represents a waiting period for a random time before making jumps. Studies highlight that the damage ends tend to be confined in certain states before associating/dissociating with the protein complexes. [200, 217].

In DaMaRiS, the strand breaks progress through the user defined pathways by changing state using a step-by-step approach. For example, the SSB will transition into a strand break with a PARP-1 protein attached to it, representing the break site recognition. Similarly, each transition is accompanied by user-defined time constants, which represent the average time required to make that transitions, with actual transition time randomly selected. When SSB has multiple options of transitions, the model randomly chooses the transition time and shortest time out of those is likely to be selected. This random selection follows an exponential probability function, given by the equation $p = e^{\tau/t}$; where τ represents the time constant and t represents the post irradiation time.Flow through the pathway and branch selection of the damage can be modified by altering time constants, with an increased time constant reducing the chance of branch selection. To implement the SSB repair pathway, a review of experimental data from various studies detailing the recruitment of proteins involved in SSB repair was conducted. These mechanisms were incorporated into the DaMaRiS model, with time constants adjusted to align simulation outcomes with experimental observations. The framework of SSB repair pathway and the various time constants linking each state is represented in the Figure 37.

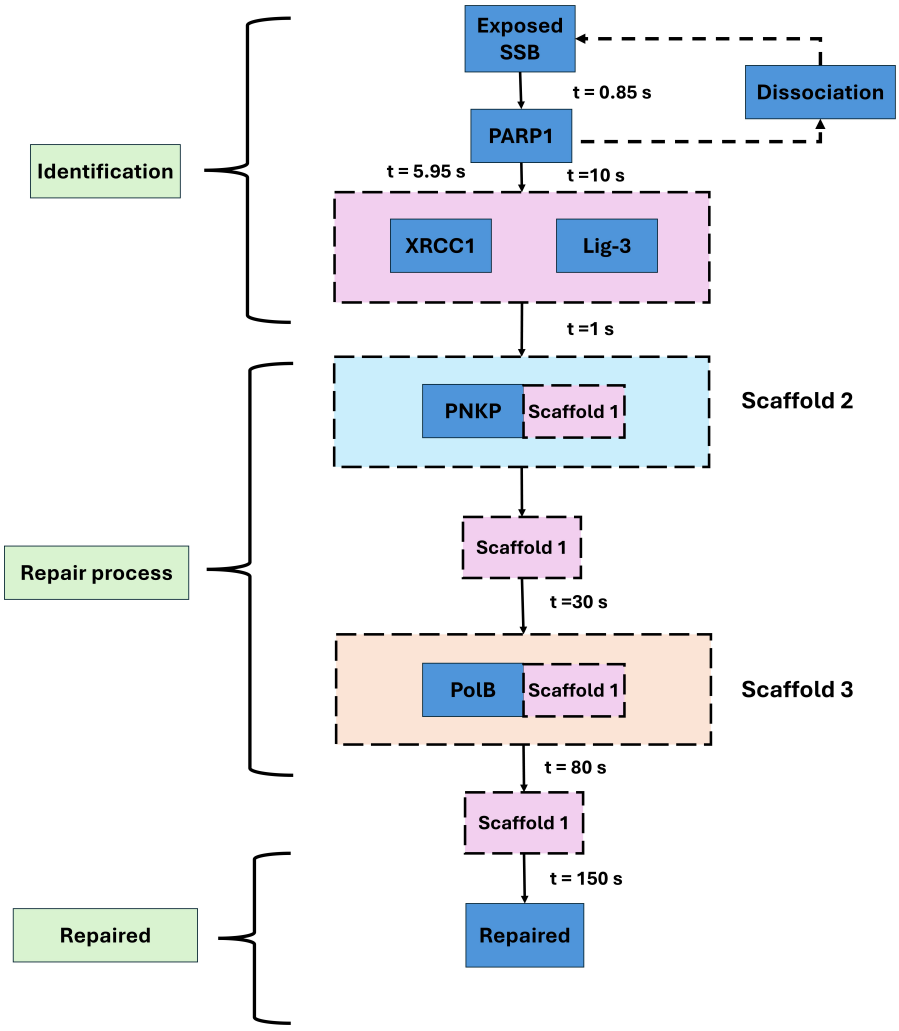


Figure 37: Schematic of the reduced SSBR pathway used in the Monte Carlo simulations along with the associated time constants.

The SSB repair pathway is divided into three distinct phases: Identification, Repair Process, and Repaired DNA. These phases ensure a systematic progression from recognizing the damage to fully repairing the DNA strand.The repair process begins with the identification of single-strand break (SSB) ends by the PARP-1 protein, which serves as the primary sensor of DNA damage. Upon binding to the SSB site, PARP-1 undergoes auto-PARylation, a process where poly(ADP-ribose) (PAR) chains are attached to itself. These PAR chains play a crucial role by facilitating the recruitment and phosphorylation of XRCC1, a scaffold protein essential for downstream repair processes. Simultaneously, the degradation of PAR chains by PARG ensures the recycling of PARP-1 along with detachment of proteins attached to the SSB, allowing it to bind to new SSB sites and maintain efficient repair across multiple damage locations. Experimental data, such as that provided by Mortusewicz et al. [218], quantifies the binding behavior of PARP-1 at SSB sites. The model assumes that once PARP-1

binds to the break site, it establishes a dynamic equilibrium with the DNA damage. This equilibrium is considered sufficiently stable that the dissociation of PARP-1 from the site can be neglected during the identification phase. After successfully marking the SSB, PARP-1 dissociates, leaving the site accessible for subsequent repair proteins to form complexes with the damage site. The recruitment of phosphorylated XRCC1 to the damage site marks the beginning of the repair phase. XRCC1 acts as a scaffold protein, facilitating the formation of the first reaction complex, which includes Ligase 3-α (Lig3α). This initial scaffold complex serves as a platform for the sequential recruitment of other repair proteins. The first enzymatic step involves the recruitment of APE1 or PNKP, which process and clean the SSB site. These enzymes remove damaged residues or repair the ends to ensure they are suitable for subsequent steps, designated in the model as scaffold 2. Further, PNKP dissociates from scaffold 2 to create space for DNA polymerase β (Pol-β). Pol-β adds the missing nucleotide to the DNA strand, effectively filling the gap left by the break (scaffold 3). The final steps involve the dissociation of Pol-β and the repair completion by Lig3α, which seals the DNA strand by forming the final phosphodiester bond. At this point, the complex reverts to the configuration of scaffold 1, representing the closure of the repair cycle. The repaired DNA strand is now restored to its intact state, ensuring genomic stability and preventing the accumulation of further damage.

This structured pathway ensures that each step is tightly regulated and proceeds in a coordinated manner. By systematically assembling and disassembling repair complexes, the cell efficiently repairs SSBs. The model incorporates experimental findings to align the simulated repair dynamics with real-world biochemical processes. Throughout these transitions, the model assumes that repair proteins dynamically interact with the SSB site, with each protein contributing to specific repair tasks. When competing factors are negligible, the repair process progresses efficiently through each scaffold complex. However, external conditions or protein competition may temporarily inhibit recruitment, reverting the system to a prior state or delaying progression. This sequence ensures a stepwise repair mechanism, where each state and scaffold complex corresponds to a distinct biochemical function, allowing for accurate simulation and modeling of SSB repair dynamics.

5.3.4 *Results*

5.3.4.1 *Quantifying the number of strand breaks in the DNA*

The first goal of this project was to quantify the number of SSB, Base Damages (BD), and DSBs induced in the cells when irradiated with protons. For this purpose, a nucleus of 10 μm diameter at G0/G1

phase with chromatin structure was defined using biological data from Hi-C experiments to investigate the entire chromatin structure (see Figure38). The specific geometry data used in this work was collected from a population of IMR90 cells by Rao et al [219]. Within the Hi-C damage model, the chromosome is represented as a continuous walk of polymer beads. Each bead represents a topologically associated domain, with size proportional to the DNA content as measured in the Hi-C data. No DNA geometry is explicitly built within these polymer beads. Instead, an energy deposition within a polymer bead is randomly sampled to determine the likelihood of the deposition occurring within an inferred DNA volume. This probability was fit to reproduce DSB and SSB yields reported in the literature [213]. Depositions within a DNA volume, passing the previously mentioned probability check, are then sampled based on the magnitude of energy deposition. A linear probability varying from 0 to 1 between 5 and 37.5 eV is used. Depositions passing this check are then randomly assigned to strand one or two of the DNA double helix, forming an isolated backbone damage. The base pair position along the chromosome is assigned randomly between the base pair constraints of the Hi-C polymer bead containing this damage. Following the simulation of all particles required to deliver a specified dose, the list of backbone damages are clustered to determine DSBs. Backbones on opposing strands and separated by 10 or fewer base pairs are grouped into a DSB[216].

Since the DNA geometry is not explicitly modelled within the Hi-C beads, and the sampling method explained above can only predict backbone damages, some backbones must be converted into base damages. In order to achieve this conversion, whilst maintaining expected yields of SSBs and DSBs, the initial number of accepted damages must be increased. This can be achieved by varying the energy deposition rule (between 5 and 37.5 eV) or by changing the geometrical sampling parameter (sensitive fraction). Since the outcome is to treat more of the nucleus volume as "sensitive" it is logical to vary the sensitive fraction. This assumes that bases, as well as the already assumed backbones, are uniformly distributed across the Hi-C polymer bead.

Incorporating the chromosome geometrical data from Hi-C measurement into MC simulations of DNA damage is a newly developed methodology by the Manchester PRECISE group, allowing the direct inference of spatial genome conformations from experimentally measured genome contacts [216]. Track structure simulations were conducted using the Geant4-DNA; TOPAS-nBio extension with the G4EmDNAPhysics list to model the DNA damage induced by a 10 MeV proton beam. A uniformly distributed circular proton source was positioned at the entrance of the bounding water box. The target DNA model was enclosed in a 50 × 50 × 50 μm water box. The

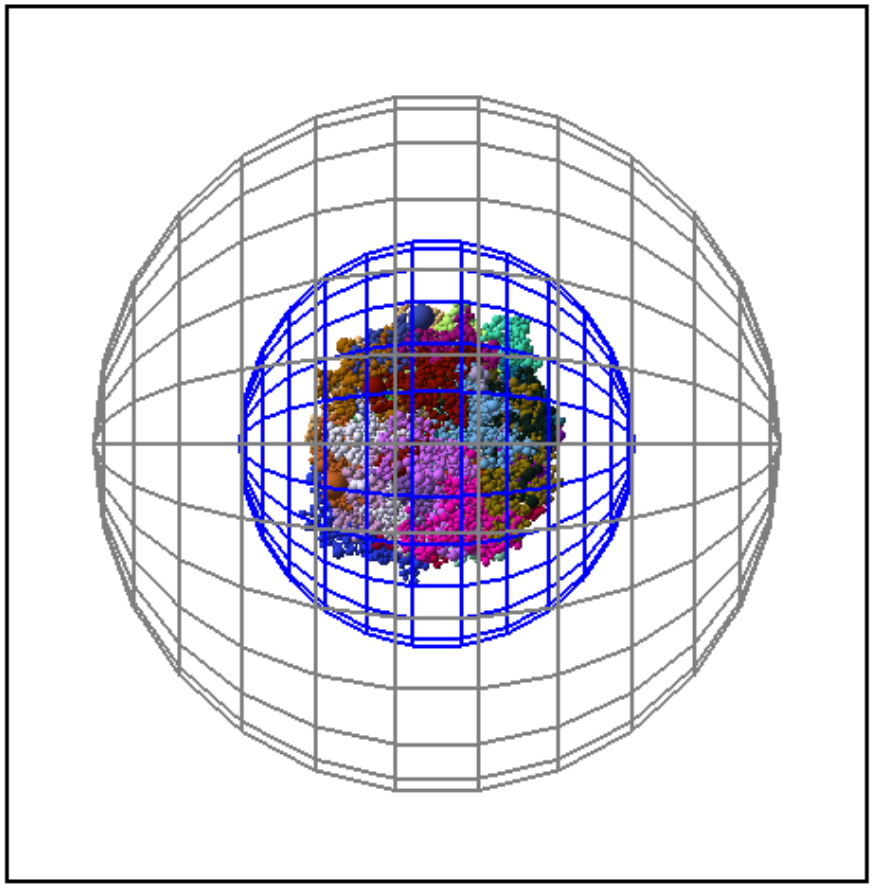


Figure 38: Chromatin structure utilized in DaMaRiS model. The colored wired center represents the chromatin material. The blue mesh is the nuclear membrane and the black mesh represent cytoplasm. While the square boxes represent the box size of the entire system under consideration.

simulation involved at least 1000 primary protons of 10 MeV. The delivered dose was regulated by adjusting the source radius, modifying the fluence, and controlling the number of particles crossing the nucleus volume. For higher dose simulations, where the beam radius was smaller than the nucleus, the number of primary particles was increased to ensure sufficient interactions [220]. The authors in [220] has classified damages into 7 types; which includes isolated base damage, isolated backbone damage (SSB), potential DSB (a backbone with an opposing base damage within 10bp), simple DSB with damaged backbones on opposite strands separated by less than 10 bp, simple DSB with at least one associated base damage within 3 bp of the DSB ends, complex DSB with multiple associated damaged backbones, complex DSB with at least one associated base damage.This present work is dependent on SSB and base damage, so only isolated

base and backbone damages were considered for analysis, though all other damage types still exist in the simulation output.

The Linear Energy Transfer (LET) of radiation is a crucial factor in determining the biological effectiveness of radiation in therapeutic settings. LET refers to the amount of energy deposited by ionizing radiation per unit length of tissue. High-LET radiation, such as α-particles, tends to cause dense ionization along its track, resulting in clustered double-strand breaks (DSBs) that are difficult for cells to repair. This high ionization leads to significant DNA damage within the target cells, while sparing the surrounding normal cells. On the other hand, low-LET radiation, such as X-rays, is sparsely ionizing and spreads its energy over a longer distance. This results in a higher frequency of single-strand breaks (SSBs) and base damage. Hence, in this study, the number of BD, SSBs and DSBs were calculated for varying LET expected for a range of proton; then these data were fit with a second order polynomial [220]. Where, LET is calculated by the fit; LET (keV/μm) = A x $(\mathrm{Energy})^{-B}$, where A = 28.8589 and B = 0.809599. For quantifying, the damages over the LET range of 1 to 40 keV/μm were calculated; thereby changing the energy likewise. Each simulation yielded different number of isolated backbone and base damages and DSBs as shown in the table (Supplementary figure). As mentioned before, this set of data was then fit to second order polynomial equation, and scaled by dose, given by:

$$\mathrm{Yield}(D, L) = D \times (aL^2 + bL + c) \tag{5}$$

where D is the proton dose in units of Gy and L is the track averaged LET in units of keV/μm. The parameters of eqn (2) are represented in Table 3

Table 3: The fitted parameters of eqn (2), for predicting the yield of each damage category as a function of proton dose and LET.

Parameter	**Isolated backbone**	**Isolated base**	**Simple DSB**
a	0.0147391	-0.0407463	-2.44×10^{-3}
b	-3.36766	1.5976	3.98×10^{-1}
c	552.28	245.949	1.56×10^{0}

Using the above listed constants, the yield of isolated backbone and base damages were obtained as a function of LET, shown in Figure 39a. In order to determine the change in sensitive fraction, the expected number of base damage is calculated. The ratio of base to backbone damage was used inform this change, i.e. for every backbone damage it is expected that there will be some number of base damages (as shown in Figure 39b). We can observe that with increasing LET value; the isolated BD and Backbone damages are either decreasing or not varying much. So a value with optimal base damage

and backbone damage was used ($N_{backbone}$ = 539 and N_{base} = 252); which was around 4 keV/µm. This change in the code converts a certain number of backbone damages produced to base damages. This results in a lower yield of SSB and DSBs. To maintain the DSB and SSB yield, more energy depositions must be accepted as occurring in DNA volumes. This was achieved by increasing the "sensitive fraction" of the modeled nucleus - the parameter that represents the likelihood of an energy deposition occurring in DNA. The simulated cell was uniformly irradiated with 100 protons at 10 MeV. The sensitive fraction was varied to reproduce the BD predicted by Equation 6. The SSB yield was assessed against the expected yields from the nominal sensitive fraction (0.141328) Table 4 shows the number of different damages obtained with varying sensitive fraction.

Table 4: Sensitive fraction vs Damages

Sensitive Fraction	**No: of BD**	**No: of SSBs**	**No: of DSBs**
0.141328 (default)	0	81	3
0.182657	49	63	5
0.152657	35	48	2
0.162657	36	52	4
0.158814	37	55	1
0.218574	82.4	83.6	7.2
0.214679	89.25	83	8.75
0.213088	86	80	10
0.213357	**85.75**	**81**	**9.75**

A sensitive fraction of 0.213357 reproduces the SSB yields expected from the nominal sensitive fraction of 0.141328. The number of DSBs has increased 3 fold compared to the nominal case. This implies that the above approach is an oversimplification, since both SSB and DSB yield cannot be maintained whilst producing the correct yield of BD. However, since the work is mainly concerned with BD and SSB processing (no DDR to DSBs is included here) the simplification was deemed to be acceptable. With this modification, the MC damage model is capable of predicting BD whilst maintaining the expected SSB yield. This finalises the DNA damage model parameters. This "clusBD" is an important characterisation when considering DDR and the direction of damages to suitable repair pathways. For this purpose, 3 assumptions were made:

1. All isolated base damages and backbone damages are considered as "isoBD" and "SSBs" respectively

2. All backbone damages which are less than 10 bp (3.14 nm) apart and are on opposite strands of same chromosome are consid-

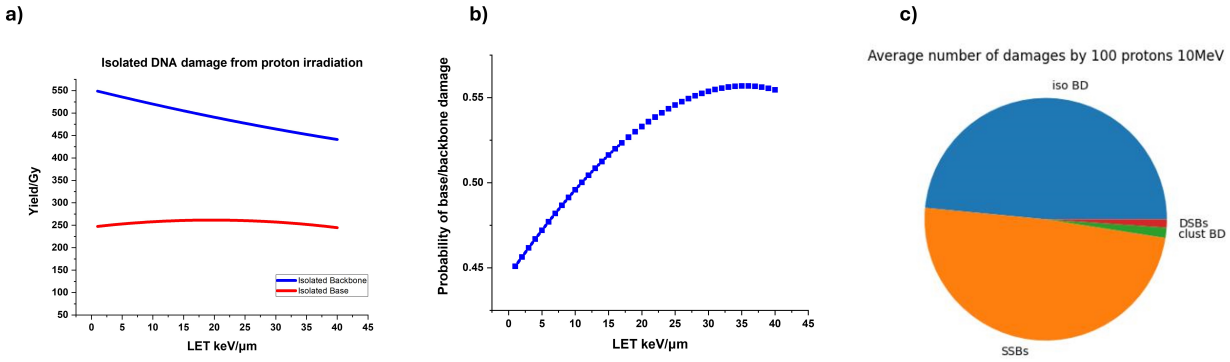


Figure 39: a) Quantity of isolated backbone damages and isolated base damages with increasing LET value b) Probability of base damage per backbone damage c) Pie chart illustrating the quantity of different type of damages obtained

ered as DSBs and are clustered together (This type of damage is not considered in the following work, due to the over prediction of DSBs (Table 4)).

3. All base damages which are less than 6 bp (2.04 nm) apart are clustered together and are termed as "clusBDs" [84].

These 3 assumptions simplify the choice of repair pathways and will be explained in the next section. The damage model uses an implementation of the DBSCAN (Density-based Spatial Clustering of Applications with Noise) algorithm following the simulation of DNA damage [221]. This algorithm searches for all the nearby points/that are closely packed which satisfy the 2nd and 3rd assumptions and clusters each of them accordingly giving rise to 2 types of clusters. All the points that are not packed together are considered as isolated damages. The system was simulated with the above mentioned parameters. With a sensitive fraction of 0.213357, 100 protons of 10MeV and LET of 4.47 keV/µm; resulted in average of 88.75,89.75,2.5 and 2 of "isoBD", "SSBs", "clusBDs" and "DSBs" respectively. Altogether, the most of the energy is being transferred into creating more of SSBs and base damages. The slight increase in the SSB formation observed, is logical considering the relative target size of a DNA backbone compared to a base. The relatively low yield of DSBs is explained by the proton energy, and resulting LET of 4.47 keV/um. However, the yield of DSB predicted by this modified damage model is irrelevant given the results of Table 4. Whilst this energy is low, representing end of proton range, the LET is not sufficiently high to create a large number of SSB clusters. Out of all the base damages only 2% of them are placed very close to each other, which is potentially risky to the cell. The rest of the isolated damages, depending upon their position in the chromosome, can be easily repaired.

It is noteworthy to speculate that, the amount of out-of-field dose during proton therapy is due to secondary neutrons; whose LET is

much lower compared to primary proton. The LET value for primary protons within the Bragg region and can go up to 8-12 keV/µm. If considering the out of field LET is less than that observed in the targeted volume [222, 223], LET value of 4 keV/µm, which could correspond to out-of-field regions. Hence the above results, could primarily mean that lower LET is associated with higher number of SSBs than DSBs or clustered base damages; consistent with the findings Goy et.al in [13, 220]. Even though their study, investigated DNA damage patterns using X-rays, which are low LET radiation. While their findings highlight an increased number of SSBs compared to DSBs at lower LETs, applying these insights to proton therapy requires caution. Protons have a distinct LET profile, with lower LET values in out-of-field regions (e.g., 4 keV/µ m in our study) and higher LET values in the Bragg peak. Despite these differences, the trend of lower LET leading to more SSBs remains consistent and could also be a speculation for future studies.

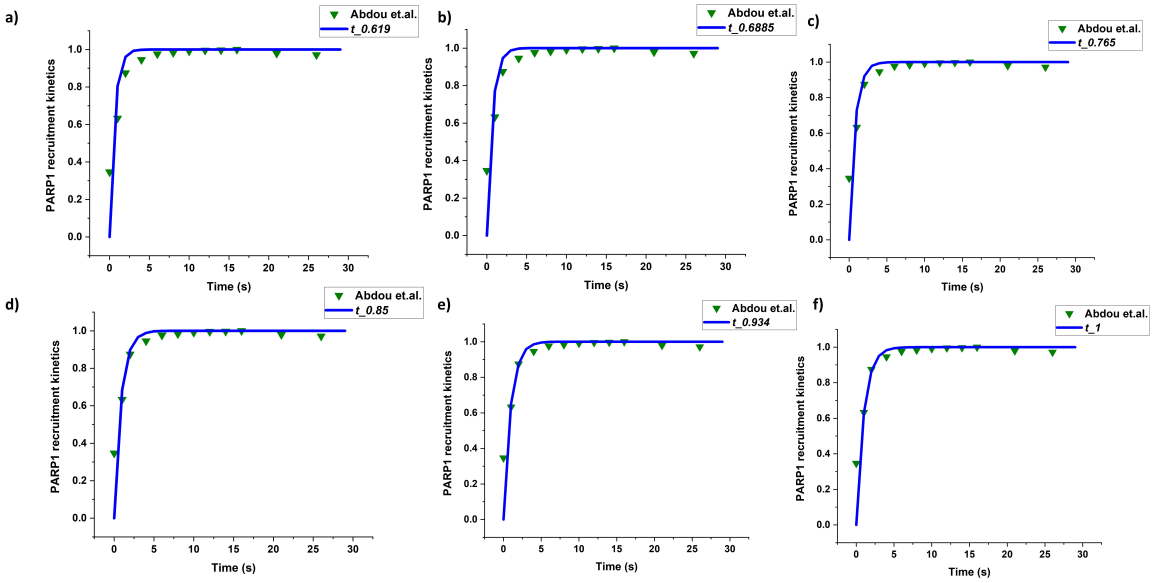
5.3.4.2 *Recruitment kinetics of SSB repair proteins*

Various literature studies exploring the kinetics of different proteins involved in the SSBR pathway were collected and the data was extracted from the graphs. Table 5 highlights the different literature utilized in this study and the related R^2 value between the model and the experimental dataset. This collated data is used to compare against model predictions. All of the experiments used laser micro irradiation to generate SSB in the cells, which predominantly resulted in the accumulation of sensor proteins (such as PARP-1, Lig-3α) and repair proteins at the damage site. The breaks induced by proton radiation, included a large number of isolated backbone damages; which requires short-patch BER/SSBR pathway to repair them (see Figure 10) [116].

Authors	**Radiation type**	**Cell-type**	**Protein**	R^2
Abdou et al. [224]	Laser micro-irradiation (405 nm)	HeLa Cells	PARP-1	0.697
Lan et al. [225]	Laser micro-irradiation (365 nm)	HeLa Cells	XRCC1	0.85822
Abdou et al. [224]	Laser micro-irradiation (405 nm)	HeLa Cells	XRCC1	0.80187
Lan et al. [225]	Laser micro-irradiation (365 nm)	HeLa Cells	Lig-3α	0.64018
Abdou et al. [224]	Laser micro-irradiation (405 nm)	HeLa Cells	Lig-3α	0.627381
Horton et al. [226]	Laser micro-irradiation (355 nm)	Mouse Embryonic fibroblasts	PNKP	0.8373
Godon et al. [227]	Laser micro-irradiation (405 nm)	HeLa Cells	Repair kinetics	0.90

Table 5: The different literature used to validate DaMaRiS Model and their R^2 value considered.

The recruitment kinetics of each protein was used to determine the time constant required for them to associate with the damage site. This was performed by changing the time constant parameter till the



Time constant	R^2 value
0.619	0.601
0.688	0.628
0.765	0.658
0.85	0.697
0.934	0.683
1	0.674

Figure 40: **(a-f)** Recruitment kinetics of PARP-1 with different time constants τ fit with the experimental data from Abdou et.al. [224]. The table shows the R^2 value obtained for DaMaRiS model fitted to experimental data. With R^2 value maximum for τ = 0.85s, it was finalized for further simulations. Using similar approach, the time constant for each protein was finalized and is reported in table 5

fit between the experimental data and the model yields a R^2 value close to 1. For instance, to finalize the τ value of PARP-1, an optimal value around 0.6 was chosen, followed by calculating the R^2 value. This is followed by increasing the τ value to 10% unless the R^2 value value close to obtained. The figure 40 depicts the plots obtained for different τ value and the associated R^2 value.

The recruitment kinetics of key SSBR pathway proteins, as modeled by DaMaRiS, demonstrate good agreement with experimental data as reported in the literature Figure 41 a–d). The damage sensor proteins are usually quickly recruited to the the damage site without any delay. Hence, a very short time constant like less than 1s is alloted for PARP-1 recruitment. The experiments were conducted in HeLa cells for both Abdou et.al [224] and Reber et al. [228]. DaMaRiS model rapidly recruits , such that it reaches 50% of the maximum within approximately 2–3 seconds and 80–90% by 10 seconds. Similarly, XRCC1 recruitment (Figure 41b) aligns well with data from Lan et al. [225], Abdou et al.[224], and Reber et al.[228], with saturation occurring by

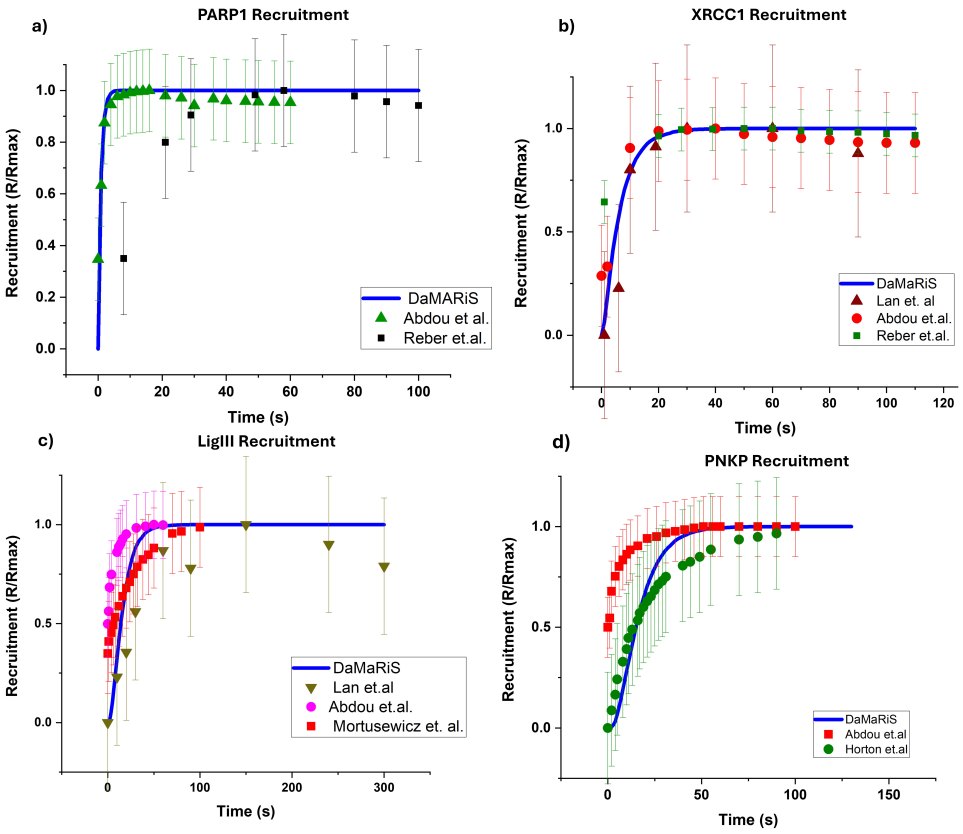


Figure 41: a)Recruitment kinetics of each protein recruited to the damage site validated with experimental data from various literature. a)PARP-1 b)XRCC1 c)Lig-3α, d) PNKP

100 seconds. Lig3α recruitment (Figure 41c) exhibits slower kinetics compared to upstream proteins, consistent with experimental data from Lan et al., Abdou et al., and Mortusewicz et al. [218], reaching maximum levels around 300 seconds. For PNKP recruitment (Figure 41d), the DaMaRiS model captures the trends reported by Abdou et al. and Horton et al. [226], with recruitment saturating at approximately 150 seconds. Out of the 3 or 4 studies used to validate the model, we chose the one with right R^2 value, as pointed out in the table 5 and the corresponding chosen time constants are given in the figure 37. Overall, the DaMaRiS model effectively reproduces the temporal dynamics of SSBR protein recruitment observed in experimental systems. Although, the kinetics of Pol-β recruitment is not shown due lack of experimental data corresponding to SSBR pathway; thereby the parameters were tuned such that the final quantity of repaired SSBs can be accounted, i.e. the SSB yield after 30 minutes of repair matches the yields as explained in section 5.3.4.3.

5.3.4.3 *Overall SSB Repair Kinetics*

To check whether the model could successfully orchestrate the proteins and bring about SSB repair, overall repair kinetics were calcu-

lated. SSB repair kinetics are normally quantified by performing a comet assay. The comet assay is a popular technique based on the fact that, when an electric field is applied, the damaged DNA fragments in the gel migrate towards the anode; creating a shape of the comet, hence the name [229]. Other than comet assay, accumulation kinetics of repair proteins like XRCC1, which are exclusively seen during BER/SSBR pathway [116]. To fit the model to the experimental data, experimental studies on SSB repair by protein after radiation treatment of cells. The studies include data obtained from Godon et.al [227] and Reynolds et.al. [230].

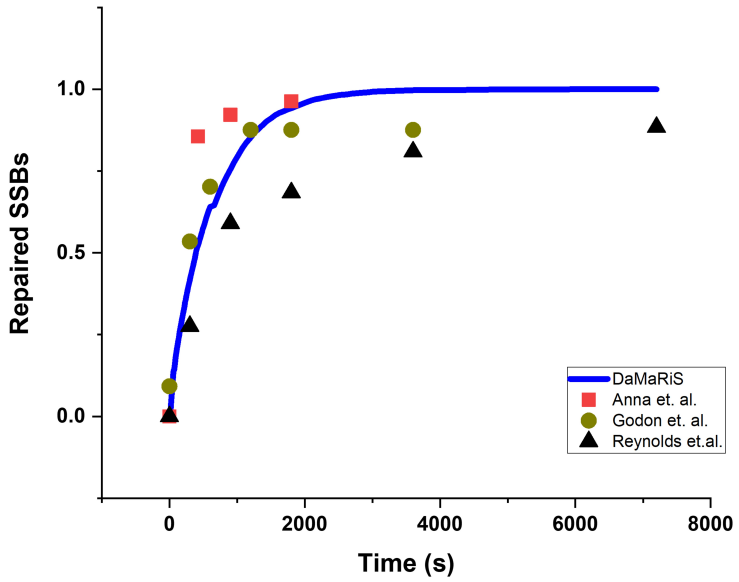


Figure 42: SSB repair kinetics for the simulated model. Entire damage population is repaired by 30mins. They were fitted with the experimental data observed by Anna et al. [231], Godon et al. [227] and Reynolds et al. [230]. Each of them having a R^2 value of 0.78,0.90, and 0.71 respectively.

As seen in Figure 42 the model agrees well with the experimental result. Both the model and the experimental results shows that the overall repair lasts up to 30 mins. One interesting thing to note is the rapidity of repair protein recruitment, e.g. XRCC1/ are completely recruited in seconds. Despite this rapid recruitment, the final step of repair process slows down and takes around 30 minutes to completely repair the damaged DNA as seen in Figure 41 resulting from a much, relatively, slower processing time of the damage site.

In conclusion, the model effectively simulates the orchestration of proteins involved in the SSB repair process, closely matching experimental observations. The repair kinetics demonstrate that while protein recruitment, such as XRCC1 and , occurs rapidly, the final stages of repair are significantly slower, requiring up to 30 minutes for com-

plete DNA restoration. This highlights the complexity and sequential nature of the BER/SSBR pathway. The strong agreement with experimental data underscores the model's reliability and its potential as a tool for further exploring DNA repair dynamics.

Although the model successfully predicted the dynamics of the SSBR pathway, it has certain limitations. Incorporating DNA geometry into the system would provide a more accurate representation of the experimental scenario, capturing the spatial constraints that influence repair efficiency. Additionally, exploring the cross-talk between SSB and DSB repair under different radiation types could offer deeper insights into therapy outcomes. X-ray radiation, which generates a high number of SSBs, predominantly activates the SSBR pathway, potentially leading cells either toward repair or senescence. This could be a disadvantage in cases where senescent cells contribute to tumor progression or therapy resistance. In contrast, proton therapy induces more DSBs, which are inherently more lethal, rendering SSBR less effective in these contexts. A comparative study of these effects could aid in optimizing radiation therapy strategies tailored to specific cancer types. Expanding this analysis to include the role of inhibitors, such as PARP inhibitors, could provide a mechanistic understanding of post-radiation therapeutic interventions. By modeling the duration of trapping, it is possible to estimate the persistence of SSBs and their potential conversion into complex DSBs, ultimately leading to cell death. A similar approach could be applied to out-of-field radiation effects, where senolytics—rather than PARP inhibitors—could be explored to selectively eliminate senescent cells and improve therapeutic efficacy. This integrated modeling framework could serve as a powerful tool for refining radiotherapy and combination treatment strategies.

5.4 SUMMARY AND OUTLOOK

In this Chapter, I investigated the role of mathematical modeling and MC simulations in understanding the DNA damage repair pathway, particularly SSBR pathway in the out-of-field area and in the targeted volume during the radiation therapy. We employed two complementary approaches to enhance our understanding. The the single-strand break repair (SSBR) pathway triggered by out-of-field radiation exposure was mathematically modeled using COPASI software. The model effectively simulates the complex biochemical cascades of the SSBR pathway, highlighting the critical decision-making steps influenced by different kinetic equations. Time-course simulations revealed key dynamics, such as the impact of reactive oxygen species on availability and the recruitment behavior of XRCC1, consistent with experimental observations [13]. The model also highlights the preferential activation of the p16 pathway over the p53 pathway, leading to

senescence rather than apoptosis, which is experimentally observed in fibroblasts. Additionally, the sensitivity of the model to kinetic equations and parameter values emphasizes the need for accurate parameter estimation in future studies.

The deterministic approach to depict how cell decisions can be affected by changing the concentrations of each protein species in the pathway, but still the lack of experimental setup was a disadvantage to get meaningful results. Therefore a complementary approach for modeling the SSBR pathway was performed using the GEANT4-TOPAS framework. At first, the number of single strand breaks and base damages induced in the chromatin were quantified, using assumptions which segregated isolated backbone damages as SSB, isolated base damages, double strand breaks and clustered base damages. Running clustering algorithms using DBSCAN in Geant-4 DNA, we observed that lower LETs are linked to a higher number of single-strand breaks compared to double-strand breaks or clustered base damages. Once quantified, the focus was shifted to introducing the SSBR pathway to the SSBs produced within the system. Therefore, the recruitment kinetics of key SSBR pathway proteins were studied. The recruitment kinetics of key SSBR pathway proteins, as modeled by DaMaRiS, showed good agreement with experimental data, with proteins like , XRCC1, and Lig3α being recruited at varying time points, thus highlighting the sequential nature of the repair process. The DaMaRiS model effectively reproduced the temporal dynamics observed in experimental systems, validating its reliability for exploring DNA repair dynamics. In summary, the model successfully simulates the orchestration of proteins involved in the SSBR process, with rapid recruitment followed by slower final repair stages. This comprehensive analysis highlights the complexity of the SSBR pathway and its sensitivity to modeling parameters.

6

CONCLUSIONS AND FUTURE PERSPECTIVES

The thesis titled "Computer Simulation and Experiments of Radio-induced Cell Senescence" is a work that entirely focuses on exploring the connections between cellular senescence and DNA damage. The state-of-art of this work presented in the Chapter 1 gives a comprehensive overview on how the presence of single-strand breaks (SSB) triggers the SSBR/BER pathway, and how the unrepaired fraction of SSB can lead the cell to make critical decisions between apoptosis or cellular senescence. The crucial decisions are associated majorly with the identification of the DNA break in the upstream, and processing of the break in the downstream of the triggered pathway (see Figure 34). Hence, this work revolves around three main aspects of the DNA damage repair pathway: **(i)** intricate dynamics between the SSB damage and the PARP-1 protein, notably, how the protein identifies the damage; **(ii)** influence of DNA damage on chromatin structure and dynamics; and, **(iii)** decision-making process in SSBR pathways. I approached items (i) and (ii) by utilizing all-atom molecular dynamics (MD) simulations, and item (iii) by mathematical modeling and Monte Carlo (MC) simulation, supported with experimental results wherever possible.

The SSB damage produced during cancer radiotherapy treatments usually ends up with removing an entire nucleotide from the DNA sequence, leaving the broken ends with 3′-OH 5′-dRP terminations. These unprocessed ends are identified by damage sensor proteins like PARP-1. There are numerous structural and experimental studies on PARP-1 identifying SSB damage. However, there is scarce literature on understanding the intricate dynamics behind this identification process. Hence, in Chapter 2 and 3, I investigated the interaction of Zinc-finger domains Zn1-Zn2 of PARP-1 protein with SSB, in a DNA hairpin loop as used in several experiments, and in a mononucleosome including full histone tails, respectively. The MD simulations in the DNA hairpin revealed that its conformational flexibility, driven by thermal fluctuations, allows for significant bending and gap openings independent of direct PARP-1 contact. Zn2 was found to bind to the 3′ end of the SSB preferentially, initiating DNA bending and gap widening, followed by Zn1 interacting with the 5′ end, leading to the final kinked structure, nearly coincident with the experimentally observed state. Hence, the experimental complex was confirmed as a global thermodynamic minimum, supporting the reliability of our simulation protocol. However, direct docking of Zn1-Zn2 on a SSB on a straight, still closed hairpin was less effective due to ad-

ditional protein-DNA contacts preventing sufficient hairpin bending. The study also explored whether PARP-1 functions as a monomer or dimer, showing Zn2's higher affinity for SSB ends and its critical role in initiating interactions, with Zn1 playing a secondary role. Additionally, spontaneous fluctuations of the DNA hairpin before PARP-1 binding likely enhances the observed binding affinity, providing a qualitative support to the idea of a 'bend-then-bind' mechanism, in which the damage site spontaneously exposes itself to the action of the repair enzyme. This is a seminal idea that we have been exploring in our research groups, in the attempt at unraveling alternative mechanisms by which proteins can rapidly and effectively identify their target, in the vast complexity of the genome.

Facing a more realistic condition that could approach the context of nuclear chromatin, the next study was performed on a whole nucleosome with multiple Zn-finger domains of PARP-1 interacting with multiple SSB break sites. The leading idea was to highlight the role of mechanical deformation of the damage site, as a way for the damage to become 'easily' visible to the repair enzyme. Hence, I focused on the nucleosome structural deformation and the interactions of histone tails with the PARP-1 protein and the damage site. The findings reveal that SSBs induce localized DNA openings, with fluctuations in the width of the break ("gap-opening") driven by both thermal and mechanical effects. The presence of Zn1-Zn2 significantly enhances these openings, particularly at the damaged sites. These interactions enable critical contacts, such as Leu-151 proximity to the 3′ end, and hydrogen bonding with the DNA backbone, facilitating SSB recognition. While achieving the "trademark" interactions of PARP-1 observed in hairpin experiments proved challenging in the nucleosomal context, the interactions of Zn1-Zn2 with histone tails and the DNA backbone effectively opened the DNA at the break site. These results suggests that Zn1-Zn2 may not solely be responsible for SSB identification in complex nucleosomal assemblies, but additional interactions from histone tails and DNA backbone fuels the identification process efficiently. Furthermore, Zn1-Zn2 presence prompted structural reorganization in the nucleosome, increasing the splitting between adjacent gyres of the DNA near the SSB, and highlighting its role in chromatin remodeling. Altogether, this part of the study underscored the unique structural contributions of Zn1-Zn2 in nucleosomal remodeling, and its pivotal role in facilitating PARP-1's function in recognizing and responding to DNA damage, offering valuable insights into the molecular mechanisms underlying genome stability and repair.

In a further step of this work, I moved to study the evolution of SSB and DSB within the chromatin structure, as another aspect of the main question, namely, which mechanism(s) may possibly help the repair proteins to quickly identify the damage sites, as opposed to the inefficient, purely randomly search along the DNA sequence.

By studying damage in the DNA linkers joining two nucleosomes, this work is among the first to demonstrate how complex defects within linker DNA can induce significant structural alterations in nucleosomal architecture. The results revealed that SSBs in the linker DNA, under just a minimal force to induce a nucleosome relative displacement, induce a pronounced bending and kink defects at the break site, whereas linker DNA without damage remained mechanically resilient under the same conditions. Notably, linker of different lengths including SSBs exhibited extensive mechanical deformation, suggesting that a minimum linker length (>22-24 bp) is necessary for inducing the mechanical instability. Even more interestingly, the presence of a DSB in the same dinucleosome structures caused dramatic bending and structural changes, driven solely by thermal fluctuations, with the final conformations remaining stable. These findings suggest that the complexity of the defect directly correlates with the degree of structural deformation. From a biological perspective, such structural alterations in damaged linker DNA may serve as "markers", exalting (and greatly accelerating) damage recognition by repair enzymes, such as Ku70/80, glycosylases, or PARP-1, which could preferentially target regions of structural instability rather than conducting an exhaustive search, randomly spanning back and forth all the chromatin DNA. Furthermore, the results raise compelling questions about the cumulative effects of clustered DNA damage within chromatin, where multiple defects might drive extensive chromatin remodeling events, potentially influencing gene expression, replication, or repair pathways. These structural disruptions could act as triggers for cellular outcomes such as senescence or apoptosis, as explored in the final chapter of this thesis. The interplay between damage-induced chromatin remodeling and cellular responses underscores the broader implications of this work, highlighting the necessity of further studies to unravel how molecular damage propagates to higher-order chromatin organization and influences critical biological processes. Insights from these investigations could ultimately inform the development of targeted therapies for DNA repair-related diseases, including cancer.

Finally, in Chapter 5, I explored the role of mathematical modeling and Monte Carlo (MC) simulations in understanding the DNA damage repair pathway, particularly the single-strand break repair (SSBR) pathway, both within and outside the targeted volume during radiation therapy. Using COPASI software, I developed a mathematical model of coupled differential equations, to simulate the SSBR pathway triggered by out-of-field radiation exposure, capturing the complex biochemical cascades and decision-making steps influenced by kinetic equations. Time-course simulations revealed key dynamics, such as the impact of reactive oxygen species on PARP-1 availability and XRCC1 recruitment, aligning with experimental observations.

The model also highlighted the preferential activation of the p16 pathway over p53, leading to senescence rather than apoptosis, consistent with fibroblast studies. However, the results remain purely theoretical because of the present lack of experimental data to validate the model.

In a second part, I introduced the SSB-repair model in the GEANT4-TOPAS simulation environment, for which a large amount of experimental data about PTV irradiation with energetic particles is available. This framework quantified single-strand breaks (SSBs) and base damages in chromatin, revealing that lower linear energy transfer (LET) radiation induces more SSBs than double-strand breaks or clustered base damages. Furthermore, the recruitment kinetics of SSBR proteins, modeled using DaMaRiS, matched experimental data, with proteins like PARP-1, XRCC1, and Lig3α recruited sequentially, reflecting the repair process's temporal dynamics. The model successfully simulated the orchestration of SSBR proteins, showing rapid recruitment followed by slower repair stages, and underscored the pathway's complexity and sensitivity to modeling parameters. Together, these approaches provided a comprehensive understanding of SSBR dynamics and its response to radiation-induced damage.

This thesis paves the way for a broad spectrum of future investigations, encompassing both experimental and computational approaches. The current work focused on DNA, nucleosome and (small-scale) chromatin dynamics and mechanics in the presence of DNA damage, providing a foundation for further exploration. A particularly intriguing direction would be to investigate the effects of PARP-1 binding to DNA damage sites in the dinucleosome, elucidating the mechanisms of interaction and the extent of linker DNA bending under varying lengths and damage conditions (SSB/DSB). Additionally, it would be valuable to examine how the presence of a damage sensor protein influences local chromatin structure and remodeling processes. Furthermore, comparative studies involving various damage sensor proteins could yield significant insights. Research has demonstrated that damage sensor proteins involved in mismatch repair (MMR), nucleotide excision repair (NER), and single-strand break repair (SSBR) induce DNA bending or kinking during their recognition and repair activities [232, 233]. Comparing these mechanisms across different repair pathways could reveal common or unique structural and dynamic features at the basis of damage recognition. The complexity of such studies could range from analyzing single DNA defects to more complex damage scenarios. These investigations would enhance our understanding of the efficiency with which damage sensor proteins recognize DNA lesions, providing critical insights into the molecular basis of DNA repair. Such knowledge could ultimately inform the design of targeted protein inhibitors, such as

PARP-inhibitors, which are used to disrupt repair pathways in cancer cells, thereby improving therapeutic strategies for cancer treatment.

Another compelling computational aspect is the exclusive integration of out-of-field, dose-induced SSBR within the Geant4-TOPAS framework. Incorporating a detailed DNA geometry would enhance the accuracy of the model, by accounting for spatial constraints that influence repair efficiency. Additionally, investigating the interplay between SSB and DSB repair across different radiation types (photons, charged particles) could provide deeper insights into therapeutic outcomes. X-ray radiation, which generates a high number of SSBs, primarily activates the SSBR pathway, potentially leading to either repair or senescence. This could be detrimental in cases where senescent cells contribute to tumor progression or resistance to therapy. Conversely, proton therapy induces more DSBs, which are inherently more lethal, making SSBR less effective in these scenarios. A comparative study of these effects could help refine radiation therapy strategies tailored to specific cancer types.

Expanding this analysis to include the role of inhibitors, such as PARP inhibitors, could offer a mechanistic understanding of therapeutic strategy. For example, by modeling the duration of PARP-1 trapping, it would be possible to estimate the persistence of SSBs and their likelihood of transforming into complex DSBs, ultimately leading to cell death. This integrated modeling framework could be a valuable tool for optimizing radiotherapy and combination treatment strategies. It could provide critical insights into whether out-of-field radiation doses, typically encountered during radiotherapy, induce SSBs and DSBs in the surrounding tissue. By quantifying the amount and fraction of breaks, the model could help unravel the mechanisms behind cellular responses to radiation – particularly why cells exposed to out-of-field dose tend to favor senescence over apoptosis [13]. Ultimately, these findings should be further validated through experimental data, strengthening the correlation between simulation outcomes and biological responses.

APPENDIX

A

SUPPLEMENTARY FIGURES

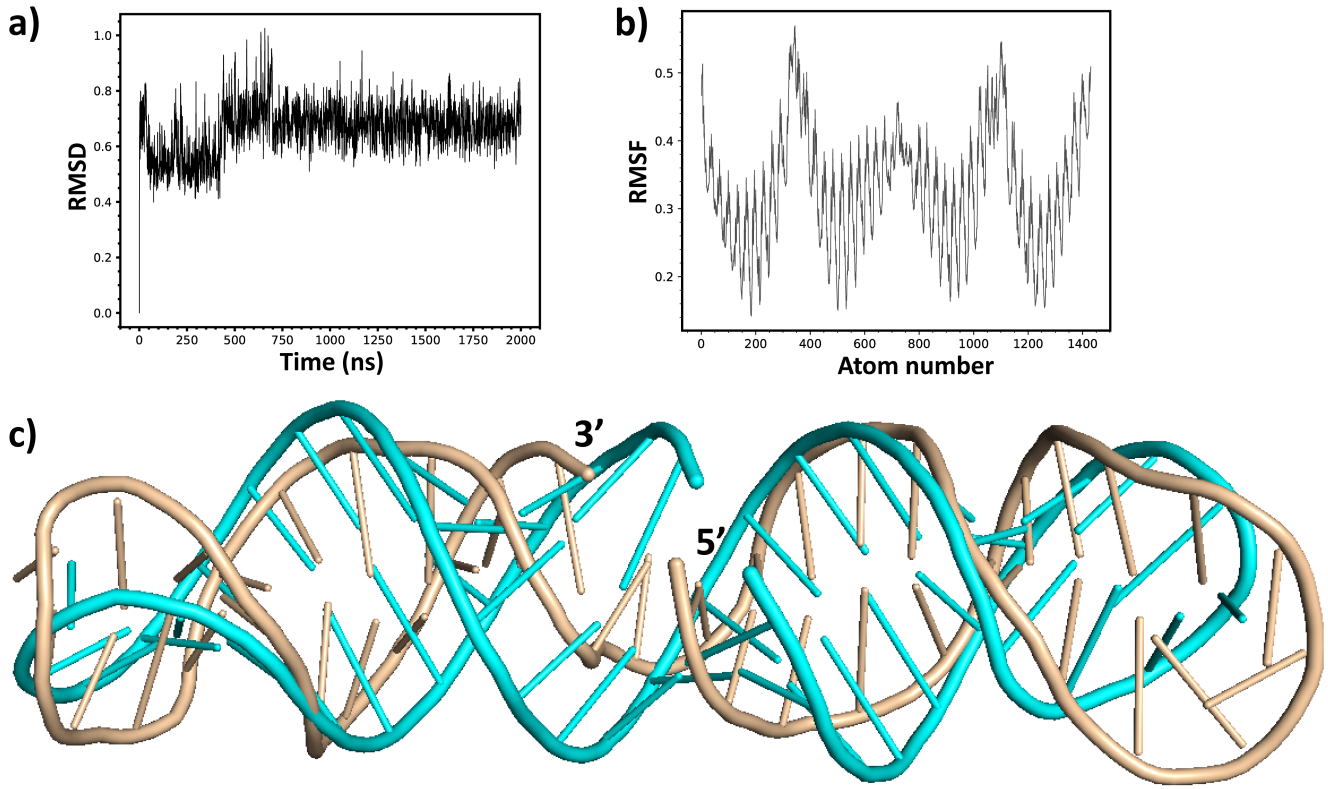


Figure A1: **(a)** RMSD values of isolated DNA hairpin in configuration **1** over 2µs MD simulation. **(b)** RMSF values on a residue-by-residue basis. **(c)** Superimposed structure of the hairpin loop best cluster after 2µs simulation (gold), and the initial structure (cyan).

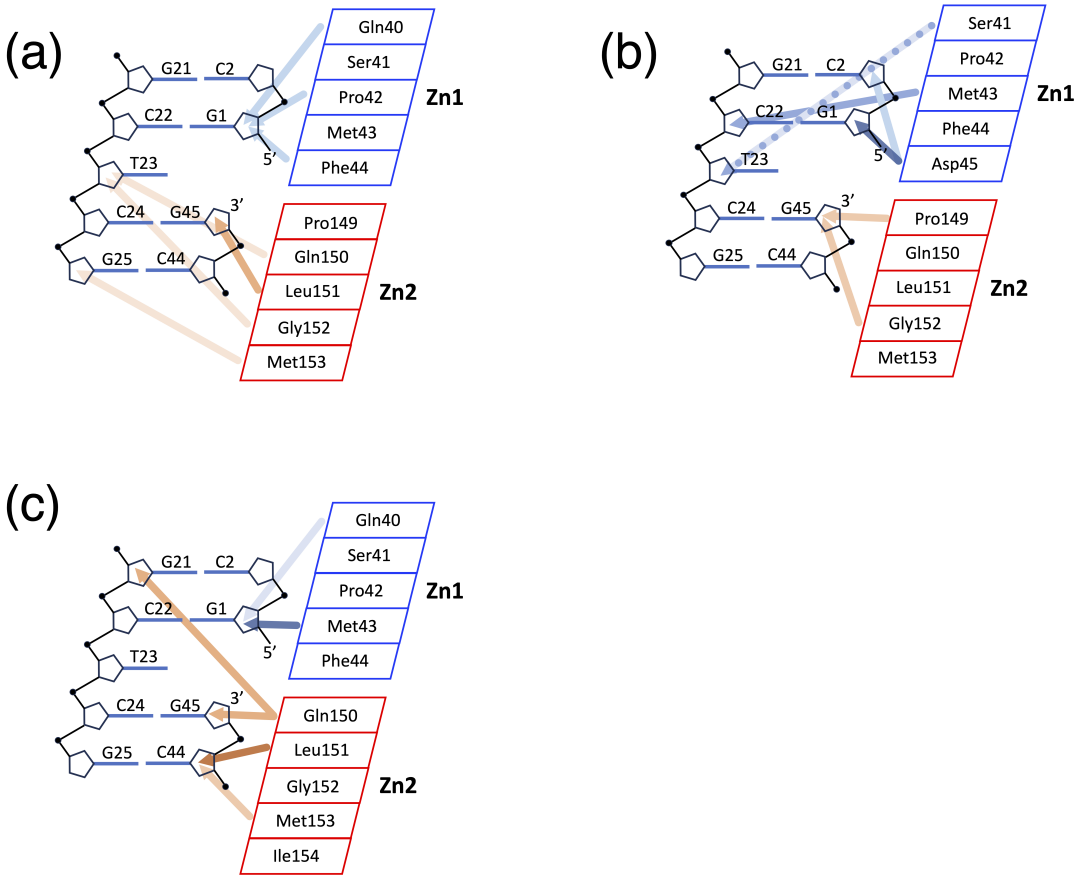


Figure A2: Contact maps of the PARP1-DNA hairpin assembly (Zn1 blue, Zn2 red) including an SSB (missing nucleotide at position 46). Arrow-color intensity according to increasingly close bonds. **(a)** Docking + MD of PARP-1 Zn1 and Zn2 domains against a DNA hairpin in the NMR experimental configuration **1**. **(b)** Docking + MD PARP-1 Zn1 and Zn2 domains against the straight DNA hairpin configuration **2**. **(c)** After MD of PARP-1 manually fitted by CHIMERA to the straight DNA hairpin **2**.

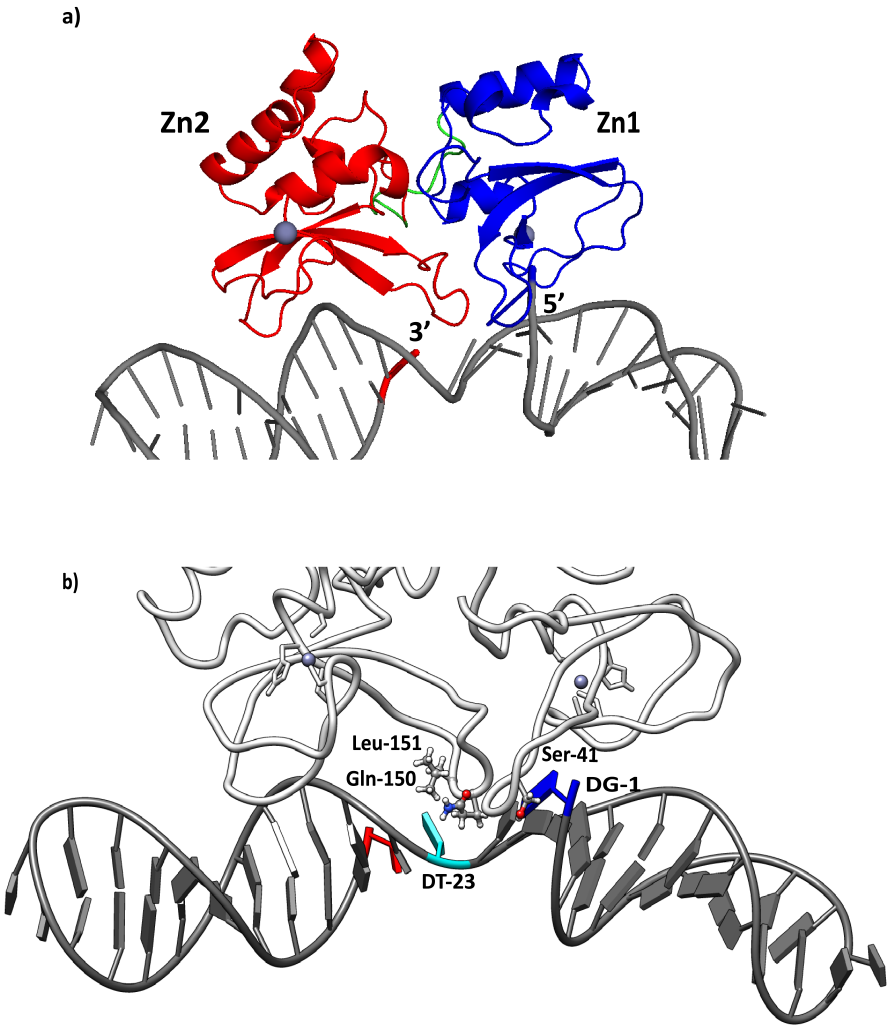


Figure A3: **(a)** Zn2 and Zn1 interaction with the straight hairpin, after the bare docking protocol followed by 0.5 µs MD. **(b)** Zn2 interacting with the T23 (unpaired nucleotide) via Leu-151 and Gln-150, while Ser-41 interacting closely with 5′ (DG1) end.

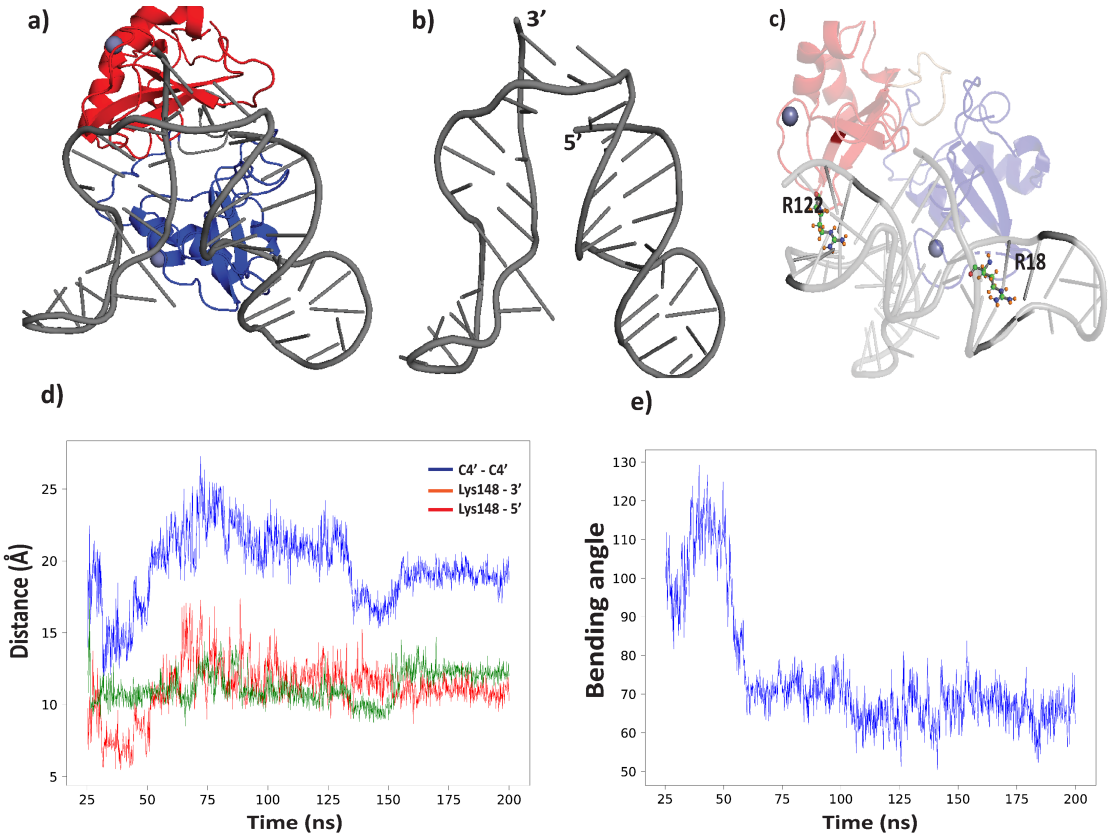


Figure A4: **(a)** Best cluster obtained after 200ns-MD simulation with manually adjusted Zn1-Zn2 starting from the straight hairpin. Notice Zn2 (red) interacting with both SSB ends, and Zn1 (blue) locking the bent DNA conformation. **(b)** Snapshot of DNA bending during the simulation. **(c)** Snapshot of Arg-18 and Arg-122 penetrating into the minor groove of DNA and imposing the bending. **(d)** Evolution of SSB gap opening (blue, C4′-C4′ distance); red,green=distance between Zn2's Lys-148 and the 5′ or 3′ ends throughout the 200-ns simulation. **(e)** Evolution of DNA bending angle over the same 200 ns.

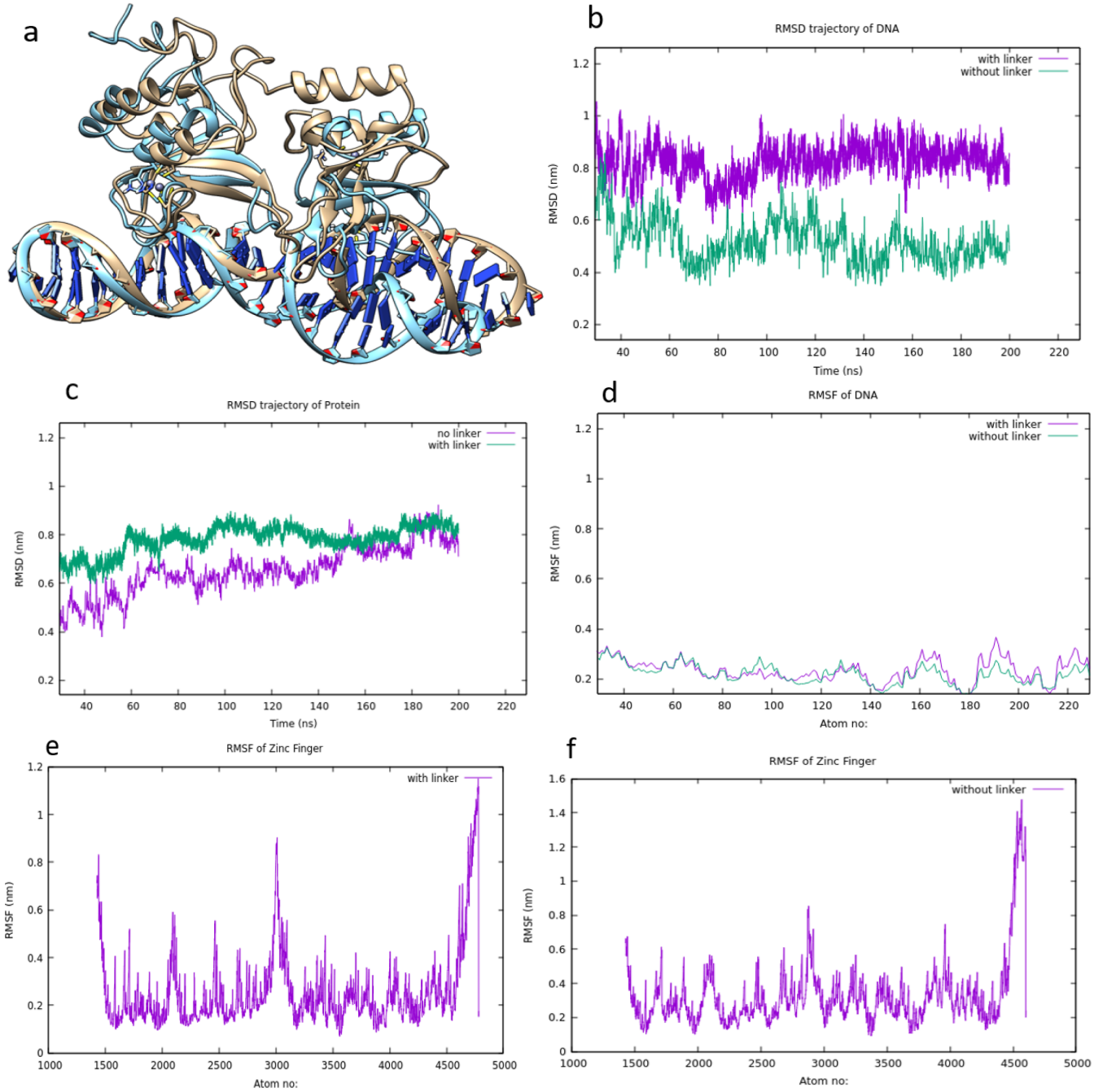


Figure A5: **(a)** Superimposed structure of the best cluster of Zn1-Zn2 interacting with DNA after 0.5 µs MD simulation, with (gold) vs. without (cyan) the 13-aa linker joining the zinc fingers. **(b)** RMSD trajectories of DNA, and **(c)** of Zn1-Zn2, purple= with linker, green= without linker. **(d)** RMSF of DNA by residue, with and without linker. **(e,f)** RMSF of Zn1-Zn2 with vs. without linker.

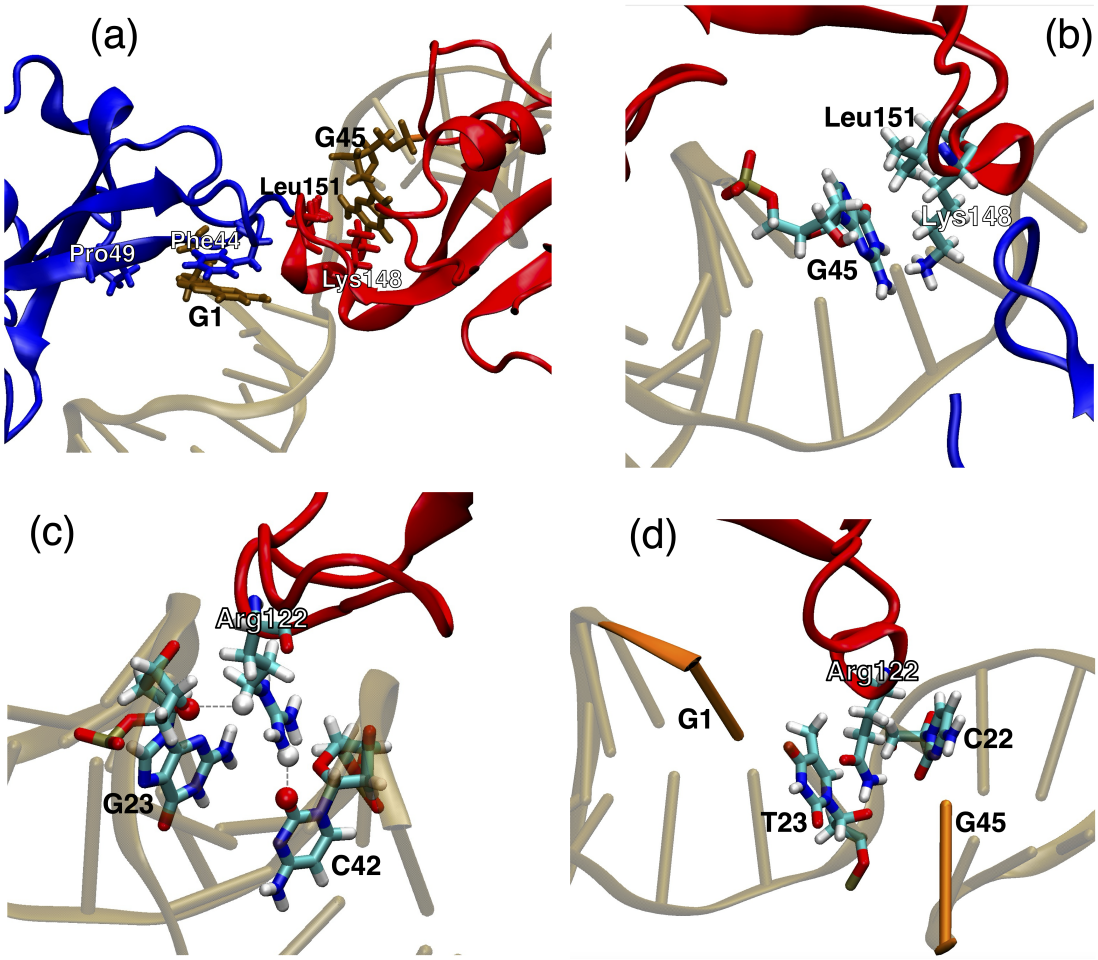


Figure A6: Interactions between Zn1-Zn2 and the hairpin in the orientation **(2)**. **(a)** Overview with Zn1 and Zn2 in blue and red ribbons, DNA in ochre (ochre sticks). Main residues interacting at 3′ and 5′ ends are highlighted in licorice representation. **(b)** Electrostatic interaction between Zn2 and the 3′ hairpin end; Zn1 is represented by blue ribbons in foreground. **(c)** Arg-122 from the loop 119-122 penetrating into the minor groove, and making H-bonds with C23 and C42. **(d)** Gln-150 penetrating deep into the SSB gap, and interacting with C22 and the unpaired T23 (Zn1 loops removed for clarity; the SSB ends, G1 and G45, are shown in orange sticks).

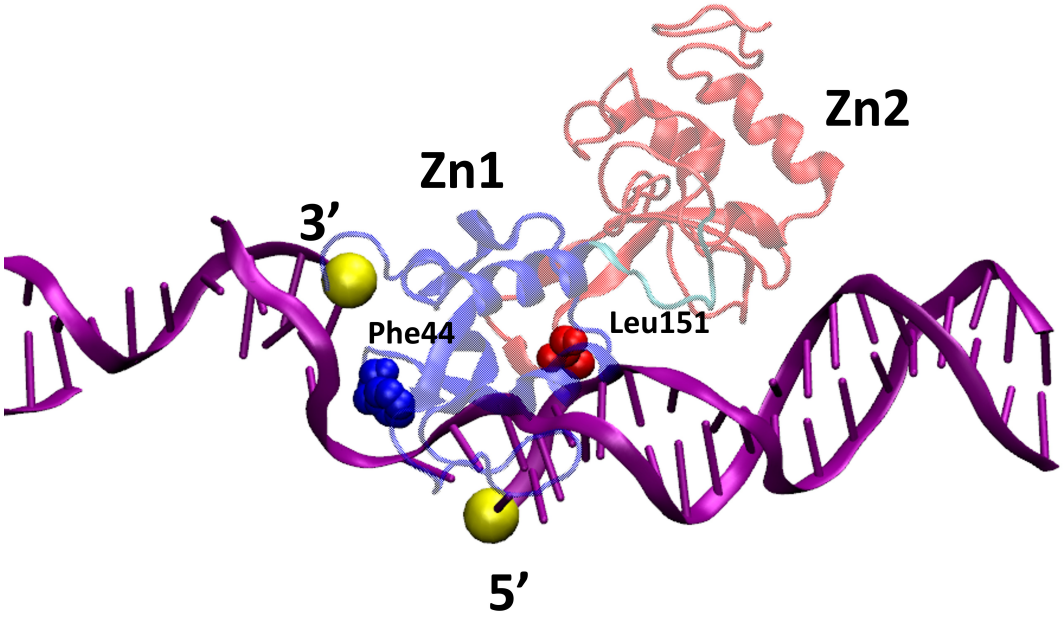


Figure A7: Docking Zn12 at a single stand break site with PydockDNA. The conformation obtained is the first ranked configuration obtained. The software couldn’t replicate the same interactions even after giving specific residues for interaction for our input structure. The best configuration obtained had Zn1 close to 3′ end and Zn2 behind the Zn1.

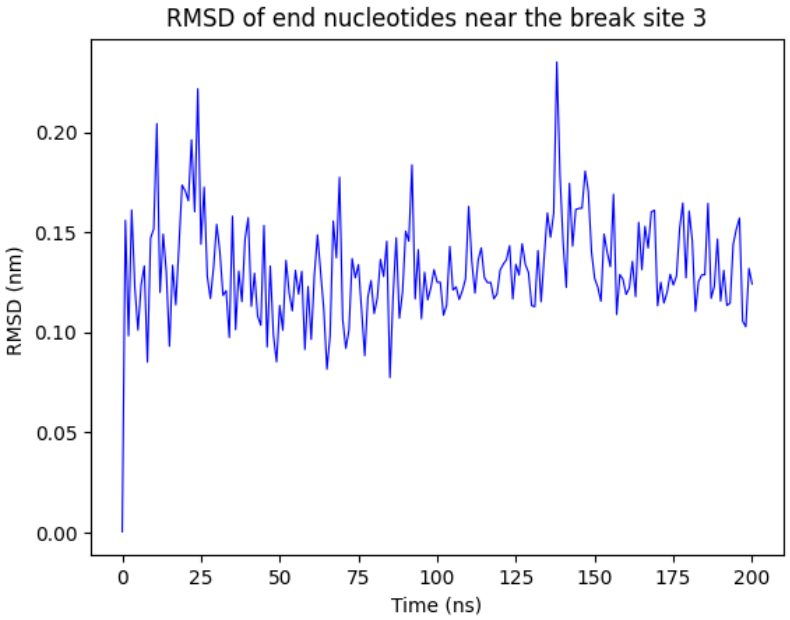


Figure A8: RMSD of end nucleotides in the break site 3

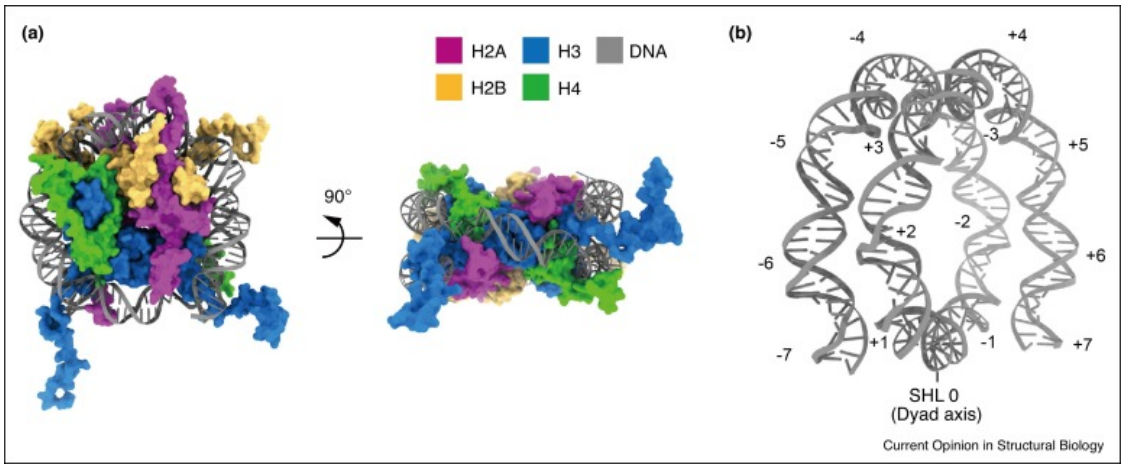


Figure A9: (a) High-resolution structure of the nucleosome core particle. Histones and DNA are shown in surface and cartoon representations, respectively. H2A, H2B, H3, H4, and DNA are colored magenta, yellow, blue, green, and grey, respectively. (b) The superhelical locations (SHLs) of the nucleosome. The nucleosomal DNA is left-handedly wrapped around the histone octamer surface. The nucleosomal dyad (center) position is numbered as SHL 0. Each SHL location is numbered as SHL -1 to -7 and SHL +1 to +7 from SHL 0, with ~10 base-pair periodicity. The figure is taken from [234].

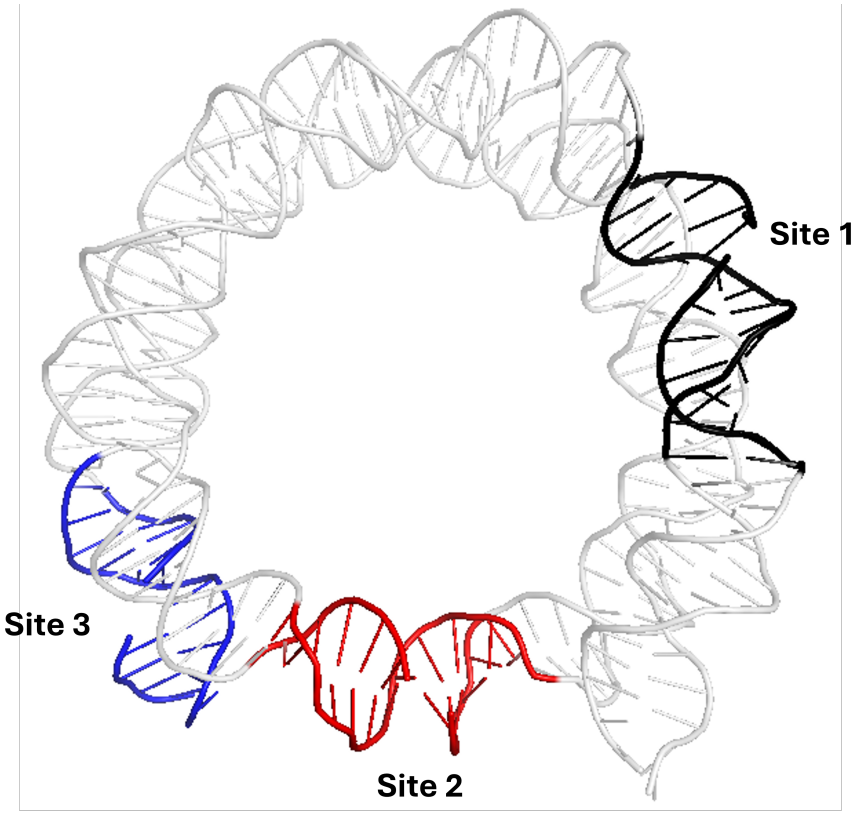


Figure A10: Position and numbering of the SSB introduced in the DNA wrapped around the nucleosome for H-bond analysis.

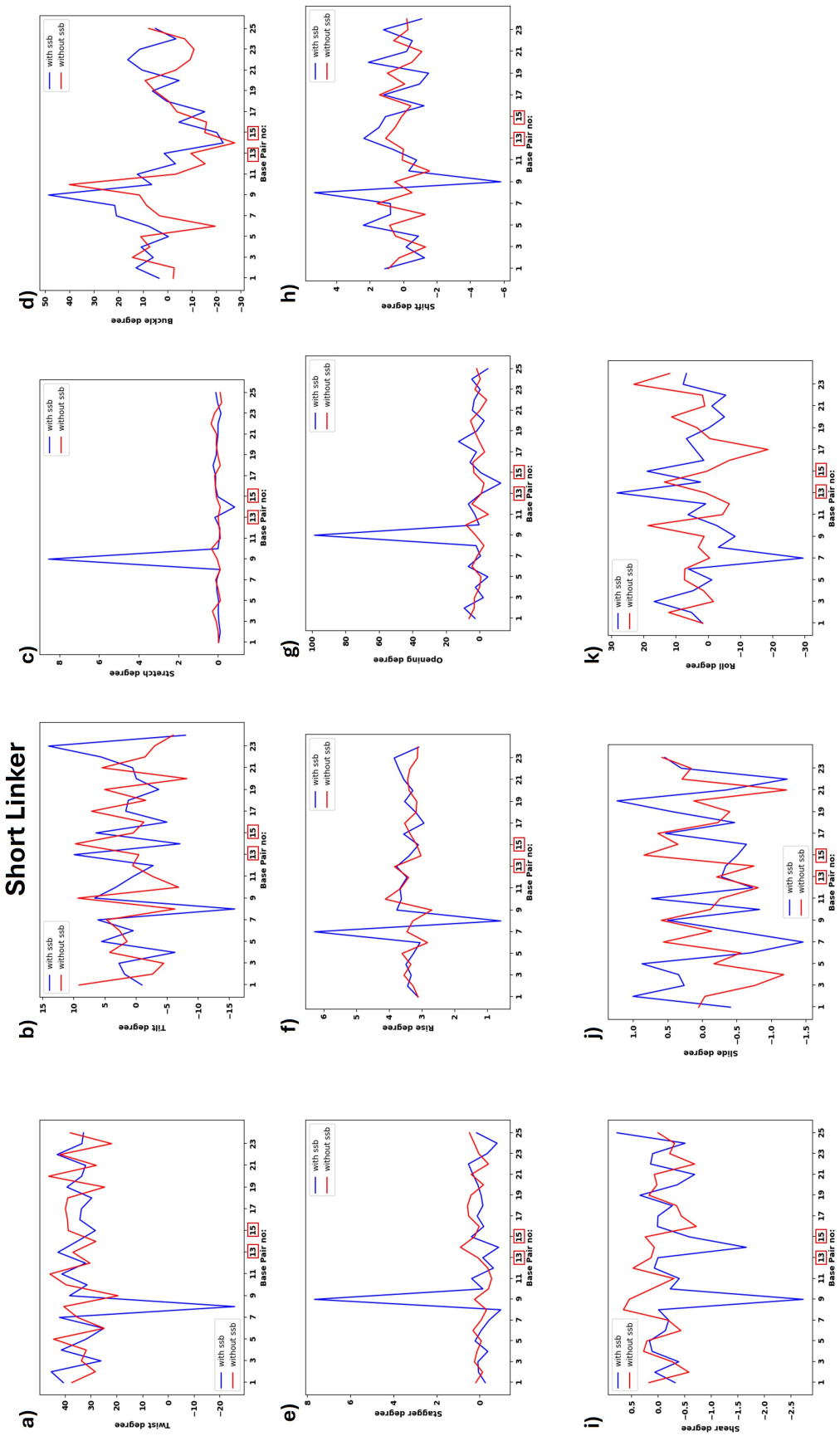


Figure A11: Different helical parameters for short linker with and without SSB.

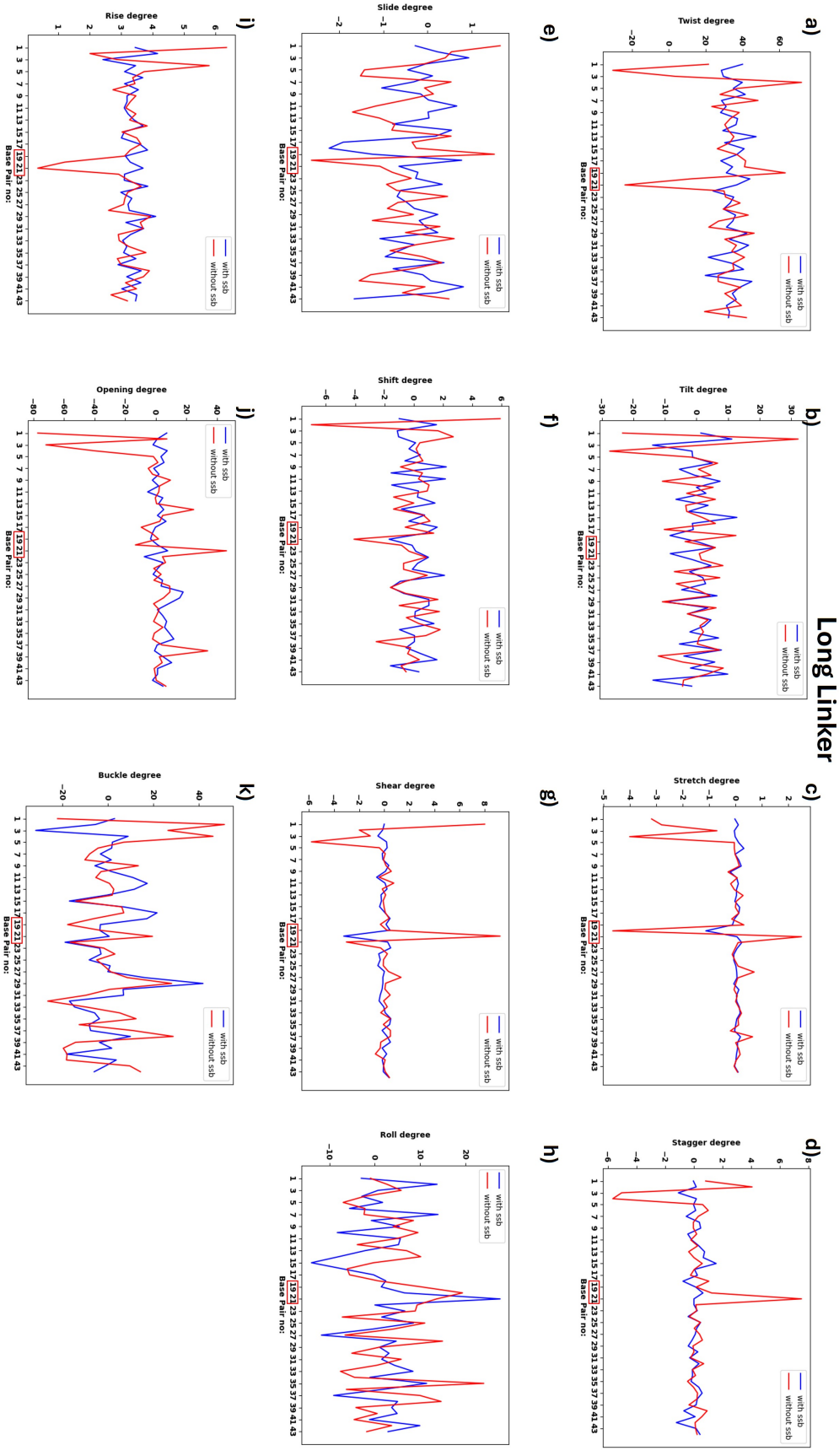


Figure A12: Different helical parameters for long linker with and without SSB.

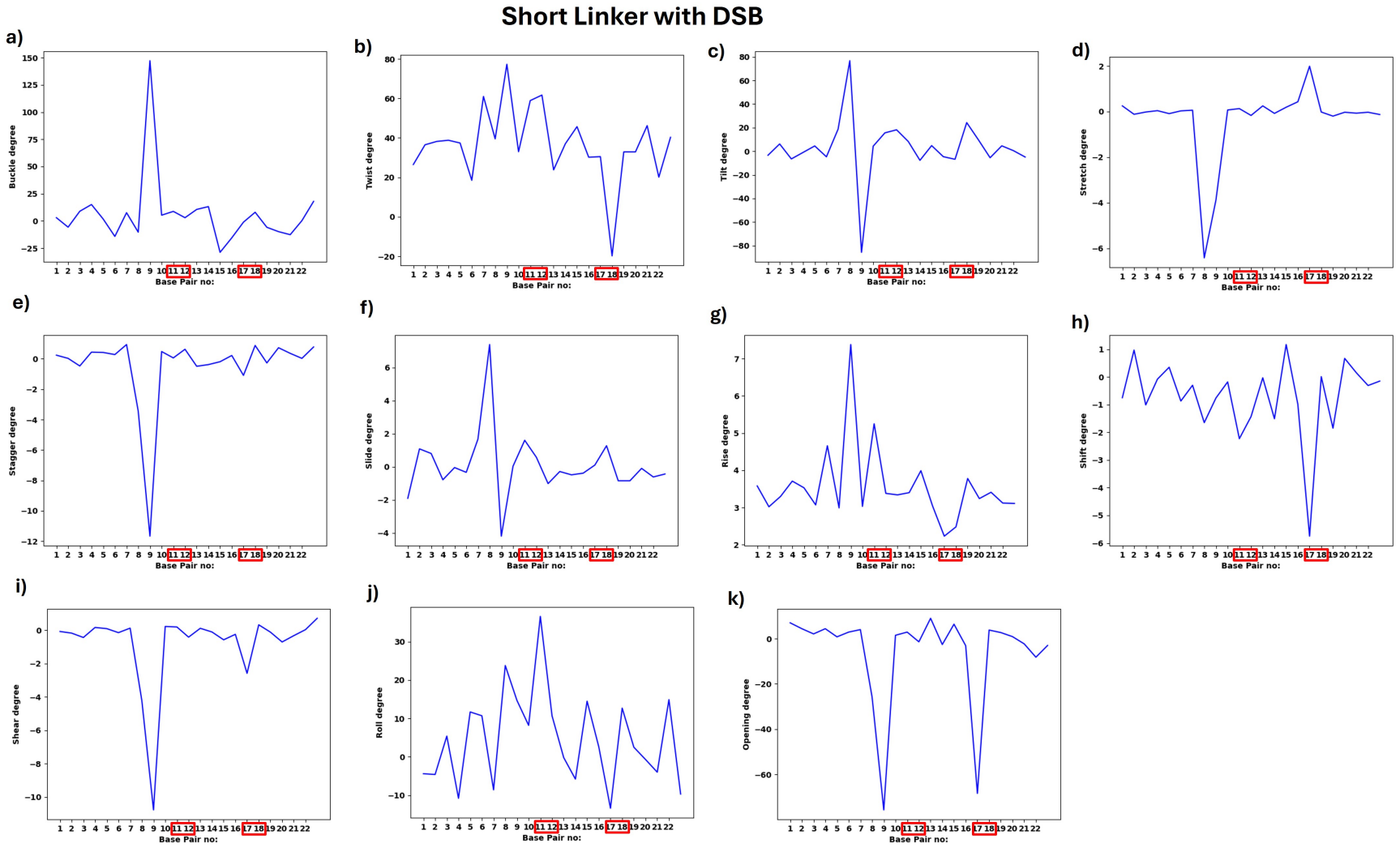


Figure A13: Different helical parameters for short linker with a DSB.

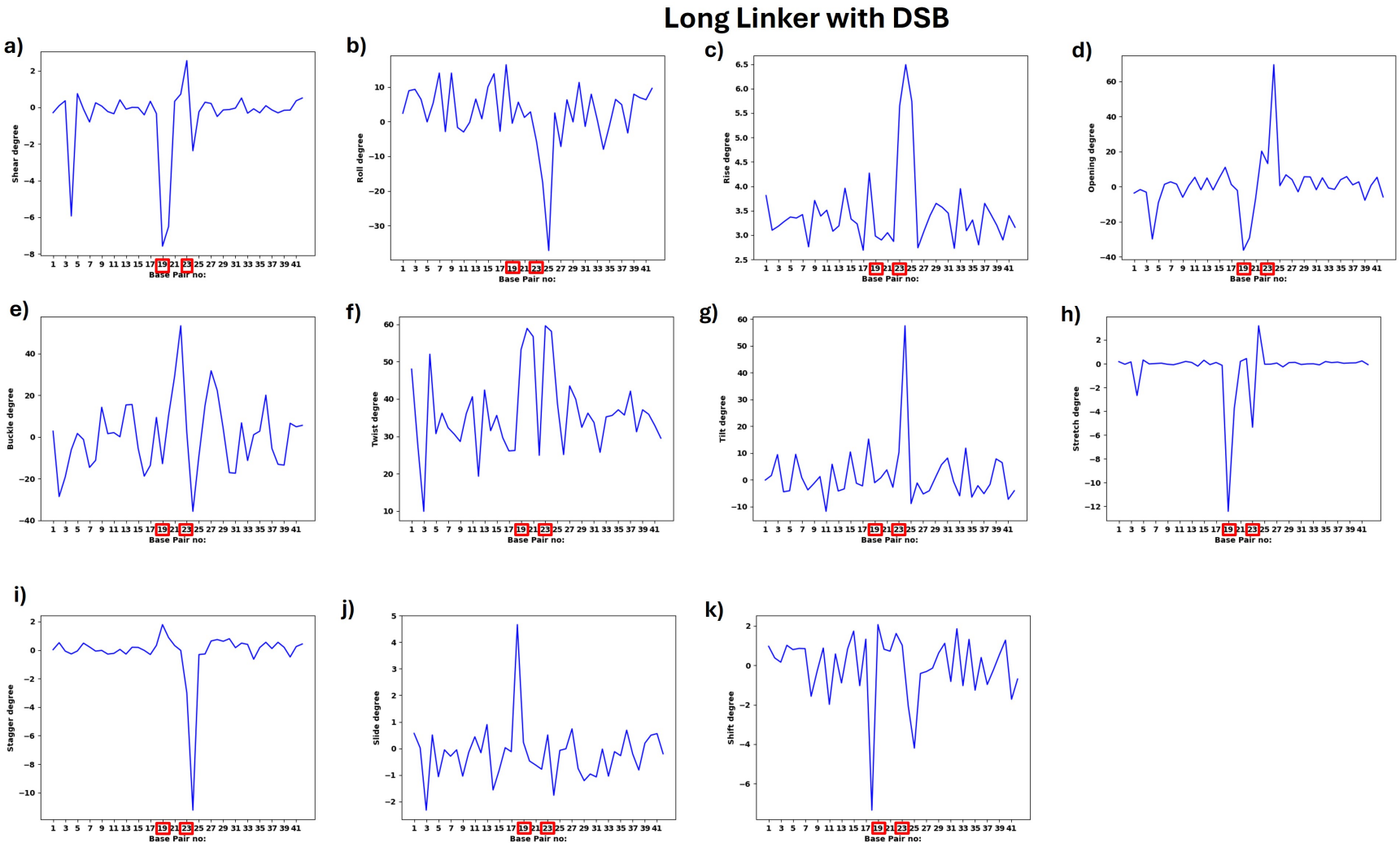


Figure A14: Different helical parameters for long linker with a DSB.

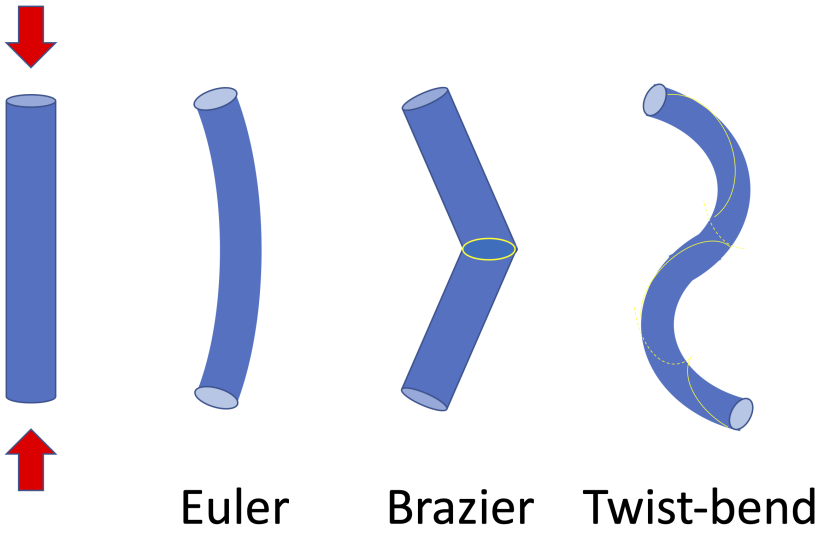


Figure A15: Mechanical instabilities in an axially compressed rod. The first bending instability (Euler) is encountered at small initial load: the axial symmetry is spontaneously broken by fluctuations, and the rod falls in a "mexican-hat"-shaped potential well. As the applied load increases the approximately circular DNA cross section may turn elliptic (or squeezed) at some point. This induces a second instability (Brazier), in which the rod develops a sharp kink. A bifurcation toward a twist-bend deformation mode can alternatively develop: for twisting to occur as a result of simple bending, the structure must have different bending stiffness along (at least) two directions not parallel to its main axis.The figure is taken from [141]

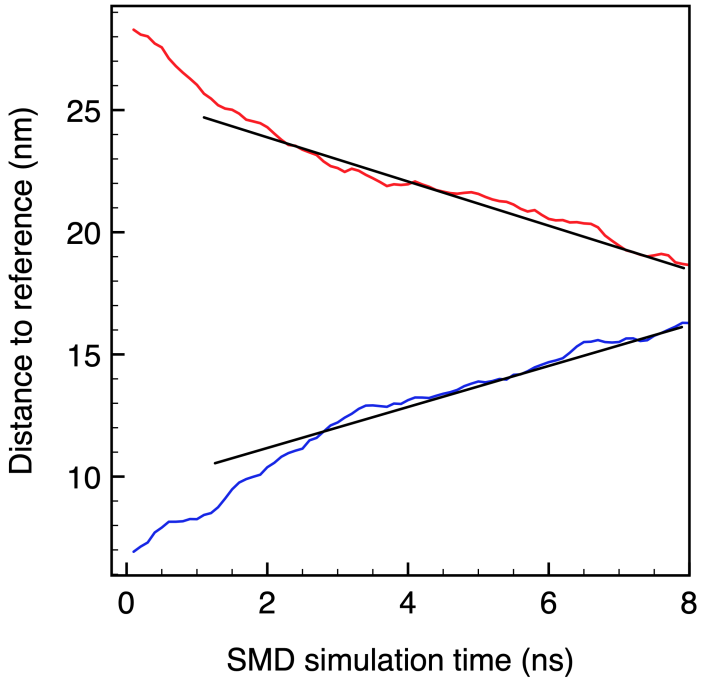


Figure A16: Evolution of distance between the two nucleosomes over time, under the applied force to their center of mass. The straight lines identify the linear-drag regime.

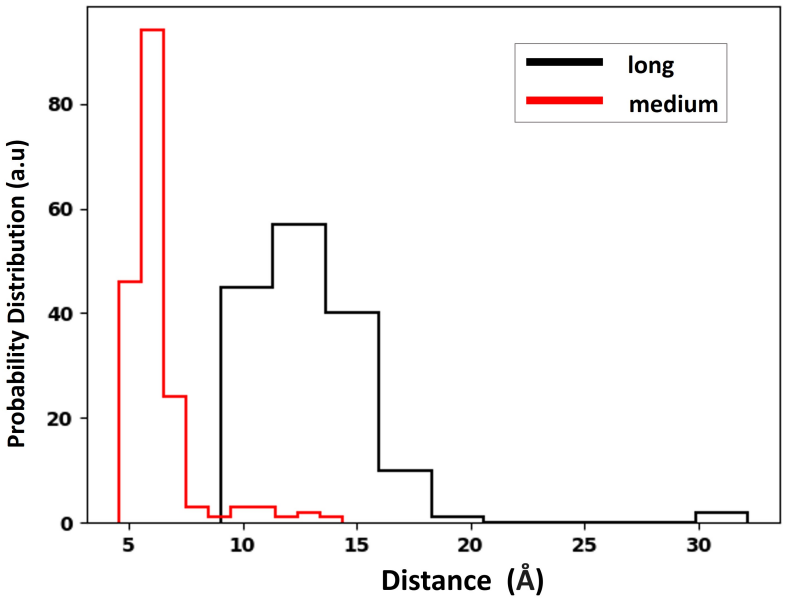


Figure A17: Probability distribution histogram of the gap-opening distance of the second strand break in DSB

LET (keV/μm)	Isolated Backbone	Isolated Base	nBases per Backbone Damage
1	548.9271	247.5059	0.45089
2	545.6036	248.9812	0.45634
3	542.3097	250.3751	0.46168
4	539.0452	251.6875	0.46691
5	535.8102	252.9183	0.47203
6	532.6046	254.0677	0.47703
7	529.4286	255.1356	0.48191
8	526.2820	256.1220	0.48666
9	523.1649	257.0269	0.49129
10	520.0773	257.8504	0.49579
11	517.0192	258.5923	0.50016
12	513.9905	259.2527	0.50439
13	510.9913	259.8317	0.50849
14	508.0216	260.3291	0.51244
15	505.0814	260.7451	0.51624
16	502.1706	261.0795	0.51990
17	499.2894	261.3325	0.52341
18	496.4376	261.5040	0.52676
19	493.6153	261.5940	0.52996
20	490.8224	261.6025	0.53299
21	488.0591	261.5295	0.53586
22	485.3252	261.3750	0.53856
23	482.6208	261.1390	0.54109
24	479.9459	260.8215	0.54344
25	477.3004	260.4226	0.54562
26	474.6845	259.9421	0.54761
27	472.0980	259.3801	0.54942
28	469.5410	258.7367	0.55104
29	467.0134	258.0118	0.55247
30	464.5154	257.2053	0.55371
31	462.0468	256.3174	0.55474
32	459.6077	255.3480	0.55558
33	457.1981	254.2971	0.55621
34	454.8180	253.1647	0.55663
35	452.4673	251.9508	0.55684
36	450.1461	250.6554	0.55683
37	447.8544	249.2785	0.55661
38	445.5922	247.8201	0.55616
39	443.3594	246.2803	0.55549
40	441.1562	244.6589	0.55459

Table 6: LET values and corresponding damage metrics

B

PROTOCOLS

B.1 MOLECULAR SIMULATION PROTOCOLS

B.1.1 *Molecular construction of the DNA hairpin*

The blueprint of a 45-nucleotide long RNA, with the same experimental DNA sequence of Ref.[128] and all thymines replaced by uracils, was submitted to the ROAD interactive website (RNA-Origami Automated Design tool (ROAD), a computer-aided design software to build automated 3D-structures of nucleic acids [133]). Then we substituted back uracils with thymines in the resulting 3D-structure, using CHIMERA [235]. The hairpin DNA structure comprises a missing nucleotide at the center, mimicking an SSB (see Figure 4d in Chapter 2), flanked at 3′ and 5′ by Guanine in positions 1 and 45, and an unpaired Thymine in position 23 facing the missing SSB nucleotide. We also made sure that the DNA forms a closed loops with 2 unpaired nucleotides in positions 11, 34 and 35, at both hairpin ends. This DNA structure was then utilized as a starting configuration, in parallel to the experimental one fro RCSB PDB **2N8A**, to elucidate the first steps of the PARP-1 Zn-finger interactions with the DNA. While there could be minor differences between RNA 3D structure and the putative ssDNA folded into hairpin, based on previous studies following this same protocol in our group [236] we are confident that the subsequent equilibration by MD at finite temperature is able to wash out any structural instability, and produces a stable DNA folded structure.

For the comparison with straight DNA, I used CHIMERA[235] to design a long stretch of free DNA with 100 bp long, with a missing nucleotide in one strand at position 101 to represent the SSB. Like for the hairpin, the break in this long DNA is also flanked by Guanine on both $5' - 3'$ sides, along with Thymine facing the missing nucleotide.

B.1.2 *Assembly of the Zn finger domains of PARP-1*

Zinc-finger proteins are among the most abundant proteins in eukaryotic genomes. Their functions are extraordinarily diverse and include DNA recognition, RNA packaging, transcriptional activation, regulation of apoptosis, protein folding and assembly, and lipid binding.[237]. The zinc-finger structure is maintained by a zinc ion, which coordinates cysteine and histidine in different combinations. In classical C2H2 zinc-finger proteins, two cysteines in one chain and two

histidines in another one are coordinated by a single zinc ion. Crystallographic studies revealed that classical zinc-finger domains have two β-sheets and one α-helix. The PARP-1 protein has three Zn-fingers, with the structure variant CCHC; Zn1 and Zn2 have the typical secondary structure 2β-1α. The necessary parametrization to include Zn-finger structures in AMBER is described below.

The structure of the Zn1-Zn2 finger region of PARP-1 was taken from the experimental NMR data [128], deposited in the RCSB PDB archive under the entry **2N8A**. For the sake of comparison, we also took the corresponding fragment from the AlphaFold2 repository, entry **P09874** (see Figure 4) [238, 239]. The structure obtained from **2N8A** consists mainly of the Zn1-Zn2 fragment of PARP-1 (residues 1-201). AlphaFold2 instead predicted the entire structure of PARP-1 with all its domains: in the AlphaFold2 model, all major domains of PARP-1 are comparable to the corresponding solution structure obtained from NMR studies within a reasonable confidence interval, except the disordered linker stretch of 13 amino acids connecting the Zn1 and Zn2 domains.

B.1.3 *Molecular structure of the mononucleosome*

The initial molecular configuration of the nucleosome is taken from the RSCB Protein Database, entry **1KX5** [56]. This is an x-ray structure of the entire histone octamer with 147 DNA base-pairs resolved at an average RMS of 1.94 Å, reconstituted from human nuclear extract expressed in *E. coli*; only 6 histone residues were unidentified in this experimental structure, with respect to the known histone sequences, which was completed with the Swiss-Model utility of SwissProt [240] to obtain a complete nucleosome including the full-length histone tails. This 147-bp DNA is a palindromic sequence, originally chosen to maximize the degree of ordering and increase the x-ray spatial resolution; to obtain a model structure useful for the MD simulations, I removed all the crystallization water molecules and ions from the published structure; DNA bases are numbered from 1 to 147 in each chain, one running clockwise and the other counter-clockwise, the dyad being located at basis 74 of each chain.

DNA is wrapped left-handed about the histone core, making two nearly complete turns that join at the dyad symmetry point; the two DNA turns define two circles lying in two ideally parallel planes, with a superhelical symmetry axis (SHL) perpendicular to the center of the circles; the relaxed DNA double helix makes a complete twist around its double-helical axis, about every 10.4 bp; correspondingly, 14 contact points between DNA and the histone core proteins can be identified within the nucleosome structure, loosely situated at the minor groove locations facing inwards. Based on these geometrical

features, the sites where to place DNA damage in the form of either a SSB or a DSB were defined, in structurally significant positions.

B.1.4 *Molecular structure of the dinucleosome*

I used two identical copies of the nucleosome **1KX5**. One end of a nucleosome was continued with a straight dsDNA of different lengths (20, 24 and 42 bp), and was joined at the end of the other nucleosome copy. All the adjustments to joining the DNA fragments together were performed using PyMol and Chimera softwares. The linker-DNA bp sequences are randomly chosen. Damage in the form of either a SSB, or a DSB, was introduced approximately in the middle of each linker-DNA, at the position at which a stretch of ds(TAAT) sequence is also located.

B.1.5 *Docking protocols*

Docking of proteins and nucleic acids is still a rapidly evolving field of research. We tested different methods available as online server platforms: HADDOCK [241, 242], HDOCK [243] and the very recent pyDockDNA [244]. In our analyses, only HADDOCK was capable of providing consistent results comparable to the available experimental information. This is evident in both the straight DNA hairpin and the bent DNA extracted from the experimental NMR structure, especially when assessed against the Zn1-Zn2 fragment of PARP-1. HADDOCK bootstrap parameters were chosen such that we could typically aim at the regions including the trademark interactions between Zn1/Zn2 and the SSB [128]. Best-fit clusters of candidate high-scoring poses were taken from the docking process.

Furthermore, we performed interpolation morphing between the straight and bent DNA conformations, using CHIMERA. Such morphing simulations provided a preliminary look at how the protein and DNA might change their conformations during the SSB recognition process. From the first step, it seems that no steric clashes are prohibiting the rearrangement towards the final (experimental) conformation. However, the results of docking displayed a large variety of candidate conformations, also including inverted contacts, that is Zn1 close to the 3′ end and Zn2 close to the 5′ end of the hairpin, contrary to the experimental evidence. Hence, starting from the docked configuration, we also tried to improve the contacts by refinement, adjusting the contacts by minimal rigid-body translations and rotations, with subsequently relaxing and cyclic annealing by MD of the "post-docked" configurations. These adjusted conformations will be also used in the following as candidate starting structures for the SSB recognition.

B.2 AMBER14 FORCEFIELD PARAMETERS

The MD simulations performed in this thesis utilizes AMBER14 force field, which already includes PARM-BCS1 extension for nucleic acids. But, it lacked the parameters required for Zinc ion interaction with the PARP1 protein. Also, the 5′ ends of the SSB gap is occupied by a dRP molecule rather than Phosphate molecule. Hence, both of these parameters were added to the forcefield file before running the simulation in GROMACS. The table below highlights the files that have been changed and the parameters being added.

Zinc Chemistry from Macchiagodena et al.[245]: parameters added to the non-bonded force field parameters file **ffnonbonded.itp**

Name	Atomic Num	Mass	Charge	Ptype	Sigma	Epsilon
NZ (Nitrogen)	7	14.01	0.0000	A	3.35000e-01	7.11280e-01
SZ (Sulphur)	16	32.06	0.0000	A	3.66359e-01	1.04600e+00
OZ (Oxygen)	8	16.00	0.0000	A	3.05992e-01	8.78640e-01
Zn (Zinc)	30	65.40	0.0000	A	2.59452e-01	5.23000e-02

Table 7: Atomic properties of different elements from Ref.[245]

5′ **dRP residue**: this special residue consists of a deoxyribose molecule with a PO_4^{3-} at its terminus. To add this residue to the force field parameter, follow these steps:

1. Make a new set of parameters in the dna.rtp file (which is residue type file for DNA) for the new nucleotide. Start by naming the nucleotide differently. For example, I modified Guanine to contain a dRP end, thereby named as DGP.

2. Copy all the parameters in the [DG] set in the dna.rtp file and paste it beneath the newly created [DGP] heading.

3. Now to add extra Oxygen to the PO_3 specifications, add new atoms in the atom section beneath [DGP]. Then utilize the value for O1P and name it as O3P.

4. Next step is to add a hydrogen atom to saturate the oxygen, make a new parameter namely H5P, which have same values as any other hydrogen combined with Oxygen.

5. The next step is to create bonds between the added new atoms. Those will include a bond P–O1P, P–O2P, P–O3P and O3P–H5P.

6. The finalized list of parameters beneath [DGP] looks like the one in the table 9

7. Next step is to add the parameters in hydrogen database file **dna.htb**, **ffbonded.itp** and **ffnonbonded.itp** parameter files. In

dna.hdb add the 4 atoms that the new hydrogen have bonds with, while the ffbonded.itp file will include the bond length, dihedral values and angles between the newly created atoms. The absence of any of one of these parameters can result in a fatal error during running GROMACS commands. Hence, the user must know the parameters that are missing and needed to be added to the respective files. The parameters added to the HDB files are given in the table 8.

8. The next step is to include the name of the nucleotide [DGP] in the **residuetypes.atp** file as a DNA.

9. The last step is to modify the naming of the dRP containing nucleotide I(in this case, one or more Guanine) in your pdb with DGP. This makes sure that GROMACS considers the values corresponding to a DGP nucleotide while compiling.

DGP	9					
1	2	H5P	O3P	P	O2P	
2	6	H5′	C5′	O5′	C4′	
1	5	H4′	C4′	C5′	O4′	C3′
1	5	H1′	C1′	O4′	N9	C2′
1	1	H8	C8	N9	N7	
1	1	H1	N1	C6	C2	
2	3	H2	N2	C2	N1	
1	5	H3′	C3′	C4′	C2′	O3′
2	6	H2′	C2′	C1′	C3′	

Table 8: DGP parameters in dna.hdb file

B.2.1 *Molecular dynamics simulations and analysis*

All the MD simulations and most data analysis, were performed with the GROMACS 2020.4 package [174, 175]. The Amber14 force-field was used in the simulations [246], with extra parameters for Zn ion interaction from Macchiagodena et al. [245]; notably, Amber14 already includes the lastest PARM-BCS1 extension for nucleic acids [247]. To ensure the primary stability of the post-docked configurations, we first performed energy-minimization using a steepest-descent algorithm, followed by {NVT}-{NPT} equilibration for 150ps, to bring the systems at 310K and 1 bar pressure. The ensembles of DNA-proteins were solvated in water boxes of different sizes, according to the model systems used, along with Na^+ and Cl^- ions for neutralizing the system charge and maintaining a physiological 0.1M salt concentration.

In some instances it was necessary to perform rapid annealing and quenching cycles, to improve the mutual positions of protein and

[atom]			
P	P	1.16590	1
O1P	O2	-0.77610	2
O2P	O2	-0.77610	3
O3P	O2	-0.77610	4
H5P	HO	0.44220	5
O5′	OS	-0.49540	6
C5′	CI	-0.00690	7
H5′1	H1	0.07540	8
H5′2	H1	0.07540	9
C4′	CT	0.16290	10
H4′	H1	0.11760	11
O4′	OS	-0.36910	12
C1′	CT	0.03580	13
H1′	H2	0.17460	14
N9	N*	0.05770	15
C8	CK	0.07360	16
H8	H5	0.19970	17
N7	NB	-0.57250	18
C5	CB	0.19910	19
C6	C	0.49180	20
O6	O	-0.56990	21
N1	NA	-0.50530	22
H1	H	0.35200	23
C2	CA	0.74320	24
N2	N2	-0.92300	25
H21	H	0.42350	26
H22	H	0.42350	27
N3	NC	-0.66360	28
C4	CB	0.18140	29
C3′	CE	0.07130	30
H3′	H1	0.09850	31
C2′	CT	-0.08540	32
H2′1	HC	0.07180	33
H2′2	HC	0.07180	34
O3′	OS	-0.52320	35

[bonds]	
P	O1P
P	O2P
P	O3P
P	O5′
O3P	H5P
O5′	C5′
C5′	H5′1
C5′	H5′2
C5′	C4′
C4′	H4′
C4′	O4′
C4′	C3′
O4′	C1′
C1′	H1′
C1′	N9
C1′	C2′
N9	C8
N9	C4
C8	H8
C8	N7
N7	C5
C5	C6
C5	C4
C6	O6
C6	N1
N1	H1
N1	C2
C2	N2
C2	N3
N2	H21
N2	H22
N3	C4
C3′	H3′
C3′	C2′
C3′	O3′
C2′	H2′1
C2′	H2′2

[impropers]			
C4	C8	N9	C1′
C5	N1	C6	O6
C6	C2	N1	H1
C2	H21	N2	H22
N9	N7	C8	H8
N2	N1	C2	N3

Table 9: DGP parameters in the dna.rtp file

DNA, for example after a manual adjustment necessary to remove a steric clash. Each cycle was typically performed by ramping up the temperature from 310K to 400K, and then back to 310K, in steps of 10K for 500 ps each. Usually, during the annealing cycle the DNA was frozen in its conformation and only the protein was left free to adjust.

Coulomb forces were summed with particle-mesh Ewald sum (PME), using a real-space cutoff of 1.2 nm, equal to the cut-off radius of shifted Van der Waals potentials. Rigid bonds for the water molecules were used, which allowed integration of the equations of motion with a time step of 2 fs for the thermal equilibration phases, and 1 or 2 fs for production runs.

Trajectory clustering analysis was performed using a special subprogram of GROMACS. The MD simulation stores 'frames' containing the positions, velocities and forces of all particles in the system, at prescribed intervals (typically every 10 to 50 ps, or longer: a 1 µs-long MD run can store as much as 100,000 frames of about 2Mbytes each, resulting in data files with size of hundreds of Gbytes); the subprogram calculates a matrix of root-mean-squared displacements (RMSD) between each pair of frames, by comparing the positions of each atom in the pair; then, RMSD values are grouped according to a cut-off criterion, and clusters of similar frames, typically separated by a small enough RMSD, are detected. The cluster(s) with the highest number of members, and those which are more frequently sampled over the entire simulation time, provide an indication of the 'best-average' molecular configurations, thus allowing to bypass the noise of the short-time atomic fluctuations.

Free energy analysis and contact surface calculations were estimated by the PDBePisa web-based utility [248], by extracting a few selected configurations in PDB format from the MD trajectory. The resulting values are approximated, however they can provide a qualitative appraisal of the binding energies involved in the protein-DNA interaction.

B.3 GROMACS PULL CODE

The GROMACS already has an existing pull-code with various option to pull a group in the system, using various algorithms for force application. The code applies a force or constraints on the user-defined collective variables. These forces acts on a fictitious spring attached to the to the pulling atom (that could be the center of mass of a whole group).

With the "umbrella pulling" that was used in this work, a harmonic potential (a "spring") is applied between the centers of mass of two groups; therefore, the force is proportional to the displacement. It is important to realize that in all cases, the moving end of the fictitious spring has a displacement in time that does not necessarily coincide with that of the pulled group of atoms. Namely, if the spring constant is too large compared to the intrinsic elasticity of the atom group, what is actually measured is the force-displacement relation on the spring and not on the atoms; if, on the other hand, the spring constant is too soft, the atoms' movement will be erratic, and usually follow the

displacement of the spring in sudden steps. The value of the pulling spring constant and velocity must therefore be calibrated according to each different situation.

The typical run of the pull code applied to our dinucleosome system looks like the list given below (the "pbc atom" for each variation changes according to the system); group 1 chain A and group2 chain B correspond to the two whole nucleosomes, while group 3 chain C corresponds to the linker DNA.

```
;Pull_code
pull = yes
pull-ncoords = 2
pull-ngroups = 3
pull-group1-name = Chain_A
pull-group2-name = Chain_B
pull-group3-name = Chain_C
pull-group1-pbcatom = 13315
pull-group2-pbcatom = 29055
pull-group3-pbcatom = 36752
pull-pbc-ref-prev-step-com = yes
pull-print-com = yes

;Pulling_first_nucleosome
pull-coord1-groups = 3 1
pull-coord1-type = umbrella
pull-coord1-geometry = direction
pull-coord1-vec = 0 1 0
pull-coord1-dim = N Y N
pull-coord1-rate = 0.0005
pull-coord1-k = 10

;Pulling_second_nucleosome
pull-coord2-groups = 3 2
pull-coord2-type = umbrella
pull-coord2-geometry = direction
pull-coord2-vec = 0 -1 0
pull-coord2-dim = N Y N
pull-coord2-rate = 0.0005
pull-coord2-k = 10
```

B.4 COPASI: COmplex PAthway SImulator

COmplex PAthway SImulator is a software application designed by Hoops et. al to simulate complex and elaborate biological networks and its dynamics[197]. It addresses the growing need in the biological community for software tools that allow access to diverse simulation and modeling methods as well as support for the usage of these methods. The standard methods used in the field comprises mainly of:

- Deterministic approach - integration of ordinary differential equations (ODEs).
- Stochastic simulation of reaction networks[249].
- Computation of steady states and their stability (e.g: Computing elementary nodes)[250].
- Sensitivity analysis & Metabolic control analysis[251, 252]
- Optimization and parameter estimation.

One of the key strengths of COPASI lies in its extensive range of features, which cater to a wide array of biochemical network analyses. The software supports both deterministic and stochastic simulations, allowing researchers to select the most suitable method for their specific study. This dual capability is particularly valuable when dealing with systems where either type of simulation might be appropriate depending on the biological context. Additionally, COPASI includes functionalities for parameter estimation, enabling users to refine their models based on experimental data, and metabolic control analysis, which helps identify control points within metabolic pathways. The software also supports elementary flux mode analysis, providing insights into the possible pathways through a biochemical network. This versatility makes COPASI a highly adaptable tool for various research needs in systems biology.

The user interface of COPASI is designed to be intuitive and accessible, lowering the barrier to entry for users with varying levels of expertise in computational biology. The graphical user interface (GUI) features drag-and-drop capabilities for constructing biochemical networks, along with visual tools for mapping and analyzing these networks. This visual approach not only simplifies the process of setting up simulations but also aids in the interpretation of results. Comprehensive documentation and tutorials are provided to assist users in maximizing the software's potential, ensuring that researchers can effectively leverage COPASI for their research.

B.4.1 *Kinetic Models in COPASI*

Hill Kinetics: Hill's equation realistically describes many transcriptional level functions. It describes both activation and inhibition kinetics and is derived from considering the equilibrium binding of the transcription factor to its promoter site. The function is dependent on three parameters: K, the activation coefficient with units of concentration; β, the maximal promoter activity; and n the coefficient governing the step-like input function, the larger n the more steep is the sigmoid shape of the funciton [253].

If one considers a pathway leading from $X \rightarrow Y$, the rate of production of Y can be written as function, $Y = f(X)$ taking the following form:

$$f(X) = \pm\beta \cdot \left(\frac{A^n}{K^n + X^n} \right) \tag{6}$$

the '+' sign and $A = X$ holding for activation, and the '–' sign, $A = K$ for inhibition. Typical values of the steepness coefficient are $n = 1 - 4$.

Michaelis-Menten Kinetics: M.M. kinetics is a model of enzyme kinetics describing the relationship between enzyme and the substrate, in an enzyme-catalyzed reaction. It also depends on two parameters, V_{max}, the maximum rate of the reaction; and K_m, the M.M. constant. Also, it assumes the reaction is at steady state.

The equation is derived from a reversible equation between enzyme, substrate, product, and their reaction rates: an enzyme E binding to a substrate S to form a complex ES, which in turn releases a product P regenerating the original form of the enzyme, as

$$E + S \rightleftharpoons ES \rightarrow P + E. \tag{7}$$

The final expression for the reaction rate, depending on the substrate concentration $[S]$, is:

$$V_0 = V_{max} \cdot \left(\frac{[S]}{K_m + [S]} \right) \tag{8}$$

Mass-action kinetics: Describes the behavior of reactants and products in any chemical reaction, in which the rate of a chemical reaction is directly proportional to the concentration of reactants. For example, for a reaction $A + B \rightarrow C$, for concentrations $[A]$ and $[B]$ it is:

$$\text{Rate of reaction} \propto [A] \cdot [B] \tag{9}$$

$$\frac{d[C]}{dt} = k[A][B] \tag{10}$$

B.4.2 *Differential equations for the entire out-of-field dose-triggered SSBR cascade pathway*

1. **Rate of Change of ROS:**

$$\frac{d([ROS]\cdot V_1)}{dt} = -V_1\cdot\left(k1_{(4)}\cdot[ROS]\right) + V_1\cdot\left(k1_{(1)}\cdot[IR]\right) - \\ -V_1\cdot\left(k1_{(1.1)}\cdot([ROS]-[ROS_1])\right) - V_1\cdot\frac{k_{(5)}\cdot K_{(5)}}{K_{(5)}+[ROS]}$$

2. **Rate of Change of SSB:**

$$\frac{d([SSB]\cdot V_1)}{dt} = +V_1\cdot\left(k1_{(4)}\cdot[ROS]\right) - V_1\cdot\left(k1_{(10)}\cdot[SSB]\right) - \\ -V_1\cdot\left(k_{(6)}\cdot[SSB]\cdot[PARP1]\right)$$

3. **Rate of Change of Unrepaired SSBs:**

$$\frac{d([SSBir]\cdot V_1)}{dt} = +V_1\cdot\left(k1_{(10)}\cdot[SSB]\right) - V_1\cdot\left(k_{(13)}\cdot[SSBir]\cdot[XRCC1]\right)$$

4. **Rate of Change of PARP1:**

$$\frac{d([PARP1]\cdot V_1)}{dt} = +V_1\cdot\frac{k_{(7)}\cdot[Enz]}{K_{(7)}+[Enz]} - V_1\cdot\frac{k_{(8)}\cdot[PARP1]}{K_{(8)}+[PARP1]} + V_1\cdot\left(k_{(6)}\cdot[SSB]\cdot[PARP1]\right) + \\ +V_1\cdot\frac{k_{(5)}\cdot K_{(5)}}{K_{(5)}+[ROS]}$$

5. **Rate of Change of SSBir & XRCC1 complex:**

$$\frac{d([SSBirXRCC1]\cdot V_1)}{dt} = -V_1\cdot\frac{k_{(14)}\cdot[SSBirXRCC1]}{K_{(14)}+[SSBirXRCC1]} + V_1\cdot\left(k_{(13)}\cdot[SSBir]\cdot[XRCC1]\right)$$

6. **Rate of Change of XRCC1:**

$$\frac{d([XRCC1]\cdot V_1)}{dt} = +V_1\cdot\left(k1_{(12)}\cdot[PARS]\right) - V_1\cdot\left(k_{(13)}\cdot[SSBir]\cdot[XRCC1]\right)$$

7. **Rate of Change of PARS:**

$$\frac{d([PARS]\cdot V_1)}{dt} = +V_1\cdot\frac{k_{(9)}\cdot K_{(9)}}{K_{(9)}+[PARG]} + V_1\cdot\frac{k_{(8)}\cdot[PARP1]}{K_{(8)}+[PARP1]} - V_1\cdot\left(k1_{(12)}\cdot[PARS]\right)$$

8. **Rate of Change of Enzymes:**

$$\frac{d([Enz] \cdot V_1)}{dt} = -V_1 \cdot \frac{k_{(7)} \cdot [Enz]}{K_{(7)} + [Enz]} + V_1 \cdot \frac{k_{(3)} \cdot K_{(3)}}{K_{(3)} + [IR]}$$

9. **Rate of Change of PARG:**

$$\frac{d([PARG] \cdot V_1)}{dt} = +V_1 \cdot \left(k1_{(2)} \cdot [IR]\right) - V_1 \cdot \frac{k_{(9)} \cdot K_{(9)}}{K_{(9)} + [PARG]}$$

10. **Rate of Change of p38MAPK:**

$$\frac{d([p38MAPK] \cdot V_1)}{dt} = +V_1 \cdot \frac{k_{(14)} \cdot [SSBirXRCC1]}{K_{(14)} + [SSBirXRCC1]} - V_1 \cdot \left(k1_{(15)} \cdot [p38MAPK]\right) - \\ -V_1 \cdot \left(k1_{(16)} \cdot [p38MAPK]\right)$$

11. **Rate of Change of BAX:**

$$\frac{d([BAX] \cdot V_1)}{dt} = +V_1 \cdot \left(k1_{(17)} \cdot [p53]\right) - V_1 \cdot \left(k1_{(20)} \cdot [BAX]\right) - V_1 \cdot \left(k_{(19)} \cdot [p16]\right)$$

12. **Rate of Change of p53:**

$$\frac{d([p53] \cdot V_1)}{dt} = +V_1 \cdot \left(k1_{(15)} \cdot [p38MAPK]\right) - V_1 \cdot \left(k1_{(17)} \cdot [p53]\right)$$

13. **Rate of Change of Apoptosis**

$$\frac{d([Apop] \cdot V_1)}{dt} = +V_1 \cdot \left(k1_{(20)} \cdot [BAX]\right)$$

14. **Rate of Change of p16:**

$$\frac{d([p16] \cdot V_1)}{dt} = +V_1 \cdot \left(k1_{(18)} \cdot [ROS]\right)$$

15. **Rate of Change of Rb:**

$$\frac{d([Rb] \cdot V_1)}{dt} = +V_1 \cdot \left(k1_{(19)} \cdot [p16]\right)$$

16. **Rate of Change of Senescence:**

$$\frac{d([Senescence] \cdot V_1)}{dt} = +V_1 \cdot \left(k1_{(21)} \cdot [Rb]\right)$$

B.5 SDD format for TOPAS

Field	Value	Notes	Type
1	SDD version	Version number of SDD definition	String (“SDDv1.0”)
2	Software	Program name, version, and access link if any	String (free text)
3	Author	Corresponding author, date, references	String (free text)
4	Simulation details	Description of simulation settings and parameters	String (free text)
5	Source	Description of source properties	String (free text)
6	Source type	Monoenergetic, distribution, phase space, GCR, etc.	Int
7	Incident particles	Definition of primary incident irradiation particle(s) in PDG code format	Int(s)
8	Mean particle energy	Mean incident energy for each particle in MeV	Float(s)
9	Energy distribution	Full energy distribution specification	String(s) + Floats
10	Particle fraction	Fraction of fluence of each particle in field	Float(s)
11	Dose or fluence	Define dose or fluence in each exposure, or note that the simulation was for a single track	Int + Float (+Float)
12	Dose rate	Dose rate of irradiation field	Float
13	Irradiation target	Description of simulated cell or target (DNA) region and microenvironment	String (free text)
14	Volumes	Shape parameter plus X, Y, Z extents (μm)	2×(Int + 6 Float)
15	Chromosome sizes	Number and base pair size of chromosomes	Int + (Int) Floats
16	DNA density	Density of base pairs in volume (MBP/μm^3)	Float
17	Cell cycle phase	Cell cycle phase index and progression	Int + Float
18	DNA structure	Additional field to define DNA structure	2 Ints
19	In vitro/in vivo	Experimental condition	Int
20	Proliferation status	Proliferative or quiescent and status details	Int + String (free text)
21	Microenvironment	Temperature (°C) and molar O_2 concentration	2 Float
22	Damage definition	Define how types of damage were determined	1 Float + 1 Int + 1 Bool + 2 Float
23	Time	Time point at which damages were recorded	Float
24	Damage and primary count	Number of distinct damage lesions scored and primaries simulated	2 Int
25	Data entries	Number of fields included in the data section	14 Bool
26	Additional information	Field for additional relevant information	String (free text)
27	***EndOfHeader***	Empty field to mark end of header	—

Table 10: Field by Field Summary of the Header Fields and their Type

Field	Value	Notes	Type	Req?
1	Classification	Is damage associated with new primary particle or new exposure? And event ID.	2 Ints	Y, N
2	X, Y, Z	Spatial X, Y, Z coordinates and extent (μm)	3×3 Floats	*
3	Chromosome IDs	ID of chromosome/chromatid where damage occurred and on which arm (long/short) or specification of non-nuclear DNA type.	4 Int	*
4	Chromosome position	Location of damage within chromosome	Float	*
5	Cause	Cause of damage - direct or indirect and number	3 Int	N
6	Damage types	Types of damage at site (Base damage, SSB, DSB)	3 Int	**
7	Full break spec	Full description of strand break structure	Special	**
8	DNA sequence	DNA Base Sequence around break site	Special	N
9	Lesion time	Time of each damage induction (in ns)	Special	N
10	Particle types	PDG list of particles	Int(s)	N
11	Energies	List of kinetic energies for each particle (in MeV)	Float(s)	N
12	Translation	Starting position of each particle (in μm)	Floats	N
13	Direction	Starting direction of each particle (unit vector)	Floats	N
14	Particle time	Starting time of each particle (in ns)	Floats	N

Table 11: Value by Value Definition of the Data Fields to Score DNA Damages

B.6 FINALIZED TIME CONSTANT FOR SSBR PATHWAY

The input text format of the entire SSBR pathway is given below:

```
i:Ch/DaMaRiS/numberOfConnections = 11

#Change 1
s:Ch/DaMaRiS/ProteinKinetic1/changeFrom = "DSBEnd"
sv:Ch/DaMaRiS/ProteinKinetic1/changeTo = 1 "DSBEnd_PARP1"
d:Ch/DaMaRiS/ProteinKinetic1/changeTime = 0.85 s
b:Ch/DaMaRiS/ProteinKinetic1/requireClean = "false"
b:Ch/DaMaRiS/ProteinKinetic1/cleanBackbone = "false"
b:Ch/DaMaRiS/ProteinKinetic1/cleanBase = "false"

#Change 2
s:Ch/DaMaRiS/ProteinKinetic2/changeFrom = "DSBEnd_PARP1"
sv:Ch/DaMaRiS/ProteinKinetic2/changeTo = 1 "DSBEnd_XRCC1"
d:Ch/DaMaRiS/ProteinKinetic2/changeTime = 5.95 s
b:Ch/DaMaRiS/ProteinKinetic2/requireClean = "false"
b:Ch/DaMaRiS/ProteinKinetic2/cleanBackbone = "false"
b:Ch/DaMaRiS/ProteinKinetic2/cleanBase = "false"
```

```
#Change 3
s:Ch/DaMaRiS/ProteinKinetic3/changeFrom = "DSBEnd_XRCC1"
sv:Ch/DaMaRiS/ProteinKinetic3/changeTo = 1 "DSBEnd_Lig3"
d:Ch/DaMaRiS/ProteinKinetic3/changeTime = 10 s
b:Ch/DaMaRiS/ProteinKinetic3/requireClean = "false"
b:Ch/DaMaRiS/ProteinKinetic3/cleanBackbone = "false"
b:Ch/DaMaRiS/ProteinKinetic3/cleanBase = "false"

#Change 3
s:Ch/DaMaRiS/ProteinKinetic4/changeFrom = "DSBEnd_Lig3"
sv:Ch/DaMaRiS/ProteinKinetic4/changeTo = 1 "DSBEnd_Scaffold1"
d:Ch/DaMaRiS/ProteinKinetic4/changeTime = 0.0 s
b:Ch/DaMaRiS/ProteinKinetic4/requireClean = "false"
b:Ch/DaMaRiS/ProteinKinetic4/cleanBackbone = "false"
b:Ch/DaMaRiS/ProteinKinetic4/cleanBase = "false"

#Change 5
s:Ch/DaMaRiS/ProteinKinetic5/changeFrom = "DSBEnd_Scaffold1"
sv:Ch/DaMaRiS/ProteinKinetic5/changeTo = 1 "DSBEnd_APE1"
d:Ch/DaMaRiS/ProteinKinetic5/changeTime = 52.0 s
b:Ch/DaMaRiS/ProteinKinetic5/requireClean = "false"
b:Ch/DaMaRiS/ProteinKinetic5/cleanBackbone = "false"
b:Ch/DaMaRiS/ProteinKinetic5/cleanBase = "false"

#Change 6
s:Ch/DaMaRiS/ProteinKinetic6/changeFrom = "DSBEnd_APE1"
sv:Ch/DaMaRiS/ProteinKinetic6/changeTo = 1 "DSBEnd_Scaffold2"
d:Ch/DaMaRiS/ProteinKinetic6/changeTime = 0.0 s
b:Ch/DaMaRiS/ProteinKinetic6/requireClean = "false"
b:Ch/DaMaRiS/ProteinKinetic6/cleanBackbone = "false"
b:Ch/DaMaRiS/ProteinKinetic6/cleanBase = "false"

#Change 6
s:Ch/DaMaRiS/ProteinKinetic7/changeFrom = "DSBEnd_Scaffold2"
sv:Ch/DaMaRiS/ProteinKinetic7/changeTo = 1 "DSBEnd_Scaffold1"
d:Ch/DaMaRiS/ProteinKinetic7/changeTime = 5.0 s
b:Ch/DaMaRiS/ProteinKinetic7/requireClean = "false"
b:Ch/DaMaRiS/ProteinKinetic7/cleanBackbone = "false"
b:Ch/DaMaRiS/ProteinKinetic7/cleanBase = "false"

#Change 7
s:Ch/DaMaRiS/ProteinKinetic8/changeFrom = "DSBEnd_Scaffold1"
sv:Ch/DaMaRiS/ProteinKinetic8/changeTo = 1 "DSBEnd_PolB"
d:Ch/DaMaRiS/ProteinKinetic8/changeTime = 30.0 s
b:Ch/DaMaRiS/ProteinKinetic8/requireClean = "false"
b:Ch/DaMaRiS/ProteinKinetic8/cleanBackbone = "false"
```

```
b:Ch/DaMaRiS/ProteinKinetic8/cleanBase = "true"

#Change 9
s:Ch/DaMaRiS/ProteinKinetic9/changeFrom = "DSBEnd_PolB"
sv:Ch/DaMaRiS/ProteinKinetic9/changeTo = 1 "DSBEnd_Scaffold3"
d:Ch/DaMaRiS/ProteinKinetic9/changeTime = 0.0 s
b:Ch/DaMaRiS/ProteinKinetic9/requireClean = "false"
b:Ch/DaMaRiS/ProteinKinetic9/cleanBackbone = "false"
b:Ch/DaMaRiS/ProteinKinetic9/cleanBase = "true"

#Change 10
s:Ch/DaMaRiS/ProteinKinetic10/changeFrom = "DSBEnd_Scaffold3"
sv:Ch/DaMaRiS/ProteinKinetic10/changeTo = 1 "DSBEnd_Scaffold1"
d:Ch/DaMaRiS/ProteinKinetic10/changeTime = 80.0 s
b:Ch/DaMaRiS/ProteinKinetic10/requireClean = "false"
b:Ch/DaMaRiS/ProteinKinetic10/cleanBackbone = "false"
b:Ch/DaMaRiS/ProteinKinetic10/cleanBase = "true"

#Change 11
s:Ch/DaMaRiS/ProteinKinetic11/changeFrom = "DSBEnd_Scaffold1"
sv:Ch/DaMaRiS/ProteinKinetic11/changeTo = 1 "DSB_Fixed"
d:Ch/DaMaRiS/ProteinKinetic11/changeTime = 150.0 s
b:Ch/DaMaRiS/ProteinKinetic11/requireClean = "false"
b:Ch/DaMaRiS/ProteinKinetic11/cleanBackbone = "true"
b:Ch/DaMaRiS/ProteinKinetic11/cleanBase = "false"

i:Ch/DaMaRiS/numberOfReactions = 3

#Reaction 1
sv:Ch/DaMaRiS/Reaction1/reactants = 2 "DSBEnd_XRCC1"
"DSBEnd_Lig3"
sv:Ch/DaMaRiS/Reaction1/products = 1 "DSBEnd_Scaffold1"
d:Ch/DaMaRiS/Reaction1/reactionRange = 25.0 nm

#Reaction 2 sv:Ch/DaMaRiS/Reaction2/reactants = 2 "DSBEnd_APE1"
"DSBEnd_Scaffold1"
sv:Ch/DaMaRiS/Reaction2/products = 1 "DSBEnd_Scaffold2"
d:Ch/DaMaRiS/Reaction2/reactionRange = 25.0 nm

#Reaction 3 sv:Ch/DaMaRiS/Reaction3/reactants = 2 "DSBEnd_PolB"
"DSBEnd_Scaffold1"
sv:Ch/DaMaRiS/Reaction3/products = 1 "DSBEnd_Scaffold3"
d:Ch/DaMaRiS/Reaction3/reactionRange = 25.0 nm
```

BIBLIOGRAPHY

BIBLIOGRAPHY

[1] Alexis Carrel and Albert H Ebeling. "Age and multiplication of fibroblasts." In: *The Journal of experimental medicine* 34.6 (1921), pp. 599–623.

[2] Lijing Jiang. "Alexis Carrel's Immortal Chick Heart Tissue Cultures (1912-1946)." In: *Embryo Project Encyclopedia* (2012).

[3] Leonard Hayflick and Paul S Moorhead. "The serial cultivation of human diploid cell strains." In: *Experimental cell research* 25.3 (1961), pp. 585–621.

[4] Calvin B Harley, A Bruce Futcher, and Carol W Greider. "Telomeres shorten during ageing of human fibroblasts." In: *Nature* 345.6274 (1990), pp. 458–460.

[5] James D Watson. "Origin of concatemeric T7DNA." In: *Nature New Biology* 239.94 (1972), pp. 197–201.

[6] Titia De Lange. "How telomeres solve the end-protection problem." In: *Science* 326.5955 (2009), pp. 948–952.

[7] Justyna Mikuła-Pietrasik, Arkadiusz Niklas, Paweł Uruski, Andrzej Tykarski, and Krzysztof Książek. "Mechanisms and significance of therapy-induced and spontaneous senescence of cancer cells." In: *Cellular and Molecular Life Sciences* 77 (2020), pp. 213–229.

[8] Ruchi Kumari and Parmjit Jat. "Mechanisms of cellular senescence: cell cycle arrest and senescence associated secretory phenotype." In: *Frontiers in cell and developmental biology* 9 (2021), p. 645593.

[9] Ewa Sikora, Grazyna Mosieniak, and Malgorzata Alicja Sliwinska. "Morphological and functional characteristic of senescent cancer cells." In: *Current drug targets* 17.4 (2016), pp. 377–387.

[10] Akshaya Chandrasekaran, Maria del Pilar Sosa Idelchik, and J Andrés Melendez. "Redox control of senescence and age-related disease." In: *Redox biology* 11 (2017), pp. 91–102.

[11] Joe Nassour, Sébastien Martien, Nathalie Martin, Emeric Deruy, Elisa Tomellini, Nicolas Malaquin, Fatima Bouali, Laure Sabatier, Nicolas Wernert, Sébastien Pinte, et al. "Defective DNA single-strand break repair is responsible for senescence and neoplastic escape of epithelial cells." In: *Nature communications* 7.1 (2016), p. 10399.

[12] Corinne Abbadie, Olivier Pluquet, and Albin Pourtier. "Epithelial cell senescence: an adaptive response to pre-carcinogenic stresses?" In: *Cellular and Molecular Life Sciences* 74.24 (2017), pp. 4471–4509.

[13] Erwan Goy, Maxime Tomezak, Caterina Facchin, Nathalie Martin, Emmanuel Bouchaert, Jerome Benoit, Clementine de Schutter, Joe Nassour, Laure Saas, Claire Drullion, et al. "The out-of-field dose in radiation therapy induces delayed tumorigenesis by senescence evasion." In: *Elife* 11 (2022), e67190.

[14] Vassilis Gorgoulis, Peter D Adams, Andrea Alimonti, Dorothy C Bennett, Oliver Bischof, Cleo Bishop, Judith Campisi, Manuel Collado, Konstantinos Evangelou, Gerardo Ferbeyre, et al. "Cellular senescence: defining a path forward." In: *Cell* 179.4 (2019), pp. 813–827.

[15] Francesca Veronesi, Deyanira Contartese, Laura Di Sarno, Veronica Borsari, Milena Fini, and Gianluca Giavaresi. "In vitro models of cell senescence: A systematic review on musculoskeletal tissues and cells." In: *International Journal of Molecular Sciences* 24.21 (2023), p. 15617.

[16] Kazuhito Naka, Akira Tachibana, Kyoji Ikeda, and Noboru Motoyama. "Stress-induced premature senescence in hTERT-expressing ataxia telangiectasia fibroblasts." In: *Journal of Biological Chemistry* 279.3 (2004), pp. 2030–2037.

[17] Simona Parrinello, Enrique Samper, Ana Krtolica, Joshua Goldstein, Simon Melov, and Judith Campisi. "Oxygen sensitivity severely limits the replicative lifespan of murine fibroblasts." In: *Nature cell biology* 5.8 (2003), pp. 741–747.

[18] Shuai Xiao, Dongmin Qin, Xueyang Hou, Lingli Tian, Yeping Yu, Rui Zhang, Hao Lyu, Dong Guo, Xing-Zhen Chen, Cefan Zhou, et al. "Cellular senescence: a double-edged sword in cancer therapy." In: *Frontiers in oncology* 13 (2023), p. 1189015.

[19] G H Stein, L F Drullinger, A Soulard, and V Dulic. "Differential roles for cyclin-dependent kinase inhibitors p21 and p16 in the mechanisms of senescence and differentiation in human fibroblasts." In: *Molecular Cell Biology* 19 (1999), pp. 2109–2117.

[20] Marta Varela-Eirín and Marco Demaria. "Cellular senescence." In: *Current Biology* 32.10 (2022), R448–R452.

[21] Mikhail V Blagosklonny. "Cellular senescence: when growth stimulation meets cell cycle arrest." In: *Aging (Albany NY)* 15.4 (2023), p. 905.

[22] Mikhail V Blagosklonny. "Cell senescence, rapamycin and hyperfunction theory of aging." In: *Cell Cycle* 21.14 (2022), pp. 1456–1467.

[23] Sohee Cho and Eun Seong Hwang. "Fluorescence-based detection and quantification of features of cellular senescence." In: *Methods in cell biology* 103 (2011), pp. 149–188.

[24] Hong Zhao and Zbigniew Darzynkiewicz. "Biomarkers of cell senescence assessed by imaging cytometry." In: *Cell Senescence: Methods and Protocols* (2013), pp. 83–92.

[25] Niki Chondrogianni, Fiona LL Stratford, Ioannis P Trougakos, Bertrand Friguet, A Jennifer Rivett, and Efstathios S Gonos. "Central role of the proteasome in senescence and survival of human fibroblasts: induction of a senescence-like phenotype upon its inhibition and resistance to stress upon its activation." In: *Journal of Biological Chemistry* 278.30 (2003), pp. 28026–28037.

[26] David Bernard, Karo Gosselin, Didier Monte, Chantal Vercamer, Fatima Bouali, Albin Pourtier, Bernard Vandenbunder, and Corinne Abbadie. "Involvement of Rel/Nuclear Factor-κB Transcription Factors in Keratinocyte Senescence." In: *Cancer Research* 64.2 (2004), pp. 472–481.

[27] Karo Gosselin et al. "Senescent keratinocytes die by autophagic programmed cell death." In: *American Journal of Pathology* 174 (2009), pp. 423–435.

[28] Wendy O'Brien, Göran Stenman, and Ruth Sager. "Suppression of tumor growth by senescence in virally transformed human fibroblasts." In: *Proceedings of the National Academy of Sciences* 83.22 (1986), pp. 8659–8663.

[29] Clemens A Schmitt, Boshi Wang, and Marco Demaria. "Senescence and cancer—role and therapeutic opportunities." In: *Nature reviews Clinical oncology* 19.10 (2022), pp. 619–636.

[30] Zhiyong Han, Wenyi Wei, Stephen Dunaway, James W Darnowski, Paul Calabresi, John Sedivy, Eric A Hendrickson, Kannan V Balan, Panayotis Pantazis, and James H Wyche. "Role of p21 in apoptosis and senescence of human colon cancer cells treated with camptothecin." In: *Journal of Biological Chemistry* 277.19 (2002), pp. 17154–17160.

[31] Nicolás Herranz, Suchira Gallage, Massimiliano Mellone, Torsten Wuestefeld, Sabrina Klotz, Christopher J Hanley, Selina Raguz, Juan Carlos Acosta, Andrew J Innes, Ana Banito, et al. "mTOR regulates MAPKAPK2 translation to control the senescence-associated secretory phenotype." In: *Nature cell biology* 17.9 (2015), pp. 1205–1217.

[32] Maria Peiris-Pagès, Federica Sotgia, and Michael P Lisanti. "Chemotherapy induces the cancer-associated fibroblast phenotype, activating paracrine Hedgehog-GLI signalling in breast cancer cells." In: *Oncotarget* 6.13 (2015), p. 10728.

[33] T. Rich, R.L. Allen, and A.H. Wyllie. "Defying death after DNA damage." In: *Nature* 407.6805 (2000), pp. 777–783.

[34] Natalia Mitin, Kirsten A Nyrop, Susan L Strum, Anne Knecht, Lisa A Carey, Katherine E Reeder-Hayes, E Claire Dees, Trevor A Jolly, Gretchen G Kimmick, Meghan S Karuturi, et al. "A biomarker of aging, p16, predicts peripheral neuropathy in women receiving adjuvant taxanes for breast cancer." In: *NPJ breast cancer* 8.1 (2022), p. 103.

[35] Xiaohong Peng, Yi Wu, Uilke Brouwer, Thijmen van Vliet, Boshi Wang, Marco Demaria, Lara Barazzuol, and Rob P Coppes. "Cellular senescence contributes to radiation-induced hyposalivation by affecting the stem/progenitor cell niche." In: *Cell death & disease* 11.10 (2020), p. 854.

[36] Andreas Domen, Christophe Deben, Jasper Verswyvel, Tal Flieswasser, Hans Prenen, Marc Peeters, Filip Lardon, and An Wouters. "Cellular senescence in cancer: clinical detection and prognostic implications." In: *Journal of Experimental & Clinical Cancer Research* 41.1 (2022), p. 360.

[37] Emilie Rovillain, Louise Mansfield, Christopher J Lord, Alan Ashworth, and Parmjit S Jat. "An RNA interference screen for identifying downstream effectors of the p53 and pRB tumour suppressor pathways involved in senescence." In: *BMC genomics* 12 (2011), pp. 1–12.

[38] Tarek Abbas and Anindya Dutta. "p21 in cancer: intricate networks and multiple activities." In: *Nature Reviews Cancer* 9.6 (2009), pp. 400–414.

[39] Graziella Pellegrini, Elena Dellambra, Patrizia Paterna, Osvaldo Golisano, Carlo E Traverso, Paolo Rama, Pedro Lacal, and Michele De Luca. "Telomerase activity is sufficient to bypass replicative senescence in human limbal and conjunctival but not corneal keratinocytes." In: *European journal of cell biology* 83.11-12 (2004), pp. 691–700.

[40] Parvathy A. Sarma, Corinne Abbadie, Yvan de Launoit, and Fabrizio Cleri. "Cell senescence and the DNA single-strand break damage repair pathway." In: *DNA* 4 (2024), pp. 530–552.

[41] Gregor Mendel. "Versuche uber pflanzen-hybriden." In: *Vorgelegt in den Sitzungen* (1865).

[42] Johann Friedrich Miescher. *Ueber die chemische Zusammensetzung der Eiterzellen*. 1871.

[43] OT Avery, CM Macleod, and M McCarty. "Studies on the chemical nature of the substance inducing transformation of pneumococcal types: induction of transformation by a deoxyribonucleic acid fraction isolated from Pneumococcus type III." In: *J. Exp. Med* 79 (1944), pp. 137–158.

[44] AD Herschey and M Chase. "Independent functions of viral protein and nucleic acid in growth of bacteriophage J." In: *Gen. Physiol* 36.1 (1952), pp. 39–56.

[45] Erwin Chargaff. "Chemical specificity of nucleic acids and mechanism of their enzymatic degradation." In: *Experientia* 6.6 (1950), pp. 201–209.

[46] Linus Pauling and Robert B Corey. "A proposed structure for the nucleic acids." In: *Proceedings of the National Academy of Sciences* 39.2 (1953), pp. 84–97.

[47] Leslie Pray. "Discovery of DNA structure and function: Watson and Crick." In: *Nature Education* 1.1 (2008), p. 100.

[48] PA Levene. "THE STRUCTURE OF YEAST NUCLEIC ACID. III. AMMONIA HYDROLYSIS." In: *Studies from the Rockefeller Institute for Medical Research* 30 (1919), p. 231.

[49] Ralf Dahm. "Discovering DNA: Friedrich Miescher and the early years of nucleic acid research." In: *Human genetics* 122 (2008), pp. 565–581.

[50] James D Watson and Francis HC Crick. "Molecular structure of nucleic acids: a structure for deoxyribose nucleic acid." In: *Nature* 171.4356 (1953), pp. 737–738.

[51] B Alberts, A Johnson, J Lewis, et al. *Molecular biology of the cell. New York: Garland Science; 2002. The structure and function of DNA*.

[52] Kensal E. Van Holde. *Chromatin*. Springer-Verlag, 1988.

[53] Alan P. Wolffe. *Chromatin: Structure and Function*. Academic Press, 1999.

[54] Ada L. Olins and Donald E. Olins. "Spheroid chromatin units (v bodies)." In: *Science* 183.4122 (1974), pp. 330–332.

[55] Christopher L. Woodcock, Judith P. Safer, and John E. Stanchfield. "Structural repeating units in chromatin: I. Evidence for their general occurrence." In: *The Journal of Cell Biology* 71.3 (1976), pp. 424–435.

[56] C.A. Davey, D.F. Sargent, K. Luger, A.W. Maeder, and T.J. Richmond. "Solvent mediated interactions in the structure of the nucleosome core particle at 1.9 A resolution." In: *J. Mol. Biol.* 319 (2002), pp. 1097–1113.

[57] Ada L. Olins and Donald E. Olins. "Chromatin history: our view from the bridge." In: *Nature Reviews Molecular Cell Biology* 4.10 (2003), pp. 809–814.

[58] David R. Hewish and Lesley A. Burgoyne. "Chromatin substructure. The digestion of chromatin DNA at regularly spaced sites by a nuclear deoxyribonuclease." In: *Biochemical and Biophysical Research Communications* 52.2 (1973), pp. 504–510.

[59] Martin Noll. "Subunit structure of chromatin." In: *Nature* 251.5471 (1974), pp. 249–251.

[60] Roger D. Kornberg. "Chromatin structure: a repeating unit of histones and DNA." In: *Science* 184.4139 (1974), pp. 868–871.

[61] Kensal E. Van Holde, C. G. Sahasrabuddhe, and B. R. Shaw. "A model for particles containing histones." In: *Nucleic Acids Research* 1.11 (1974), pp. 1579–1586.

[62] Pierre Oudet, Michel Gross-Bellard, and Pierre Chambon. "Electron microscopic and biochemical evidence that chromatin structure is a repeating unit." In: *Cell* 4.4 (1975), pp. 281–300.

[63] Karolin Luger, Armin W. Mäder, Robin K. Richmond, Dave F. Sargent, and Timothy J. Richmond. "Crystal structure of the nucleosome core particle at 2.8 Å resolution." In: *Nature* 389.6648 (1997), pp. 251–260.

[64] Christopher L. Woodcock. "Higher-order structure and dynamics of chromatin." In: *Molecular Cell* 15.4 (2005), pp. 525–535.

[65] Jan Bednar, Robert A. Horowitz, Sergey A. Grigoryev, L. Mark Carruthers, Jeffrey C. Hansen, Arie J. Koster, and Christopher L. Woodcock. "Nucleosomes, linker DNA, and linker histone form a unique structural motif that directs the higher-order folding and compaction of chromatin." In: *Proceedings of the National Academy of Sciences* 95.24 (1998), pp. 14173–14178.

[66] Kirti Prakash. "The periodic and dynamic structure of chromatin." PhD thesis. 2016.

[67] J Allan, PG Hartman, C Crane-Robinson, and FX Aviles. "The structure of histone H1 and its location in chromatin." In: *Nature* 288.5792 (1980), pp. 675–679.

[68] Wolfgang Fischle, Yuguang Wang, and C. David Allis. "Enzymatic activity associated with class I histone deacetylases is required for heterochromatin silencing and genome stability." In: *Nature* 437.7062 (2005), pp. 432–435.

[69] Charles L. Smith and Craig L. Peterson. "Chromatin remodeling: nucleosome dynamics during epigenetic regulation and noncoding DNA transactions." In: *Nature Reviews Molecular Cell Biology* 6.12 (2005), pp. 930–942.

[70] V Barra and D Fachinetti. "The dark side of centromeres: types, causes and consequences of structural abnormalities implicating centromeric DNA." In: *Nature communications* 9.1 (2018), p. 4340.

[71] Malathi G Nayak, Anice George, MS Vidyasagar, Stanley Mathew, Sudhakar Nayak, Baby S Nayak, YN Shashidhara, and Asha Kamath. "Quality of life among cancer patients." In: *Indian journal of palliative care* 23.4 (2017), p. 445.

[72] Elham Bidram, Yasaman Esmaeili, Hadi Ranji-Burachaloo, Nuha Al-Zaubai, Ali Zarrabi, Alastair Stewart, and Dave E Dunstan. "A concise review on cancer treatment methods and delivery systems." In: *Journal of Drug Delivery Science and Technology* 54 (2019), p. 101350.

[73] Chinna Babu Dracham, Abhash Shankar, and Renu Madan. "Radiation induced secondary malignancies: a review article." In: *Radiation oncology journal* 36.2 (2018), p. 85.

[74] Elizabeth Cohen-Jonathan, Eric J Bernhard, and W Gillies McKenna. "How does radiation kill cells?" In: *Current opinion in chemical biology* 3.1 (1999), pp. 77–83.

[75] Emil Mladenov and George Iliakis. "Induction and repair of DNA double strand breaks: the increasing spectrum of non-homologous end joining pathways." In: *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* 711.1-2 (2011), pp. 61–72.

[76] Agnes Schipler and George Iliakis. "DNA double-strand–break complexity levels and their possible contributions to the probability for error-prone processing and repair pathway choice." In: *Nucleic acids research* 41.16 (2013), pp. 7589–7605.

[77] Clemens von Sonntag. "Recent Trends in Radiation Chemistry." In: *Wishart, JF* (2010), p. 543.

[78] John F Ward. "DNA damage produced by ionizing radiation in mammalian cells: identities, mechanisms of formation, and reparability." In: *Progress in nucleic acid research and molecular biology* 35 (1988), pp. 95–125.

[79] Hans E Krokan, Finn Drabløs, and Geir Slupphaug. "Uracil in DNA - occurrence, consequences and repair." In: *Oncogene* 21 (2002), pp. 8935–8948.

[80] Clemens Sonntag. *Free-radical-induced DNA damage and its repair: a chemical perspective*. Springer Science & Business Media, 2006, pp. 379–390.

[81] Marisa L. Taverna Porro and Marc M. Greenberg. "Double-Strand Breaks from a Radical Commonly Produced by DNA-Damaging Agents." In: *Chemical Research in Toxycology* 28 (2015), pp. 810–816.

[82] M. Weinfeld and K. J. Soderling. "32P-postlabeling detection of radiation-induced DNA damage: identification and estimation of thymine glycols and phosphoglycolate termini." In: *Biochemistry* 30.5 (1999), pp. 1091–1097.

[83] T. Zhou, K. Akopiants, S. Mohapatra, P. S. Lin, K. Valerie, D. A. Ramsden, S. P. Lees-Miller, and L. F. Povirk. "Tyrosyl-DNA phosphodiesterase and the repair of 3'-phosphoglycolate-terminated DNA double-strand breaks." In: *DNA Repair* 8.8 (2009), pp. 901–911.

[84] Evelyne Sage and Naoya Shikazono. "Radiation-induced clustered DNA lesions: Repair and mutagenesis." In: *Free Radical Biology and Medicine* 107 (2017), pp. 125–135.

[85] AJ Chalmers and RD Carruthers. "Radiobiology summaries: DNA damage and repair." In: *Clinical Oncology* 33.5 (2021), pp. 275–278.

[86] Rui-Xue Huang and Ping-Kun Zhou. "DNA damage response signaling pathways and targets for radiotherapy sensitization in cancer." In: *Signal transduction and targeted therapy* 5.1 (2020), p. 60.

[87] Alexandre Maréchal and Lee Zou. "DNA damage sensing by the ATM and ATR kinases." In: *Cold Spring Harbor perspectives in biology* 5.9 (2013), a012716.

[88] Tomas Lindahl. "Instability and decay of the primary structure of DNA." In: *nature* 362.6422 (1993), pp. 709–715.

[89] O Tounekti, A Kenani, Net al Foray, S Orlowski, and Ll M Mir. "The ratio of single-to double-strand DNA breaks and their absolute values determine cell death pathway." In: *British journal of cancer* 84.9 (2001), pp. 1272–1279.

[90] Andrei Kuzminov. "Single-strand interruptions in replicating chromosomes cause double-strand breaks." In: *Proceedings of the National Academy of Sciences* 98.15 (2001), pp. 8241–8246.

[91] Keith W Caldecott. "Mammalian DNA single-strand break repair: an X-ra (y) ted affair." In: *Bioessays* 23.5 (2001), pp. 447–455.

[92] Aroumougame Asaithamby, Burong Hu, and David J. Chen. "Unrepaired clustered DNA lesions induce chromosome breakage in human cells." In: *Proceedings of the National Academy of Sciences* 108.20 (2011), pp. 8293–8298.

[93] Thomas L. Clarke and Raul Mostoslavsky. "DNA repair as a shared hallmark in cancer and ageing." In: *Molecular Oncology* 16.18 (2022), pp. 3352–3379.

[94] Ben Clasie, Andrew Wroe, Hanne Kooy, Nicolas Depauw, Jay Flanz, Harald Paganetti, and Anatoly Rosenfeld. "Prolonged persistence of mutagenic DNA lesions in somatic cells." In: *Medical physics* 37.1 (2010), pp. 311–321.

[95] Wenjian Ma, Christopher J Halweg, Daniel Menendez, and Michael A Resnick. "Differential effects of poly (ADP-ribose) polymerase inhibition on DNA break repair in human cells are revealed with Epstein–Barr virus." In: *Proceedings of the National Academy of Sciences* 109.17 (2012), pp. 6590–6595.

[96] Michael M Vilenchik and Alfred G Knudson. "Endogenous DNA double-strand breaks: production, fidelity of repair, and induction of cancer." In: *Proceedings of the National Academy of Sciences* 100.22 (2003), pp. 12871–12876.

[97] Mojtaba Mokari, Mohammad Hassan Alamatsaz, Hossein Moeini, and Reza Taleei. "A simulation approach for determining the spectrum of DNA damage induced by protons." In: *Physics in Medicine & Biology* 63.17 (2018), p. 175003.

[98] E. A. Kouzminova and A. Kuzminov. "Fragmentation of replicating chromosomes triggered by uracil in DNA." In: *Journal of Molecular Biology* 355 (2006), pp. 20–33.

[99] Keith W Caldecott. "DNA single-strand break repair and human genetic disease." In: *Trends in Cell Biology* 32.9 (2022), pp. 733–745.

[100] Keith W Caldecott. "Causes and consequences of DNA single-strand breaks." In: *Trends in Biochemical Sciences* 49.1 (2024), pp. 68–78.

[101] Svetlana V Khoronenkova and Grigory L Dianov. "ATM prevents DSB formation by coordinating SSB repair and cell cycle progression." In: *Proceedings of the National Academy of Sciences* 112.13 (2015), pp. 3997–4002.

[102] Anna R Poetsch. "The genomics of oxidative DNA damage, repair, and resulting mutagenesis." In: *Computational and structural biotechnology journal* 18 (2020), pp. 207–219.

[103] AB Robertson, A Klungland, T Rognes, and I Leiros. "DNA repair in mammalian cells: Base excision repair: the long and short of it." In: *Cellular and molecular life sciences* 66 (2009), pp. 981–993.

[104] David Svilar, Eva M Goellner, Karen H Almeida, and Robert W Sobol. "Base excision repair and lesion-dependent subpathways for repair of oxidative DNA damage." In: *Antioxidants & redox signaling* 14.12 (2011), pp. 2491–2507.

[105] Eva Petermann, Mathias Ziegler, and Shiao Li Oei. "ATP-dependent selection between single nucleotide and long patch base excision repair." In: *DNA repair* 2.10 (2003), pp. 1101–1114.

[106] Arne Klungland and Tomas Lindahl. "Second pathway for completion of human DNA base excision-repair: reconstitution with purified proteins and requirement for DNase IV (FEN1)." In: *The EMBO journal* (1997).

[107] Paola Fortini and Eugenia Dogliotti. "Base damage and single-strand break repair: mechanisms and functional significance of short-and long-patch repair subpathways." In: *DNA repair* 6.4 (2007), pp. 398–409.

[108] Rachel Abbotts and David M Wilson III. "Coordination of DNA single strand break repair." In: *Free Radical Biology and Medicine* 107 (2017), pp. 228–244.

[109] Eric J. Hall and Amato J. Giaccia. *Radiobiology for the Radiologist*. 6th. Lippincott Williams & Wilkins, 2006. ISBN: 978-0-7817-4151-4.

[110] Igor Koturbash, Jonathan Loree, Kristy Kutanzi, Clayton Koganow, Igor Pogribny, and Olga Kovalchuk. "In vivo bystander effect: cranial X-irradiation leads to elevated DNA damage, altered cellular proliferation and apoptosis, and increased p53 levels in shielded spleen." In: *International Journal of Radiation Oncology* Biology* Physics* 70.2 (2008), pp. 554–562.

[111] M Najafi, R Fardid, GH Hadadi, and M Fardid. "The mechanisms of radiation-induced bystander effect." In: *Journal of biomedical physics & engineering* 4.4 (2014), p. 163.

[112] Tom K. Hei, Hongning Zhou, and Vladimir N. Ivanov. "Mechanism of radiation-induced bystander effects: A unifying model." In: *Journal of Pharmacology and Experimental Therapeutics* 328 (2011), pp. 10–23.

[113] Amy Berrington de Gonzalez, Ethel Gilbert, Rochelle Curtis, Peter Inskip, Ruth Kleinerman, Lindsay Morton, Preetha Rajaraman, and Mark P Little. "Second solid cancers after radiation therapy: a systematic review of the epidemiologic studies of the radiation dose-response relationship." In: *International Journal of Radiation Oncology* Biology* Physics* 86.2 (2013), pp. 224–233.

[114] Nadine Chofor, Dietrich Harder, K Willborn, and B Poppe. "Innovative approaches to reduce radiation-induced secondary cancer risk." In: *Strahlentherapie und Onkologie* 188.9 (2012), pp. 672–677.

[115] Nadine Chofor, Dietrich Harder, B Poppe, Ernst-Peter Rührnschopf, and K Willborn. "The impact of the scatter dose on the risk of radiation-induced second cancers in radiotherapy of the prostate." In: *Radiation Oncology* 6.1 (2011), pp. 1–9.

[116] Keith W Caldecott. "Single-strand break repair and genetic disease." In: *Nature Reviews Genetics* 9.8 (2008), pp. 619–631.

[117] Keith W. Caldecott. "Protein ADP-ribosylation and the cellular response to DNA strand breaks." In: *DNA Repair* 19 (2014), pp. 108–113.

[118] Cecilia E. Str'om, Oliver Mortusewicz, David Finch, Jason L. Parsons, Anne Lagerqvist, Fredrik Johansson, Niklas Schultz, Klaus Erixon, Grigory L. Dianov, and Thomas Helleday. "CK2 phosphorylation of XRCC1 facilitates dissociation from DNA and single-strand break formation during base excision repair phosphorylation of XRCC1 facilitates dissociation from DNA and single-strand break formation during base excision repair." In: *DNA Repair* 10.9 (2011), pp. 961–969.

[119] Mohammed A. Al-Mohanna, Hassan H. Al-Khalaf, and Nada Al-Yousef. "Apoptosis and Bax expression are inhibited by p16INK4a in breast cancer cells." In: *Cellular Signalling* 16 (2004), pp. 681–688.

[120] Parvathy AP Sarma, Corinne Abbadie, and Fabrizio Cleri. "Cooperative dynamics of PARP-1 zinc-finger domains in the detection of DNA single-strand breaks." In: *Scientific Reports* 14.1 (2024), p. 23257.

[121] Safwen Ghediri, Parvathy A. Sarma, Corinne Abbadie, Ralf Blossey, and Fabrizio Cleri. "How repair proteins identify DNA damage in the nucleosome." In: *chemRkiv preprint* (2025).

[122] W. Lilyestrom, M. J. van der Woerd, N. Clark, and K. Luger. "Structural and biophysical studies of human PARP-1 in complex with damaged DNA." In: *J. Mol. Biol.* 395 (2010), pp. 983–994.

[123] A.A.E. Ali, G. Timinszky, R. Arribas-Bosacoma, M. Kozlowski, P.O. Hassa, M. Hassler, A.G. Ladurner, L.H. Pearl, and A.W. Oliver. "The zinc-finger domains of PARP1 cooperate to recognize DNA strand breaks." In: *Nature Struc. Mol. Biol.* 19 (2012), pp. 685–692.

[124] Jesper V. Olsen et al. "Quantitative phosphoproteomics reveals widespread full phosphorylation site occupancy during mitosis." In: *Science Signal.* 3 (2010), ra3.

[125] J. Rudolph, U.M. Muthurajan, M. Palacio, J. Mahadevan, G. Roberts, A.H. Erbse, P.N. Dyer, and K. Luger. "The BRCT domain of PARP1 binds intact DNA and mediates intrastrand transfer." In: *Mol. Cell* 81 (2021), pp. 4994–5006.

[126] S. Bilokapic, M. J. Suskiewicz, I. Ahel, and M. Halic. "Bridging of DNA breaks activates PARP2-HPF1 to modify chromatin." In: *Nature* 503 (2020), pp. 609–613.

[127] G. Gaullier, G. Roberts, U.M. Muthurajan, S. Bowerman, J. Rudolph, J. Mahadevan, A. Jha, P.S. Rae, and K. Luger. "Bridging of nucleosome-proximal DNA double-strand breaks by PARP2 enhances its interaction with HPF1." In: *PlosONE* 15 (2020), e0240932.

[128] S. Eustermann, W. F. Wu, M. F. Langelier, J. C. Yang, L. E. Easton, A. A. Riccio, J. M. Pascal, and D. Neuhaus. "Structural basis of detection and signaling of DNA single-strand breaks by human PARP-1." In: *Mol. Cell* 60 (2015), pp. 742–754.

[129] J. M. Dawicki-McKenna, M. F. Langelier, J. E. DeNizio, A. A. Riccio, C. D. Cao, K. R. Karch, M. McCauley, J. D. Steffen, B. E. Black, and J. M. Pascal. "PARP-1 activation requires local unfolding of an autoinhibitory domain." In: *Mol. Cell* 60 (2015), pp. 755–768.

[130] Marie-France Langelier, Jamie L. Planck, Swati Roy, and John M. Pascal. "Crystal Structures of Poly(ADP-ribose) Polymerase-1 (PARP-1) Zinc Fingers Bound to DNA." In: *J. Biol. Chem.* 286 (2011), pp. 10690–10701.

[131] Eric Le Cam, Fred Fack, Josiane Ménissier de Murcia, Jean A.H. Cognet, Agnés Barbin, Vassilis Sarantoglou, Bernard Révet, Etienne Delain, and Gilbert de Murcia. "Conformational Analysis of a 139 Base-pair DNA Fragment Containing a Single-stranded Break and Its Interaction with Human Poly(ADP-ribose) Polymerase." In: *J. Mol. Biol.* 235 (1994), pp. 1062–1071.

[132] A. Sefer, E. Kallis, T. Eilert, C. Röcker, O. Kolesnikova, D. Neuhaus, S. Eustermann, and J. Michaelis. "Structural dynamics of DNA strand break sensing by PARP-1 at a single-molecule level." In: *Nature Comm.* 13 (2022), p. 6569.

[133] Cody Geary, Guido Grossi, Ewan K. S. McRae, Paul W. K. Rothemund, and Ebbe S. Andersen. "RNA-origami design tools enable cotranscriptional folding of kilobase-sized nanoscaffolds." In: *Nature Chem.* 13 (2021), pp. 549–558.

[134] Carlos Bustamante, Zev Bryant, and Steven B. Smith. "Ten years of tension: single-molecule DNA mechanics." In: *Nature* 421 (2010), pp. 423–427.

[135] Jonathan S. Mitchell, Jaroslaw Glowacki, Alexandre E. Grandchamp, Robert S. Manning, and John H. Maddocks. "Sequence-dependent persistence lengths of DNA." In: *J. Chem. Theory Comput.* 13 (2017), pp. 1539–1555.

[136] Zong Sean Juo, Thang Kien Chiu, Paul M. Leiberman, Igor Baikalov, Arnold J. Berk, and Richard E. Dickerson. "How Proteins Recognize the TATA Box." In: *J. Mol. Biol.* 261 (1996), pp. 239–254.

[137] L. Pardo, M. Campillo, D. Bosch, N. Pastor, and H. Weinstein. "Binding mechanisms of TATA box-binding proteins: DNA kinking is stabilized by specific hydrogen bonds." In: *Biophys. J.* 78 (2000), pp. 1988–1996.

[138] M. Mondal, S. Mukherjee, and D. Bhattacharyya. "Contribution of phenylalanine side-chain intercalation to the TATA-box binding protein–DNA interaction: molecular dynamics and dispersion-corrected density functional theory studies." In: *J. Mol. Mod.* 20 (2014), p. 2499.

[139] Elizaveta E. Alemasova and Olga I. Lavrik. "Poly(ADP-ribosyl)–ation by PARP1: reaction mechanism and regulatory proteins." In: *Nucl. Acids Res.* 47 (2019), pp. 3811–3827.

[140] Katie A. Wilson, Jennifer L. Kellie, and Stacey D. Wetmore. "DNA–protein π-interactions in nature: abundance, structure, composition and strength of contacts between aromatic amino acids and DNA nucleobases or deoxyribose sugar." In: *Nucl. Acids Res.* 42 (2014), pp. 6726–6741.

[141] Fabrizio Cleri, Stefano Giordano, and Ralf Blossey. "Nucleosome Array Deformation in Chromatin is Sustained by Bending, Twisting and Kinking of Linker DNA." In: *J. Mol. Biol.* 435 (2023).

[142] M. K. Sarangi, V. Zvoda, M. N. Holte, N. A. Becker, J. P. Peters, L. J. Maher, and A. Ansari. "Evidence for a bind-then-bend mechanism for architectural DNA binding protein yNhp6A." In: *Nucl. Acids Res.* 47 (2019), pp. 2781–2791.

[143] Abhinandan Panigrahi, Hemanth Vemuri, Madhur Aggarwal, Kartheek Pitta, and Marimuthu Krishnan. "Sequence specificity, energetics and mechanism of mismatch recognition by DNA damage sensing protein Rad4/XPC." In: *Nucl. Acids Res.* 48 (2020), pp. 2246–2257.

[144] Samuel B Yoshua, George D Watson, Jamieson A L Howard, Victor Velasco-Berrelleza, Mark C Leake, and Agnes Noy. "Integration host factor bends and bridges DNA in a multiplicity of binding modes with varying specificity." In: *Nucl. Acids Res.* 49 (2021), pp. 8684–8698.

[145] Maria V Sukhanova, Sanae Abrakhi, Vandana Joshi, David Pastre, Mikhail M Kutuzov, Rashid O Anarbaev, Patrick A Curmi, Loic Hamon, and Olga I Lavrik. "Single molecule detection of PARP1 and PARP2 interaction with DNA strand breaks and their poly (ADP-ribosyl) ation using high-resolution AFM imaging." In: *Nucleic acids research* 44.6 (2016), e60–e60.

[146] Natalya V Maluchenko, Dmitry K Nilov, Sergey V Pushkarev, Elena Y Kotova, Nadezhda S Gerasimova, Mikhail P Kirpichnikov, Marie-France Langelier, John M Pascal, Md Sohail Akhtar, Alexey V Feofanov, et al. In: *International Journal of Molecular Sciences* 22.22 (2021), p. 12127.

[147] Daniel Sultanov, Nadezhda Gerasimova, Kseniya Kudryashova, Natalya Maluchenko, Elena Kotova, Marie-France Langelier, John Pascal, Mikhail Kirpichnikov, Alexey Feofanov, and Vasily Studitsky. "Unfolding of core nucleosomes by PARP-1 revealed by spFRET microscopy." In: *AIMS genetics* 4.01 (2017), pp. 021–031.

[148] Nicholas J Clark, Michael Kramer, Uma M Muthurajan, and Karolin Luger. "Alternative modes of binding of poly (ADP-ribose) polymerase 1 to free DNA and nucleosomes." In: *Journal of Biological Chemistry* 287.39 (2012), pp. 32430–32439.

[149] Deepti Sharma, Louis De Falco, Sivaraman Padavattan, Chang Rao, Susana Geifman-Shochat, Chuan-Fa Liu, and Curt A Davey. "PARP1 exhibits enhanced association and catalytic efficiency with γH2A. X-nucleosome." In: *Nature Communications* 10.1 (2019), p. 5751.

[150] Natacha Gillet, Elise Dumont, and Emmanuelle Bignon. "DNA damage and repair in the nucleosome: insights from computational methods." In: *Biophysical Reviews* (2024), pp. 1–12.

[151] Fabrizio Cleri, Fabio Landuzzi, and Ralf Blossey. "Mechanical evolution of DNA double-strand breaks in the nucleosome." In: *PLOS Comp. Biol.* 14 (2018), e1006224.

[152] G. G. Poirier, G. de Murcia, J. Jongstra-Bilen, C. Niedergang, and P. Mandel. "Poly(ADP-ribosyl)ation of polynucleosomes causes relaxation of chromatin structure." In: *Proc. National Ac. Science USA (PNAS)* 79 (1982), pp. 3423–3427.

[153] Raga Krishnakumar, Matthew J Gamble, Kristine M Frizzell, Jhoanna G Berrocal, Miltiadis Kininis, and W Lee Kraus. "Reciprocal binding of PARP-1 and histone H1 at promoters specifies transcriptional outcomes." In: *Science* 319.5864 (2008), pp. 819–821.

[154] Uma M Muthurajan, Maggie RD Hepler, Aaron R Hieb, Nicholas J Clark, Michael Kramer, Tingting Yao, and Karolin Luger. "Automodification switches PARP-1 function from chromatin architectural protein to histone chaperone." In: *Proceedings of the National Academy of Sciences* 111.35 (2014), pp. 12752–12757.

[155] J. I. Friedman and J. T. Stivers. "Detection of damaged DNA bases by DNA glycosylase enzymes." In: *Biochemistry* 49 (2010), pp. 4957–4967.

[156] Isabelle Bonnet, Andreas Biebricher, Pierre-Louis Porté, Claude Loverdo, Olivier Bénichou, Raphaël Voituriez, Christophe Escudé, Wolfgang Wende, Alfred Pingoud, and Pierre Desbiolles. "Sliding and jumping of single EcoRV restriction enzymes on non-cognate DNA." In: *Nucl. Acids Res.* 36 (2008).

[157] Carina Monico, Marco Capitanio, Gionata Belcastro, Francesco Vanzi, and Francesco S. Pavone. "Optical Methods to Study Protein-DNA Interactions in Vitro and in Living Cells at the Single-Molecule Level." In: *Int. J. Molec. Sci.* 14 (2013).

[158] Kiyoto Kamagata, Yuji Itoh, and Dwiky Rendra Graha Subekti. "How p53 Molecules Solve the Target DNA Search Problem: A Review." In: *Int. J. Molec. Sci.* 21 (2020).

[159] P.H. von Hippel and O.G. Berg. "Facilitated target location in biological systems." In: *Journal of Biological Chemistry* 264 (1989), pp. 675–678.

[160] Stephen E. Halford and John F. Marko. "How do site-specific DNA-binding proteins find their targets?" In: *Nucleic Acids Research* 32 (2004), pp. 3040–3052.

[161] O.G. Berg, R.B. Winter, and P.H. von Hippel. "Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory." In: *Biochemistry* 20 (1981), pp. 6929–6948.

[162] Paul C. Blainey, Antoine M. van Oijen, Anirban Banerjee, Gregory L. Verdine, and X. Sunney Xie. "A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA." In: *Proc. National Ac. Science USA (PNAS)* 103 (2006), pp. 5752–5757.

[163] Neil M. Kad and Bennett Van Houten. "Dynamics of lesion processing by bacterial nucleotide excision repair proteins." In: *Prog. Molecular Biology Transl. Science* 110 (2012), pp. 1–24.

[164] R. H. Porecha and J. T. Stivers. "Uracil-DNA glycosylase uses DNA hopping and short-range sliding to trap extrahelical uracils." In: *Proc. National Ac. Science USA (PNAS)* 105 (2008), pp. 10791–10796.

[165] Muwen Kong et al. "Single-Molecule Imaging Reveals that Rad4 Employs a Dynamic DNA Damage Recognition Process." In: *Molecular Cell* 64 (2016), pp. 376–387.

[166] Susan E. Tsutakawa, Chi-Lin Tsai, Chunli Yan, Amer Bralic, Walter J. Chazin, Samir M. Hamdan, Orlando D. Schärer, Ivaylo Ivanov, and John A. Tainer. "Envisioning how the prototypic molecular machine TFII-H functions in transcription initiation and DNA repair." In: *DNA Repair* 96 (2020), p. 102972.

[167] Lili Liu, Muwen Kong, Natalie R. Gassman, Bret D. Freudenthal, Rajendra Prasad, Stephanie Zhen, Simon C. Watkins, Samuel H. Wilson, and Bennett Van Houten. "PARP1 changes from three-dimensional DNA damage searching to one-dimensional diffusion after auto-PARylation or in the presence of APE1." In: *Nucleic Acids Research* 45 (2017), pp. 12834–12847.

[168] H. Kenzaki and S. Takada. "Linker DNA Length is a Key to Tri-nucleosome Folding." In: *J Mol Biol.* 433 (2021), p. 166792.

[169] Thomas Brouwer, Chi Pham, Artur Kaczmarczyk, Willem-Jan De Voogd, Margherita Botto, Petra Vizjak, Felix Mueller-Planitz, and John van Noort. "A critical role for linker DNA in higher-order folding of chromatin fibers." In: *Nucleic acids research* 49.5 (2021), pp. 2537–2551.

[170] Ália Dos Santos, Alexander W Cook, Rosemarie E Gough, Martin Schilling, Nora A Olszok, Ian Brown, Lin Wang, Jesse Aaron, Marisa L Martin-Fernandez, Florian Rehfeldt, et al. "DNA damage alters nuclear mechanics through chromatin reorganization." In: *Nucleic acids research* 49.1 (2021), pp. 340–353.

[171] Saroj Baral, Sagnik Chakraborty, Peter J Steinbach, Debamita Paul, Jung-Hyun Min, and Anjum Ansari. "Evidence for intrinsic DNA dynamics and deformability in damage sensing by the Rad4/XPC nucleotide excision repair complex." In: *Nucleic Acids Research* 53.2 (2025), gkae1290.

[172] Ariel Afek et al. "DNA mismatches reveal conformational penalties in protein-DNA recognition." In: *Nature* 587 (2020), 291–296.

[173] Fabio Landuzzi, Xavier Viader-Godoy, Fabrizio Cleri, Isabelle Pastor, and Felix Ritort. "Detection of single DNA mismatches by force spectroscopy in short DNA hairpins." In: *J. Chem. Phys.* 152 (2020), p. 074204.

[174] H. Berendsen, D. van der Spoel, and R. van Drunen. "GROMACS: A message-passing parallel molecular dynamics implementation." In: *Comp. Phys. Comm.* 91 (1995), pp. 43–56.

[175] E. Lindahl, B. Hess, and D. van der Spoel. "GROMACS 3.0: a package for molecular simulation and trajectory analysis." In: *J. Mol. Model.* 7 (2001), pp. 306–317.

[176] Kshitij Amar, Fuxiang Wei, Junwei Chen, and Ning Wang. "IEffects of forces on chromatin." In: *APL Bioengineering* 5.4 (2021), p. 041503.

[177] Miguel Angel González and José L. F. Abascal. "The shear viscosity of rigid water models." In: *Journal of Chemical Physics* 132 (2010), p. 096101.

[178] Animangsu Ghatak and Apurba Lal Das. "Kink Instability of a Highly Deformable Elastic Cylinder." In: *Physical Review Letters* 99 (2007), p. 076101.

[179] Emery Hsu, Daeyeon Lee, and Eli Sloutskin. "Non-Classical Euler Buckling and Brazier Instability in Cylindrical Liquid Droplets." In: *Nano Letters* 24.28 (2024), pp. 8717–8722.

[180] Wilma K. Olson et al. "A standard reference frame for the description of nucleic acid base-pair geometry." In: *J. Mol. Biol.* 313 (2001), pp. 229–237.

[181] Richard Lavery, MJHPD Moakher, John H Maddocks, D Petkeviciute, and Krystyna Zakrzewska. "Conformational analysis of nucleic acids revisited: Curves+." In: *Nucleic acids research* 37.17 (2009), pp. 5917–5929.

[182] Mark Abraham, Andrey Alekseenko, Cathrine Bergh, Christian Blau, Eliane Briand, Mahesh Doijade, Stefan Fleischmann, Vytautas Gapsys, Gaurav Garg, Sergey Gorelov, et al. "GROMACS 2023.1 Manual." In: *GROMACS: Groningen, The Netherlands* (2023).

[183] M. Crisp, Q. Liu, K. Roux, J.B. Rattner, C. Shanahan, B. Burke, P.D. Stahl, and D. Hodzic. "Coupling of the nucleus and cytoplasm: role of the LINC complex." In: *J. Cell Biol.* 172 (2006), pp. 41–53.

[184] Nestor V Torres and Guido Santos. "The (mathematical) modeling process in biosciences." In: *Frontiers in Genetics* 6 (2015), p. 169934.

[185] Uri Alon. *An Introduction to Systems Biology: Design Principles of Biological Circuits*. CRC Press, Boca Raton FL (USA), 2006.

[186] Brian P. Ingalls. *Mathematical Modeling in Systems Biology: An Introduction*. MIT Press, Boston (USA), 2010.

[187] International Human Genome Sequencing Consortium. "Initial sequencing and analysis of the human genome." In: *Nature* 409 (2001), pp. 860–921.

[188] Rachel M. Sherman and Steven L. Salzberg. "Pan-genomics in the human genome era." In: *Nature Rev. Genetics* 21 (2020), pp. 243–254.

[189] Nesma ElKalaawy and Amr Wassal. "Methodologies for the modeling and simulation of biochemical networks, illustrated for signal transduction pathways: A primer." In: *Biosystems* 129 (2015), pp. 1–18.

[190] Claire J Tomlin and Jeffrey D Axelrod. "Biology by numbers: mathematical modelling in developmental biology." In: *Nature reviews genetics* 8.5 (2007), pp. 331–340.

[191] Fabrizio Cleri. "Agent-based model of multicellular tumor spheroid evolution including cell metabolism." In: *Eur. J. Phys. E* 42 (2019), p. 112.

[192] Nicholas Metropolis and Stanislaw Ulam. "The Monte Carlo method." In: *Journal of the American Statistical Association* 44.247 (1949), pp. 335–341.

[193] Eugene Surdutovich and Andrey V Solov'yov. "Multiscale approach to the physics of radiation damage with ions." In: *The European Physical Journal D* 68.10 (2014), p. 353.

[194] Shirin Rahmanian, Reza Taleei, and Hooshang Nikjoo. "Radiation induced base excision repair (BER): a mechanistic mathematical approach." In: *DNA repair* 22 (2014), pp. 89–103.

[195] Bahrad A Sokhansanj, Garry R Rodrigue, J Patrick Fitch, and David M Wilson III. "A quantitative model of human DNA base excision repair. I. Mechanistic insights." In: *Nucleic acids research* 30.8 (2002), pp. 1817–1825.

[196] Bahrad A Sokhansanj and David M Wilson III. "Oxidative DNA damage background estimated by a system model of base excision repair." In: *Free Radical Biology and Medicine* 37.3 (2004), pp. 422–427.

[197] Stefan Hoops, Sven Sahle, Ralph Gauges, Christine Lee, Jürgen Pahle, Natalia Simus, Mudita Singhal, Liang Xu, Pedro Mendes, and Ursula Kummer. "COPASI—a complex pathway simulator." In: *Bioinformatics* 22.24 (2006), pp. 3067–3074.

[198] Majd Haj, Yann Frey, Amit Levon, and Yosef Shiloh. "The cGAS–STING, p38 MAPK, and p53 pathways link genome instability to accelerated cellular senescence in ATM-deficient murine lung fibroblasts." In: *Proc. National Ac. Science USA (PNAS)* 122.2 (2025), e2419196122.

[199] J Schuemann, AL McNamara, J Ramos-Méndez, J Perl, KD Held, H Paganetti, S Incerti, and B Faddegon. "TOPAS-nBio: an extension to the TOPAS simulation toolkit for cellular and sub-cellular radiobiology." In: *Radiation research* 191.2 (2019), pp. 125–138.

[200] John W Warmenhoven, Nicholas T Henthorn, Samuel P Ingram, Amy L Chadwick, Marios Sotiropoulos, Nickolay Korabel, Sergei Fedotov, Ranald I Mackay, Karen J Kirkby, and Michael J Merchant. "Insights into the non-homologous end joining pathway and double strand break end mobility provided by mechanistic in silico modelling." In: *DNA repair* 85 (2020), p. 102743.

[201] Hyojun Park, Harald Paganetti, Jan Schuemann, Xun Jia, and Chul Hee Min. "Monte Carlo methods for device simulations in radiation therapy." In: *Physics in Medicine & Biology* 66.18 (2021), 18TR01.

[202] Omar Rodrigo García García, Ramon Ortiz, Eduardo Moreno-Barbosa, Naoki D-Kondo, Bruce Faddegon, and Jose Ramos-Méndez. "TOPAS-Tissue: A Framework for the Simulation of the Biological Response to Ionizing Radiation at the Multi-Cellular Level." In: *International Journal of Molecular Sciences* 25.18 (2024), p. 10061.

[203] David Edward Charlton, Hooshang Nikjoo, and JL Humm. "Calculation of initial yields of single-and double-strand breaks in cell nuclei from electrons, protons and alpha particles." In: *International journal of radiation biology* 56.1 (1989), pp. 1–19.

[204] H Tomita, M Kai, T Kusama, and A Ito. "Monte Carlo simulation of DNA strand-break induction in supercoiled plasmid pBR322 DNA from indirect effects." In: *Radiation and environmental biophysics* 36 (1998), pp. 235–241.

[205] Werner Friedland, Peter Jacob, Philipp Bernhardt, Herwig G Paretzke, and Michael Dingfelder. "Simulation of DNA damage after proton irradiation." In: *Radiation research* 159.3 (2003), pp. 401–410.

[206] H Nikjoo, D Emfietzoglou, T Liamsuwan, R Taleei, David Liljequist, and S Uehara. "Radiation track, DNA damage and response—a review." In: *Reports on Progress in Physics* 79.11 (2016), p. 116601.

[207] Sébastien Incerti, Gérard Baldacchino, M Bernal, Riccardo Capra, Christophe Champion, Ziad Francis, Paul Guèye, Alfonso Mantero, Barbara Mascialino, Philippe Moretto, et al. "The geant4-dna project." In: *International Journal of Modeling, Simulation, and Scientific Computing* 1.02 (2010), pp. 157–178.

[208] Sebastien Incerti, Ioanna Kyriakou, MA Bernal, MC Bordage, Z Francis, Susanna Guatelli, V Ivanchenko, M Karamitros, N Lampe, Sang Bae Lee, et al. "Geant4-DNA example applications for track structure simulations in liquid water: a report from the Geant4-DNA Project." In: *Medical physics* 45.8 (2018), e722–e739.

[209] Ianik Plante, Artem Ponomarev, Zarana Patel, Tony Slaba, and Megumi Hada. "RITCARD: radiation-induced tracks, chromosome aberrations, repair and damage." In: *Radiation research* 192.3 (2019), pp. 282–298.

[210] Joseph Perl, Jungwook Shin, Jan Schümann, Bruce Faddegon, and Harald Paganetti. "TOPAS: an innovative proton Monte Carlo platform for research and clinical applications." In: *Medical physics* 39.11 (2012), pp. 6818–6837.

[211] Lisa Polster, Jan Schuemann, Ilaria Rinaldi, Lucas Burigo, Aimee L McNamara, Robert D Stewart, Andrea Attili, David J Carlson, Tatsuhiko Sato, José Ramos Méndez, et al. "Extension of TOPAS for the simulation of proton radiation effects considering molecular and cellular endpoints." In: *Physics in Medicine & Biology* 60.13 (2015), p. 5053.

[212] Jose Ramos-Méndez, Joseph Perl, Jan Schuemann, Jungwook Shin, Harald Paganetti, and Bruce Faddegon. "A framework for implementation of organ effect models in TOPAS with benchmarks extended to proton therapy." In: *Physics in Medicine & Biology* 60.13 (2015), p. 5037.

[213] Samuel P Ingram, Nicholas T Henthorn, John W Warmenhoven, Norman F Kirkby, Ranald I Mackay, Karen J Kirkby, and Michael J Merchant. "Hi-C implementation of genome structure for in silico models of radiation-induced DNA damage." In: *PLoS computational biology* 16.12 (2020), e1008476.

[214] Jan Schuemann, AL McNamara, JW Warmenhoven, NT Henthorn, Karen J Kirkby, Michael J Merchant, S Ingram, Harald Paganetti, Kathryn D Held, J Ramos-Mendez, et al. "A new standard DNA damage (SDD) data format." In: *Radiation research* 191.1 (2019), pp. 76–92.

[215] SP Ingram, JW Warmenhoven, NT Henthorn, EAK Smith, AL Chadwick, Neil G Burnet, Ranald I Mackay, NF Kirkby, Karen J Kirkby, and Michael J Merchant. "Mechanistic modelling supports entwined rather than exclusively competitive DNA double-strand break repair pathway." In: *Scientific Reports* 9.1 (2019), p. 6359.

[216] Yaping Qi, John William Warmenhoven, Nicholas Thomas Henthorn, Samuel Peter Ingram, Xie George Xu, Karen Joy Kirkby, and Michael John Merchant. "Mechanistic modelling of slow and fast NHEJ DNA repair pathways following radiation for Go/G1 normal tissue cells." In: *Cancers* 13.9 (2021), p. 2202.

[217] S Girst, V Hable, GA Drexler, C Greubel, C Siebenwirth, M Haum, AA Friedl, and G Dollinger. "Subdiffusion supports joining of correct ends during repair of DNA double-strand breaks." In: *Scientific reports* 3.1 (2013), p. 2511.

[218] Oliver Mortusewicz, Jean-Christophe Ame, Valerie Schreiber, and Heinrich Leonhardt. "Feedback-regulated poly (ADP-ribosyl) ation by PARP-1 is required for rapid response to DNA damage in living cells." In: *Nucleic acids research* 35.22 (2007), pp. 7665–7675.

[219] Suhas SP Rao, Miriam H Huntley, Neva C Durand, Elena K Stamenova, Ivan D Bochkov, James T Robinson, Adrian L Sanborn, Ido Machol, Arina D Omer, Eric S Lander, et al. "A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping." In: *Cell* 159.7 (2014), pp. 1665–1680.

[220] NT Henthorn, JW Warmenhoven, Marios Sotiropoulos, Adam H Aitkenhead, EAK Smith, SP Ingram, NF Kirkby, AL Chadwick, Neil G Burnet, Ranald I Mackay, et al. "Clinically rel-

evant nanodosimetric simulation of DNA damage complexity from photons and protons." In: *RSC advances* 9.12 (2019), pp. 6845–6858.

[221] Ziad Francis, Carmen Villagrasa, and Isabelle Clairand. "Simulation of DNA damage clustering after proton irradiation using an adapted DBSCAN algorithm." In: *Computer methods and programs in biomedicine* 101.3 (2011), pp. 265–270.

[222] Ben Clasie, Andrew Wroe, Hanne Kooy, Nicolas Depauw, Jay Flanz, Harald Paganetti, and Anatoly Rosenfeld. "Assessment of out-of-field absorbed dose and equivalent dose in proton fields." In: *Medical physics* 37.1 (2010), pp. 311–321.

[223] Marta Missiaggia, Giorgio Cartechini, Francesco Tommasino, Emanuele Scifoni, and Chiara La Tessa. "Investigation of in-field and out-of-field radiation quality with microdosimetry and its impact on relative biological effectiveness in proton therapy." In: *International Journal of Radiation Oncology* Biology* Physics* 115.5 (2023), pp. 1269–1282.

[224] Ismail Abdou, Guy G Poirier, Michael J Hendzel, and Michael Weinfeld. "DNA ligase III acts as a DNA strand break sensor in the cellular orchestration of DNA strand break repair." In: *Nucleic acids research* 43.2 (2015), pp. 875–892.

[225] Li Lan, Satoshi Nakajima, Yoshitsugu Oohata, Masashi Takao, Satoshi Okano, Mitsuko Masutani, Samuel H Wilson, and Akira Yasui. "In situ analysis of repair processes for oxidative DNA damage in mammalian cells." In: *Proceedings of the National Academy of Sciences* 101.38 (2004), pp. 13738–13743.

[226] Julie K Horton, Donna F Stefanick, Ming-Lang Zhao, Agnes K Janoshazi, Natalie R Gassman, Hannah J Seddon, and Samuel H Wilson. "XRCC1-mediated repair of strand breaks independent of PNKP binding." In: *DNA repair* 60 (2017), pp. 52–63.

[227] Camille Godon, Fabrice P Cordelieres, Denis Biard, Nicole Giocanti, Frederique Megnin-Chanet, Janet Hall, and Vincent Favaudon. "PARP inhibition versus PARP-1 silencing: different outcomes in terms of single-strand break repair and radiation susceptibility." In: *Nucleic acids research* 36.13 (2008), pp. 4454–4464.

[228] Julia M Reber, Jovana Božić-Petković, Michelle Lippmann, Marvin Mazzardo, Asisa Dilger, Rebecca Warmers, Alexander Bürkle, and Aswin Mangerich. "PARP1 and XRCC1 exhibit a reciprocal relationship in genotoxic stress response." In: *Cell Biology and Toxicology* 39.1 (2023), pp. 345–364.

[229] Logan Dunkenberger, Krzysztof Reiss, and Luis Del Valle. "Comet assay for the detection of single and double-strand DNA breaks." In: *Immunohistochemistry and Immunocytochemistry: Methods and Protocols* (2022), pp. 263–269.

[230] Pamela Reynolds, Sarah Cooper, Martine Lomax, and Peter O'Neill. "Disruption of PARP1 function inhibits base excision repair of a sub-set of DNA lesions." In: *Nucleic acids research* 43.8 (2015), pp. 4028–4038.

[231] Anna EO Fisher, Helfrid Hochegger, Shunichi Takeda, and Keith W Caldecott. "Poly (ADP-ribose) polymerase 1 accelerates single-strand break repair in concert with poly (ADP-ribose) glycohydrolase." In: *Molecular and cellular biology* (2007).

[232] Ana Janićijević, Kaoru Sugasawa, Yuichiro Shimizu, Fumio Hanaoka, Nils Wijgers, Miodrag Djurica, Jan HJ Hoeijmakers, and Claire Wyman. "DNA bending by the human damage recognition complex XPC–HR23B." In: *DNA repair* 2.3 (2003), pp. 325–336.

[233] Monika Sharma, Alexander V Predeus, Shayantani Mukherjee, and Michael Feig. "DNA bending propensity in the presence of base mismatches: implications for DNA repair." In: *The Journal of Physical Chemistry B* 117.20 (2013), pp. 6194–6205.

[234] Wataru Kobayashi and Hitoshi Kurumizaka. "Structural transition of the nucleosome during chromatin remodeling and transcription." In: *Current opinion in structural biology* 59 (2019), pp. 107–114.

[235] E.F. Pettersen, T.D. Goddard, C.C. Huang, G.S. Couch, D.M. Greenblatt, E.C. Meng, and T.E. Ferrin. "UCSF-Chimera, a visualization system for exploratory research and analysis." In: *J. Comput. Chem.* 25 (2004), pp. 1605–1612.

[236] Fabrizio Cleri, Marc F. Lensink, and Ralf Blossey. "DNA Aptamers Block the Receptor Binding Domain at the Spike Protein of SARS-CoV-2." In: *Frontiers Mol. Biosci.* 8 (2021), p. 713003.

[237] Matteo Cassandri, Artem Smirnov, Flavia Novelli, Consuelo Pitolli, Massimiliano Agostini, Michal Malewicz, Gerry Melino, and Giuseppe Raschellà. "Zinc-finger proteins in health and disease." In: *Cell death discovery* 3.1 (2017), pp. 1–12.

[238] John Jumper et al. "Highly accurate protein structure prediction with AlphaFold." In: *Nature* 596 (2021), pp. 583–589.

[239] Mihaly Varadi et al. "AlphaFold Protein Structure Database: massively expanding the structural coverage of protein-sequence space with high-accuracy models." In: *Nucl. Acids Res.* 50 (2022), pp. D439–D444.

[240] Séverine Duvaud, Chiara Gabella, Frédérique Lisacek, Heinz Stockinger, Vassilios Ioannidis, and Christine Durinx. "Expasy, the Swiss Bioinformatics Resource Portal, as designed by its users." In: *Nucl. Acids Res.* 49 (2021), W216–W227.

[241] Cyril Dominguez, Rolf Boelens, and Alexandre M.J.J. Bonvin. "HADDOCK: a protein-protein docking approach based on biochemical and/or biophysical information." In: *J. Am. Chem. Soc.* 125 (2006), pp. 1731–1737.

[242] G.C.P. van Zundert, J.P.G.L.M. Rodrigues, M. Trellet, C. Schmitz, P.L. Kastritis, E. Karaca, A.S.J. Melquiond, M. van Dijk, S.J. de Vries, and A.M.J.J. Bonvin. "The HADDOCK2.2 Web Server: User-Friendly Integrative Modeling of Biomolecular Complexes." In: *J. Mol. Biol.* 428 (2015), pp. 720–725.

[243] Yumeng Yan, Di Zhang, Pei Zhou, Botong Li, and Sheng-You Huang. "HDOCK: a web server for protein–protein and protein–DNA/RNA docking based on a hybrid strategy." In: *Nucl. Acids Res.* 9 (2017), p. 988996.

[244] L. A. Rodríguez-Lumbreras, B. Jiménez-García, S. Giménez-Santamarina, and J. Fernández-Recio. "pyDockDNA: A new web server for energy-based protein-DNA docking and scoring." In: *Frontiers Mol. Biosci.* 9 (2022), p. 988996.

[245] M. Macchiagodena, M. Pagliai, C. Andreini, A. Rosato, and P. Procacci. "Upgrading and Validation of the AMBER Force Field for Histidine and Cysteine Zinc(II)-Binding Residues in Sites with Four Protein Ligands." In: *J. Chem. Inf. Model.* 59 (2019), pp. 3803–3816.

[246] E. Lindahl, B. Hess, and D. van der Spoel. "ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB." In: *J. Chem. Theor. Comp.* 11 (2015), pp. 3696–3713.

[247] Ivan Ivani et al. "PARMBSC1: A REFINED FORCE-FIELD FOR DNA SIMULATIONS." In: *Nature Meth.* 13 (2016), pp. 55–58.

[248] Evgeny Krissinel and Kim Henrick. "Inference of macromolecular assemblies from crystalline state." In: *J. Mol. Biol.* 372 (2007), pp. 774–797.

[249] Daniel T. Gillespie. "A General Method for Numerically Simulating the Stochastic Time Evolution of Coupled Chemical Reactions." In: *Journal of Computational Physics* 22.4 (1976), pp. 403–434.

[250] Stefan Schuster, Thomas Dandekar, and David A. Fell. "Detection of elementary flux modes in biochemical networks: a promising tool for pathway analysis and metabolic engineering." In: *Trends in Biotechnology* 17.2 (1999), pp. 53–60.

[251] David A. Fell. *Understanding the Control of Metabolism*. London, UK: Portland Press, 1996. ISBN: 978-1855780476.

[252] Reinhart Heinrich and Stefan Schuster. *The Regulation of Cellular Systems*. New York, NY, USA: Chapman & Hall/CRC, 1997. ISBN: 978-0412989010.

[253] Uri Alon. "Design principles of biological circuits." In: *Febs J* 277 (2007), p. 11.

COLOPHON

This document was typeset using the typographical look-and-feel `classicthesis` developed by André Miede. The style was inspired by Robert Bringhurst's seminal book on typography "*The Elements of Typographic Style*". `classicthesis` is available for both LaTeX and LyX:

https://bitbucket.org/amiede/classicthesis/

Happy users of `classicthesis` usually send a real postcard to the author, a collection of postcards received so far is featured here:

http://postcards.miede.de/